Arden Handler · Joan Kennelly · Nadine Peacock
Editors

Reducing Racial/Ethnic Disparities in Reproductive and Perinatal Outcomes

The Evidence from Population-Based Interventions

Springer

Reducing Racial/Ethnic Disparities in Reproductive and Perinatal Outcomes

Arden Handler • Joan Kennelly • Nadine Peacock
Editors

Reducing Racial/Ethnic Disparities in Reproductive and Perinatal Outcomes

The Evidence from Population-Based Interventions

Springer

Editors
Arden Handler
University of Illinois, Chicago
School of Public Health
Community Health Sciences
1603 West Taylor Street
Chicago, IL, 60612-4394, USA
handler@uic.edu

Joan Kennelly
University of Illinois, Chicago
School of Public Health
Community Health Sciences
1603 West Taylor Street
Chicago, IL, 60612-4394, USA
kennelly@uic.edu

Nadine Peacock
University of Illinois, Chicago
School of Public Health
Community Health Sciences
1603 West Taylor Street
Chicago, IL, 60612-4394, USA
npeacock@uic.edu

ISBN 978-1-4419-1498-9 e-ISBN 978-1-4419-1499-6
DOI 10.1007/978-1-4419-1499-6
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2010935948

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Why a Review of the Evidence Base for Interventions to Improve Reproductive and Perinatal Health Outcomes?

Over 2 decades ago, the Institute of Medicine (IOM) released its influential report *Preventing Low Birth Weight* (IOM, 1985), galvanizing and mobilizing community, state, and federal MCH practitioners and policy-makers to improve maternal and infant health status. Despite the development of numerous programs, initiatives, and approaches to address the delivery of care during the preconceptional, prenatal, and postpartum periods, the major indicators of maternal and infant morbidity and mortality in the US have not uniformly shown marked improvement during this time (Martin, Hamilton, Sutton, et al., 2009); most notably, racial/ethnic disparities in key maternal and infant health status measures have remained persistent, and in some cases, even increased.. However, to date there has been no systematic effort to examine these interventions in a comprehensive fashion, or to specifically look at the evidence vis a vis their potential for reducing racial/ethnic disparities in reproductive and perinatal outcomes. Thus, the focus of this book.

Given that one of the major initiatives to improve reproductive and perinatal outcomes in the last 20 years has been the expansion of financial access to care, particularly during the prenatal period, a large portion of this book reviews the evidence for the public health interventions (as opposed to clinical interventions such as blood pressure checks, urinalysis, the use of risk assessment, fundal height measurement, etc.) that are incorporated into, or delivered concomitantly with prenatal care, such as depression screening and treatment, nutritional supplementation, smoking cessation programs, and prenatal case management. This book focuses on the contribution of these interventions to the overall improvement of reproductive and perinatal outcomes and their potential to reduce disparities in such outcomes between racial/ethnic groups in the United States.

We believe this book is an important undertaking, particularly since there has been an ongoing discussion of the prenatal care investment (Huntington & Connell, 1994; Fiscella, 1995; Strong, 2000). This discussion has arisen in response to the Medicaid expansions which increased the number of women with financial access to prenatal care (Kaiser Family Foundation, 2009), resulting in improved utilization, but not in associated decreases in prematurity and LBW (Martin et al., 2009). In addition, with the publication of studies showing no difference in perinatal outcomes with a reduced schedule of prenatal visits compared to a standard schedule of prenatal visits (McDuffie, Beck, Bischoff, Cross, & Orleans, 1996), and a recognition that in many Western European countries, the schedule of visits is often fewer but outcomes are better (Papiernik, 2007), it has become increasingly clear that more prenatal care (at least as measured by number of visits) in and of itself is not necessarily better.

Some of the expectation for significant positive changes in birth outcomes as the result of the Medicaid expansions was not likely justified, as many women eligible for Medicaid only due to

pregnancy do not access Medicaid and/or prenatal care early enough to allow for any potential impact (Simon & Handler, 2008). More importantly, the Medicaid expansions were not expected to have any effect on the pregnancy outcomes of the lowest income women, who were already covered by Medicaid during pregnancy, many of whom have multiple risk factors placing them at high-risk for poor birth outcomes (Guyer, 1990). Finally, beyond the numerous issues related to adequately defining and measuring prenatal care (Bell & Zimmerman, 2003; Misra & Guyer, 1998), the assumption of an independent impact of prenatal care alone on maternal and infant outcomes, disregards the current and historical context of women's lives and the established contribution of this context to reproductive health and pregnancy outcomes.

Because there is both widespread disappointment at the "failure" of the Medicaid expansions to improve pregnancy outcomes over the last 2 decades as well as widespread acknowledgement of the conceptual and measurement issues related to establishing prenatal care's effectiveness, it has been easy for some researchers and policy-makers to dismiss the relevance of increasing access and enhancing the quality of prenatal care as strategies for improving pregnancy outcomes. These circumstances provide the opportunity for us to reframe the issues pertaining to prenatal care effectiveness and advance our understanding of the contribution made by the various interventions and programs developed for women prior to, during, or soon after pregnancy, in improving their reproductive health and perinatal outcomes. A critical review of the evidence emphasizing the breadth and timing of such interventions as provided by this book, highlights the potential of a lifespan approach and creates the opportunity to consider the evidence for each of these interventions vis a vis their potential for reducing racial/ethnic disparities in reproductive and perinatal outcomes.

What's Included in This Book?

This book focuses on a systematic review of the evidence for interventions that surround a woman's childbearing years (see chapter by Kennelly for a description of methodological approaches used). It begins with a brief discussion of evidence-based medicine (EBM) and evidence-based public health (EBPH) by Handler, with a focus on the specific challenges of implementing EBPH. The principles and underlying assumptions of the scientific process to generate 'evidence' are then presented and critiqued by Aviles and Filc. Hogan, Shanahan and Rowley's chapter outlines critical and methodological issues specific to evidence generation focused on reproductive and perinatal outcomes. Subsequent chapters focus on one or more interventions to improve reproductive and/or perinatal outcomes. The chapters span the childbearing years addressing family planning and abortion, access to and use of infertility services, specific aspects of preconception care, prenatal care overall, as well as public health interventions during the prenatal period (e.g., STD and HIV screening, smoking cessation, group prenatal care, use of doulas, prenatal case management, depression screening and treatment, nutrition supplementation, and screening and treatment for substance use) that extend, enhance, and complement prenatal care. Related topics, such as genetic disease screening, and domestic violence screening and counseling during pregnancy, were originally targeted for inclusion in the book but were ultimately not able to be included.

The book also includes a chapter on intrapartum interventions prompted by the spiraling rate of C-sections and the need to examine whether certain clinical interventions which may increase or decrease maternal and infant morbidity/mortality are differentially offered to and/or used by various racial/ethnic groups. Likewise, a chapter on perinatal regionalization examines whether this system, heralded as playing a major role in reducing infant mortality in the U.S., has additional potential for reducing racial/ethnic disparities in reproductive and perinatal outcomes by focusing beyond the prenatal and perinatal periods.

What Have We Learned?

Considered together, the reviews of the evidence in this book suggest that with respect to the effectiveness of prenatal care itself, promise may lie in more integrated care models in which "enhancements" are standardized and delivered as part of comprehensive high quality care within systems that are accessible to all women, rather than as "siloed" interventions. The evidence also suggests that going beyond the prenatal period to include well-women care across the lifespan may hold significant promise and potential not only for improving reproductive and perinatal outcomes, but for reducing disparities in these outcomes as well.

More generally, the chapters in this book reveal that the depth and range of the evidence varies with respect to both the demonstrated and potential effect of each intervention to reduce racial/ethnic disparities in reproductive and perinatal outcomes. Importantly, for many interventions, information about effects on racial and ethnic disparities does not exist or can only be inferred; for the most part, the studies reviewed tend to focus on improving outcomes in one or more populations but not necessarily on approaches to reducing disparities between populations Likewise, in many cases, overall weak or modest effects might suggest potential for effectiveness but also point out difficulties related to the lack of theoretical models for how an intervention might produce an effect, inadequate or incomplete intervention implementation, lack of standardization of program models, as well as failure to move from targeted to universal implementation, thus leading to differential uptake of interventions. Additionally, several chapters caution that it is important to ensure that differential implementation of interventions (whether in quality or quantity) does not inadvertently lead to an increase in disparities, or possibly a decrease in disparities due to a worsening of outcomes for the majority population.

Despite the caveats and challenges raised by each chapter, when reviewed as an entire body of evidence for interventions to improve the reproductive health of women as well as perinatal outcomes, this book enables us to determine the "stuck points" for the field, and to identify the necessary steps for generating future evidence and improving practice to effectively address racial/ethnic disparities in reproductive and perinatal health. Importantly, this book makes clear that such an evidence-informed practice will need to recognize context and nuance, consider factors related to program/policy implementation, and appreciate the often distal relationship between public health interventions and health status outcomes. With these common understandings as the basis for action, it is our hope that this book will be a useful tool and reference for students, researchers, and practitioners alike as they pursue a wide variety of approaches to improve reproductive and perinatal health outcomes.

Chicago, IL

Arden Handler
Joan Kennelly
Nadine Peacock

References

Bell, J., & Zimmerman, F. (2003). Selection bias in prenatal care use by Medicaid recipients. *Maternal and Child Health Journal, 7*(4), 239–251.

Fiscella, K. (1995). Does prenatal care improve birth outcomes? A critical review. *Journal of Obstetrics and Gynecology, 85*, 468–479.

Guyer, B. (1990). Medicaid and prenatal care: Necessary but not sufficient. *JAMA: The Journal of the American Medical Association, 264*, 2264–2265.

Huntington, J., & Connell, F. (1994). For every dollar spent – The cost-savings argument for prenatal care. *New England Journal of Medicine, 331*, 1303–1307.

Institute of Medicine, Committee to Study the Protection of Low Birthweight, Division of Health Promotion and Disease Prevention. (1985). *Preventing low birth weight*. Washington, DC: National Academy Press.

Kaiser Family Foundation. (2009). Statehealthfacts. Income eligibility levels for pregnant women by annual income and as a percent of Federal Poverty Level (FPL). Retrieved November 6, 2009, from, http://www.statehealthfacts.kff.org/comparetable.jsp?ind=206&cat=4

Martin, J., Hamilton, B., Sutton, P., Ventura, S., Menacher, F., Kirmeyer, S. et al. (2009). Births: Final Data for 2006. *National Vital Statistics Reports, 57*(7), 1–101.

McDuffie, R., Beck, A., Bischoff, K., Cross, J., & Orleans, M. (1996). Effect of frequency of prenatal care visits on perinatal outcome among low-risk women. A randomized controlled trial. *JAMA: The Journal of the American Medical Association, 275*(11), 847–851.

Misra D.P., & Guyer B. (1998) Benefits and limitations of prenatal care: From counting visits to measuring content. *JAMA: The Journal of the American Medical Association, 279*, 1661–1662

Papiernik, E. (2007). Preventing preterm birth – is it really impossible? A comment on the IOM report on preterm birth. *Maternal and Child Health Journal, 11*, 407–10.

Simon, K., & Handler, A. (2008). Welfare reform and insurance coverage during the pregnancy period: Implications for preconception and interconception care. *Women's Health Issues, 18S*, S97–S106.

Strong, T. (2000). *Expecting trouble. The myth of prenatal care in America*. New York: New York University Press.

Acknowledgments

The idea for this book emerged from conversations between Arden Handler and Bill Tucker (Executive Editor for Health and Education at Springer) at APHA several years ago. Arden then asked her colleagues in the Maternal and Child Health Program at the University of Illinois School of Public Health, Joan Kennelly and Nadine Peacock, to join her as editors in planning and organizing the book and soliciting chapter authors. The book has endured a long gestation and it has been an incredible journey in which we, as editors, have learned a great deal through our conversations with the authors in reviewing and editing the various chapters and shaping the final product. We are very grateful to all of the authors who have contributed to this book for sharing their expertise and insights on the selected topics.

An edited book requires significant organizational skills and incredible attention to details. Such support and assistance were provided by many women whose time, energy, and efforts made this book a reality and for whom we have immense appreciation and gratitude. We give sincere thanks to Linda Factor, Anna Wiencrot, and Erica Abu-Ghallous for their meticulous proofing, formatting, and final editorial assistance with the chapters. Special thanks also to a number of our wonderful graduate students who assisted in tracking down references, checking tables, and gracefully providing assistance with any task requested of them. They include in alphabetical order, Allison Dahlke, Emily Hutter, Catherine Lind, Erin Murphy, Antoinette Price, Melissa Sherwin, Sara Wiseman, and Kendall Ziegler.

Particular thanks are also due to our editors at Springer Publisher – Bill Tucker, Khristine Queja, and Ian Marvinney – for their flexibility, support, and patience in guiding our efforts to completion. Finally, we express our appreciation to our families and friends for their encouragement and forbearance during this process.

Contents

Contributors

Luis A. Avilés
Department of Social Sciences, University of Puerto Rico, Mayagüez, PR, USA

Mary Barger
Department of Family Health Care Nursing, University of California, San Francisco, CA, USA

Amanda C. Bennett
University of Illinois, School of Public Health,
1603 W. Taylor Chicago, IL 60612, USA

Ann E.B. Borders
Division of Maternal-Fetal Medicine, Department of Obstetrics and Gynecology,
Institute for Healthcare Studies, Northwestern University Feinberg School of Medicine,
Chicago, IL, USA

Suzanne Carlberg-Racich
University of Illinois at Chicago, Midwest AIDS Training and Education Center (MATEC),
1640 W. Roosevelt Road, Suite 511, Chicago, IL 60608, USA

Noel Chávez
Department of obstetrics and Gynecology, University of Illinois, School of Public Health,
1603 W. Taylor, Chicago, IL 60612, USA

Julie Chor
Department of Obstetrics and Gynecology, John H. Stroger Jr. Hospital of Cook Country,
Chicago, IL 60612, USA

Eugene Declercq
Department of Community Health Sciences, Boston University School of Public Health,
Crosstown Center, 4th Floor, 801 Massachusetts Avenue, Boston, MA 02118, USA

Ashley Dyer
University of Illinois College of Medicine, Department of Obstetrics and Gynecology,
Chicago, IL 60612, USA

Dani Filc
Department of Political Sciences, Ben Gurion University, Tel Aviv, Israel

Sarah G. Forrestal
Mathematica Policy Research, III N. Wacker Drive,
Suite 920, Chicago, IL 60601, USA

Patricia M. Garcia
Division of Maternal-Fetal Medicine, Department of Obstetrics and Gynecology, Northwestern University Feinberg School of Medicine, Chicago, IL, USA

David C. Goodman
Center for Health Policy Research, Dartmouth Institute for Health Policy and Clinical Practice, Hanover, NH 03755, USA

Sylvia Guendelman
School of Public Health, University of California, Berkeley, CA 94720-7360, USA

Renu Gupta
University of Illinois at Chicago, Women's Mental Health Program, Department of Psychiatry, 912 S. Wood St., Chicago, IL 60612, USA

Arden Handler
University of Illinois, Chicago School of Public Health, Community Health Sciences, 1603 West Taylor Street, Chicago, IL 60612-4394, USA
handler@uic.edu

Bryna Harwood
University of Illinois College of Medicine, Department of Obstetrics and Gynecology, Chicago, IL 60612, USA

Vijaya K. Hogan
Department of Maternal and Child Health, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

L. Michele Issel
University of Illinois, School of Public Health, Community Health Sciences, 1603 W. Taylor, Chicago, IL 60612, USA

Joan Kennelly
University of Illinois, Chicago, School of Public Health, Community Health Sciences, 1603 West Taylor Street, Chicago, IL 60612-4394, USA
kennelly@uic.edu

Rieko Kishi
University of Illinois College of Nursing, Department of Women, Children and Family Health Science, 845 S Damen Avenue, Chicago, IL 60612-7350, USA

Carrie Klima
University of Illinois College of Nursing, Department of Women, Children and Family Health Science, 845 S Damen Avenue, Chicago, IL 60612-7350, USA

Michael C. Lu
Department of Community Health Sciences, David Geffen School of Medicine at UCLA, UCLA School of Public Health, Los Angeles, CA 90095-1772, USA

Ellen Mason
Division of Maternal Fetal Medicine, Department of Obstetrics and Gynecology, John H. Stroger Jr. Hospital of Cook County, Chicago, IL 60612, USA

Laura J. Miller
Women's Mental Health Division, Department of Psychiatry, Brigham and Women's Hospital, Harvard University, 75 Francis St. Boston, MA 02115, USA

Merry.-K. Moos
Maternal Fetal Medicine Division, Department of Obstetrics and Gynecology University of California, Campus Box#7516, Chapel Hill, NC 27599-7516, USA

Kathleen F. Norr
University of Illinois College of Nursing, Department of Women, Children and Family Health Science, 845 S Damen Avenue, Chicago, IL 60612-7350, USA

Nadine Peacock
University of Illinois, Chicago, School of Public Health, Community Health Sciences, 1603 West Taylor Street, Chicago, IL 60612-4394, USA

Beth F. Pellettieri
835 N. Maplewood Are. Apt. 2F, Chicago, IL 60622, USA

Diane L. Rowley
Department of Maternal and Child Health, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, 425 Rosenau Hall, 421 Pittsboro Street, Chapel Hill, NC 27599, USA

Andrea M. Scremin
Department of Psychiatry, Illinois Masonic Medical Center, 938 W. Nelson St.Chicago, IL 60657, USA

Meghan E. Shanahan
Department of Maternal and Child Health, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, 425 Rosenau Hall, 421 Pittsboro Street, Chapel Hill, NC 27599, USA

Jaime C. Slaughter
176 Honey Suckle Drive Jupiter, FL 33458, USA

Laura Stachel
School of Public Health, University of California, Berkeley, CA 94720-7360, USA

Lindsay A. Thompson
Division of General Pediatrics, University of Florida, Gainesville, FL 32608, USA

Sonya Trinh
Harvard Medical School, Brookline, MA, USA

Susan C. Vonderheid
University of Illinois College of Nursing, Department of Women, Children and Family Health Science, 845 S Damen Avenue, Chicago, IL 60612-7350, USA

Hannah Holcombe Walford
University of California San Diego, Culver City, CA 90230, USA

Judith Weiss
Department of Community Health Sciences, Boston University, School of Public Health, Crosstown Center, 3rd Floor, 801 Massachusetts Avenue, Boston, MA 02118, USA

Anna Wiencrot
Health Care Research, National Opinion Research Center, Chicago, IL 60602, USA

Richard Windsor
Department of Prevention and Community Health, School of Public Health and Health Services, George Washington U. Medical Center, 2175 K St. NW, Washington, DC 20037, USA

Whitney B. You
Division of Maternal-Fetal Medicine, Department of Obstetrics and Gynecology, Northwestern University Feinberg School of Medicine, Chicago, IL 60611, USA

Chapter 1
Introduction: Challenges in Reducing Disparities in Reproductive and Perinatal Outcomes Through Evidence-Based Public Health

Arden Handler

Introduction

Over the past decade, there has been an increased focus on the use of "evidence" to enhance practice in the delivery of health and human services. This is partially due to the ease with which data can now be accessed and turned into information (Dobrow, Goel, & Upshur, 2004). Also influential have been increases in budget deficits at all levels of government, challenging the role of the public sector as a provider of services and requiring increased justification of the use of public resources. Public health programs and interventions have come under increasing pressure to demonstrate their impact and cost-effectiveness in improving population health as reflected in major health status indicators such as the Healthy People objectives (USDHHS, 2000).

Along with increasing attention to public health performance, there is a growing awareness that to make progress in improving the health of the population, particularly to reduce intransient disparities between racial and ethnic groups, new approaches may be needed. Potential strategies may include among others, universal application of an intervention that is currently available but under-resourced, widespread endorsement and implementation of an intervention that is typically not thought of as a health intervention (e.g., social welfare, income, nutrition, or environmental strategies), and/or the development of new models for an ordinary/common intervention.

As we seek further understanding and develop new frameworks to guide our approach to reach Healthy People 2020 and beyond, it is important to take stock of our current repertoire of interventions, understand their value, carefully examine the results of relevant evaluations, and recognize that within our current body of evidence, hidden nuggets which suggest future directions for intervention may be revealed when the body of evidence is examined as a whole. This book, focused on the evidence base for public health interventions to improve reproductive and perinatal health, is written in this spirit.

Given that this is a book about the potential of *public health* approaches to reduce racial/ethnic disparities in reproductive and perinatal health outcomes, the reviews of the variety of interventions discussed within are subject to some of the unique challenges of evidence-based public health (EBPH) in contrast to those presented by evidence-based medicine (EBM). While the chapter by Aviles and Filc critically assesses the basic assumptions of scientific inquiry in generating evidence,

A. Handler (✉)
University of Illinois School of Public Health, Community Health Sciences,
1603 W. Taylor, Chicago, IL 60612, USA
e-mail: handler@uc.edu

A. Handler et al. (eds.), *Reducing Racial/Ethnic Disparities in Reproductive and Perinatal Outcomes: The Evidence from Population-Based Interventions*, DOI 10.1007/978-1-4419-1499-6_1,

the focus of this chapter, is to briefly discuss the difference between EBPH and EBM and to delineate some of the global and specific challenges that researchers and practitioners face when engaging in EBPH.

The Challenges of Evidence-Based Public Health

While there are multiple reasons for the advent of EBPH, some of which are mentioned above, the pressure for increased accountability in public health has arisen in part because of the increasing focus on the generation and use of evidence in the field of clinical medicine (Evidence-Based Medicine), public health's major partner in improving health status. Much has been written on the difference between Evidence-Based Public Heath (EBPH) and Evidence-based Medicine (EBM) (Brownson, Baker, Leet, & Gillespie, 2003; Brownson, Gurney, & Land, 1999; Dobrow et al., 2004; Heller & Page, 2002; Jenicek & Stachenko, 2003; Kohatsu, Robinson, & Torner, 2004) In short, although EBM uses a key public health science (clinical epidemiology) to produce evidence, the focus of EBM is on enhancing the ability of practitioners to engage in informed clinical decision-making at the *individual* (patient) level. EBPH, on the other hand, uses scientific principles for decision-making to improve the health of *populations*. Specifically, EBPH generates and uses evidence to evaluate existing interventions, to determine the efficiency and effectiveness of different intervention strategies, and to develop new programs and policies that have the greatest promise to improve population health.

A key distinction between EBM and EBPH centers on how evidence is created and what constitutes evidence. The randomized controlled trial (RCT), the hallmark of the evidence base for EBM, also has value in EBPH as a tool to determine the potential effectiveness of particular public health interventions and policies. However, most public health interventions address complex problems within multi-level systems and require context specific adaptations to ensure effectiveness. The conduct of an RCT to evaluate an intervention does not guarantee that there has been adequate problem identification or effective program implementation. In addition, as currently conceptualized, an RCT does not necessarily account for community and population context, essential components of EBPH practice. As such, EBPH incorporates evidence from a variety of sources and study types including RCTs. Methods for establishing and evaluating the suitability (hierarchy of evidence) of intervention and evaluation designs other than randomized controlled trials have been promulgated by the public health enterprise [e.g., the U.S. Task Force on Community Preventive Services, Guide to Community Preventive Services; (Briss et al., 2000; Zaza et al., 2000)].

Another key difference between EBM and EBPH is related to the range and types of interventions considered. If we accept the definition of public health as "what we as a society do collectively to assure the conditions in which people can be healthy" (Institute of Medicine, 2002), ensuring that the practice of public health is based on evidence is clearly an enormous undertaking, with a potentially vast range of interventions as well as outcomes to be considered.

While EBM and EBPH may focus on the same health status outcomes (e.g., injury, cancer, low birthweight), EBM typically considers the most proximal causes, and evaluates individual level interventions and treatments. On the other hand, EBPH recognizes and indeed emphasizes the multi-factorial etiology of almost all health status outcomes, and examines the effects of population level practices, programs, and policies on such. Likewise, public health interventions are usually designed to modify or ameliorate risk factors and their associated complex pathways often significantly upstream from a health status outcome. Examples of more upstream risk factors include knowledge and attitudes about a health behavior, the availability of substances such as tobacco and alcohol, whether or not one lives in a low or high income family or community, the presence of supermarkets or parks in a neighborhood, and the extent of residential segregation or

racism. Addressing such risk factors, in other words, the more distal or upstream "causes", makes establishing a causal link between a particular public health intervention and one or more health status outcomes a particularly difficult endeavor.

In addition to the challenge of the distal nature of the relationship between public health interventions and health status outcomes, the selection of outcome measures for public health interventions is often not guided by explicit theoretical frameworks or a complete understanding of the chain of "events" that lead to the anticipated effects in intended populations. Notably, even when there is an understanding of the underlying causal chain, public health professionals are often forced to measure effectiveness according to outcome measures selected by funders (e.g., reduce infant mortality by 50% in 5 years) or other external parties, rather than being able to select the most potentially sensitive structure, process and outcome measures.

Another challenge in EBPH relates to public health's emphasis on primary prevention. Proving that an intervention has prevented an outcome from occurring (e.g., rates of unplanned pregnancy) in a community is much more difficult than showing that a new medication or treatment led to a cure in an individual (e.g., a child's cancer). Further complicating EBPH, the path to "health" (health status) can be affected by a variety of characteristics of the population, health system, or the broader physical, social, economic or political environment (Victora, Habicht, & Byrce, 2004). Thus, given that the implementation of a successful population-based intervention likely varies from community to community as noted above, testing and evaluating public health interventions in any one community rarely provides a definitive answer or solution to a prevalent public health problem.

Even when there is sufficient evidence in support of an intervention in one or more populations or settings, there is not always the political will to fully implement the intervention or to commit to implementation in ways that allows tailoring to the needs of unique and diverse populations. A consequence of the latter phenomenon is that widespread or universal introduction of an intervention may inadvertently lead to continued racial/ethnic disparities in outcomes. This may occur if there is differential implementation, uptake and/or effectiveness of the intervention in various communities or populations. Because the intervention may be understood by the public health community to be effective (e.g., Back to Sleep Campaign for SIDS; education on signs and symptoms of preterm labor for prematurity), there may be insistence that the intervention be implemented as originally delivered in research studies, without the tailoring and nuance needed for adaptation within specific and varied cultural contexts, thus precluding full (or even minimal) effectiveness.

EBPH is also hindered by the lack of adequate surveillance and quality data systems to provide suitable performance measurement and ongoing population-based outcome information. "Evidence" for EBPH practice requires timely, relevant, and appropriately analyzed data generated from population interventions (Brownson et al., 1999). However, in the US, there is no commitment to the generation and maintenance of high quality data systems, evidenced by our currently underfunded and struggling vital statistics system, insufficient resources committed to our national health surveys, and inadequate support for institutions like the U.S. Census Bureau.

All of these issues and caveats plaguing public health science and thus EBPH, create a situation in which an evaluation of a population or community intervention frequently yields a finding of "no or minimal effect" with respect to improving or reducing disparities in a health status outcome. However, interpretation of such a result in EBPH is a delicate undertaking. While a review of the evidence demonstrating "no or minimal effect" might provide support for a disinvestment in a particular intervention or policy, those closely involved and familiar with the intervention delivery may be reluctant to endorse this action, emphasizing that the intervention makes sense "on the face of it" (e.g., nutrition support, depression screening, smoking cessation services). Likewise, the intervention may be supported because it is consistent with the social justice roots/philosophy of public health which recognizes that there are many basic services to which all populations should have access (e.g., STI screening and treatment, family planning, nutrition services, prenatal care). On the

other hand, if the populations who are the recipients of the program/policy might be better served by new models or strategies, complacency or commitment to a particular intervention regardless of the evidence, gives preference to the status quo and may further compromise the health of some populations.

Also contributing to the delicate balance of weighing the evidence in public health, is the fact that many public health interventions are aimed at low-income communities. Such interventions may provide a substantial portion of the infrastructure and other resources available to address a particular health problem or related health and social issues in any one community. In such instances, making an EBPH decision must be contextualized to consider not just individual program recipients but the community as an entire unit. Finding "no effect" in the reduction of infant mortality of a case-management program, for example, might lead one to argue for the termination of this program. However, the program might improve overall maternal well-being or bring as yet unmeasured benefits to the community such as the provision of jobs for lay health workers, or the development of a community advisory council that has become involved in health issues as a result of the intervention. As this example demonstrates, explicit, multilevel theoretical frameworks are necessary to guide the development and implementation of evidence-based interventions as well as the delineation of structure, process and outcome measures to assess their effectiveness.

Clearly, tension exists in EBPH practice between the ongoing funding of programs which while not "proven" to be effective, bring additional needed resources and secondary benefits to high-risk communities and the potential termination of such programs due to inability to demonstrate a discernible impact of a particular intervention approach. Importantly, as public health practice increasingly relies on 'evidence and best practice' in the development and implementation of the best program models available to communities, ensuring that resources continue to flow to at-risk communities (if an intervention is terminated) is imperative.

Similarly, balancing fidelity to the social justice roots of public health and being responsive to the evidence base has implications for the focal points and process of generating evidence. While it is important for scientific inquiry on the effectiveness and efficacy of certain interventions to continue, particularly those which have become widespread without the development of a solid evidence base for a particular outcome of concern (e.g., prenatal care and preterm birth), it is equally, if not more important to also ensure the generation of evidence related to the accessibility, quality, and acceptability of such interventions (e.g., ensuring access to and utilization of high quality prenatal care).

Conclusion

Given the challenges and dilemmas that are part and parcel of EBPH as described above, it is essential that researchers and practitioners evaluate and make the most of the evidence for particular public health interventions recognizing the following caveats:

1. "No effect" is usually not a clear-cut outcome and may have multiple interpretations and implications.
2. Interventions may be effective, but not for the measures that have been selected as the focus of evaluation; therefore, continued implementation may be justified when considered vis a vis an alternative set of measures (structure, process or outcome) that are more sensitive to the intervention.
3. Even when there is strong evidence for the effectiveness of an intervention, there is not always sufficient will or resources to support its implementation/expansion/dissemination.
4. Strong evidence against the effectiveness of the intervention may not always lead to revision or termination; this may have both positive and negative effects for the affected populations.

5. When an intervention appears to be effective, widespread dissemination may not uniformly improve health or decrease racial/ethnic disparities, if access to or uptake of the intervention is differential across population groups.
6. When access to an intervention is a matter of ensuring equity between populations, the generation of evidence may need to increase its focus on quality improvement or implementation strategies, rather than continue to focus only on the effectiveness of the intervention as currently delivered.

Acknowledging these caveats does not preclude making the best possible decisions given the current state of knowledge about any particular public health intervention. However, to maximize the ability of public health practice to improve health outcomes, future efforts to develop and implement public health interventions based on the evidence must synergistically consider the evidence as well as the context of both evidence generation and implementation.

References

Briss, P. A., Zaza, S., Pappaioanou, M., Fielding, J., Wright-De Agüero, L., Truman, B. I., et al. (2000). The Task Force on Community Preventative Services. Developing an evidence-based Guide to Community Preventive Service-Methods. *American Journal of Preventative Medicine*, *18*(1 Suppl.), 35–43.

Brownson, R., Baker, E., Leet, T., & Gillespie, K. (Eds.). (2003). *Evidence-based public health*. Oxford, New York: Oxford University Press.

Brownson, R., Gurney, J., & Land, G. (1999). Evidence-based decision-making in public health. *Journal of Public Health Management Practice*, *5*(5), 86–97.

Dobrow, M., Goel, V., & Upshur, R. E. G. (2004). Evidence-based health policy: Context and utilization. *Social Science and Medicine*, *58*, 207–217.

Heller, R., & Page, J. (2002). A population perspective to evidence based medicine: Evidence for population health. *Journal of Epidemiology and Community Health*, *56*(1), 45–47.

Institute of Medicine, Committee on the Assuring the Health of the Public in the 21st Century. (2002). *The future of the public's health in the 21st century. Board on Health Promotion and Disease Prevention*. Washington, DC: National Academics Press.

Jenicek, M., & Stachenko, S. (2003). Evidence-based public health, community medicine, preventive care. *Medical Science Monitor*, *9*(2), SR1–SR7.

Kohatsu, N., Robinson, J., & Torner, J. (2004). Evidence-based public health: An evolving concept. *American Journal of Preventive Medicine*, *27*(5), 417–421.

U.S. Department of Health and Human Services. (2000). *Healthy people 2010: Understanding and improving health* (2nd ed.). Washington, DC: U.S. Government Printing Office.

Victora, C., Habicht, J. P., & Byrce, J. (2004). Evidence-based public health: Moving beyond randomized trials. *American Journal of Public Health*, *94*(3), 400–405.

Zaza, S., Wright-de Agüero, L., Briss, P. A., Truman, B. I., Hopkins, D. P., Hennessy, M. H., et al. (2000). Data collection instrument and procedure for systematic reviews in the Guide to community Preventative Services. *American Journal of Preventative Medicine*, *18*(1 Suppl.), 44–74.

Chapter 2
Methodological Approach to Assessing the Evidence

Joan Kennelly

The approach taken in this book to guide authors in assessing the evidence for their respective topic areas was generated by the editors. It represents a combination of current recommendations for describing the state of public health evidence, assessing the quality of that evidence, including the suitability of the various studies reviewed to assess the effectiveness of their respective interventions, along with a good dose of practicality.

It was beyond the scope of this book to conduct meta-analyses or full systematic reviews of the literature on the various topics. On the other hand, it was the intent of the editors and authors to provide a thorough and comprehensive review of the literature on select interventions designed to promote reproductive and perinatal health and to identify the role of the interventions with respect to reducing racial and ethnic disparities in related outcomes. Through this review, we expected to further our collective understanding of the strength of the evidence base for the common interventions examined and their associated outcomes, as well as the underlying assumptions of such interventions and their potential for decreasing relevant population health disparities.

Although the complexity of public health interventions is well recognized, the difficulty in assessing and evaluating the impact of population based interventions is often underappreciated and misunderstood. Public health's focus on diverse populations in real life settings presents a significant set of challenges for evaluating and assessing impact. Understanding the effect of context on the design of interventions, their implementation and potential impacts, is central for an adequate and meaningful consideration of evidence for effectiveness. Unfortunately, fundamental information on the quality of interventions as well as critical details on the value and potential replication of such, are not usually included in most systematic reviews or evaluations of public health activities and programs.

Therefore, the guidance to authors and tools for reviewing the evidence that were developed by the editors for this book attempted to address some of these limitations (Appendix A). Specifically, authors were asked to focus on a particular intervention that has been assumed to have a positive influence on reproductive and perinatal outcomes, and to provide an overview of the theoretical and scientific basis of the intervention.

Authors were directed to include a spectrum of study designs including randomized control trials, observational studies, quasi-experimental designs, and expert reports, including both quantitative and qualitative methodologies and to summarize the reviewed studies in both tabular and narrative form. For each study, authors were asked not only to delineate the study

J. Kennelly (✉)
University of Illinois School of Public Health, Community Health Sciences,
1603 W. Taylor, Chicago, IL 60612, USA
e-mail: Kennelly@uic.edu

A. Handler et al. (eds.), *Reducing Racial/Ethnic Disparities in Reproductive and Perinatal Outcomes: The Evidence from Population-Based Interventions*, DOI 10.1007/978-1-4419-1499-6_2,

type and provide a description of the intervention and key findings, but to also specify the characteristics of the population studied and to list major caveats or biases that may influence the outcomes or interpretation of the study's findings, including identifiable contexts within which the intervention was designed and implemented. This information was to be included in a table which focused on the evidence for the effectiveness of the intervention with respect to major reproductive or perinatal outcomes selected by the chapter authors (see Table 2.1 template below).

Note that column eight asks for information about caveats and biases. In addition to the common use of the term caveat, some authors also used this column to provide explanations and modifying details to prevent misinterpretation and promote a more accurate understanding of the study being reviewed.

Furthermore, in an attempt to standardize the review of study quality across the variety of interventions and study designs, authors were initially asked to complete a quality checklist covering the following domains: reporting, external validity, internal validity (bias and confounding), and power. The checklist was an adaptation of the Methodological Quality Checklist developed by Downs and Black in 1998, to accommodate approaches used in most population based evaluations as opposed to clinical research. (Downs & Black, 1998) It became obvious that this checklist was not adequate for the qualitative studies that a number of authors were including in their reviews. Thus, an additional checklist was developed by the editors to provide consistency in the evaluation of study quality and evidence for qualitative studies. This checklist included specific questions related to the study's research design, sampling, data collection, data analysis, results, as well as research value, and was adapted from existing work (Beck 1993; CASP 2002; Rychetnik & Frommer 2002; Miles & Huberman, 2002). The checklists are included in Appendix B.

Importantly, while each study reviewed by authors was given a "total quality score," categorized as good, fair and poor, each study was also rated in terms of its respective "suitability." For quantitative studies, suitability related to the study's capacity to assess the effectiveness of the particular intervention, and was classified as greatest, moderate or least. This rating (Appendix A) was adopted from the Guide to Community Preventive Services (Briss et al. 1999.) Suitability of qualitative studies (Appendix B2) referred to the study's capacity to generate knowledge, facilitate interpretation of quantitative studies, as well as illuminate factors relevant to intervention's effectiveness. Studies were designated as having high, fair, or low value. This rating was adopted from previous work (Beck, 1993; Critical Appraisal Skills Program (CASP), 2002; Miles & Huberman, 2002; Rychetnik & Frommer, 2002). Authors were asked to tabulate the information from the quality checklists and suitability assessments (see Table 2.2 template below).

Table 2.1 Major outcomes associated with studies of x intervention

Health status outcome No.1								
Author, Year	Study design	Study type	Description of intervention what, how and where	Populations studied (ages included, race and ethnicity) and Sample size	Address disparities (Yes/No)	Key findings related to intervention effectiveness (OR with CI or p values reflecting the intervention-outcome relationship	Caveats/ Biases	Findings support the intervention? Yes/No For which populations?

Table 2.2 Quality rating of studies associated with x intervention

Health status outcome No. 1							
Author, Year	Reporting	External validity	Internal validity-bias	Internal validity-confounding	Power	Total quality score <14 = poor 15–19 = fair >20 = good	Suitability of study to assess effectiveness

Table 2.3 Meta-analysis table: topic area

Source	Number of studies/*N*/(% receiving xx intervention)	Findings	Contextual factors	Disparities/Comments

In addition to individual studies, a number of the chapters also include reviews of meta-analyses and other systematic literature reviews. The importance of contextual factors that might influence the quality, strength, and external validity of the meta-analyses was noted by one of our book's chapter authors, Mary Barger. Thus, a third table template developed by Dr. Barger was included for authors' use in tabulating the findings of such inquiry and to facilitate discussion in the chapter narratives. However, not every meta-analysis discussed in the chapter narratives was included in such tables.

While there is no summary score for the totality of studies reviewed in relation to a particular intervention, authors were asked to provide a narrative summary of the evidence and the potential role of the intervention to reduce racial and ethnic disparities in reproductive and perinatal outcomes. In discussing the evidence summary, authors were specifically asked to address demonstrated effects as well as context and any variability in implementation of the intervention, along with the relevance of the evidence for public health practitioners. Finally, in the absence of any quantified effects or impact, authors were encouraged to speculate on reasons why the interventions continue to hold favor in public health practice.

Although efforts to standardize a quality review and discussion of the literature across the book chapters were agreed upon and embraced by authors, the actual process of reviewing the literature across the various topics did not always lend itself to such standardization. The range of intervention topics had their own set of exceptions in terms of the types of interventions and practice that were being considered, as well as the relevant studies and evaluations that had been carried out. There was also considerable variation in the availability of the desired information from the primary studies. This affected the extent to which some authors were able to address the issue of reducing racial and ethnic disparities for a particular intervention, as well as speculate on the relevance of the study findings for specific population groups or the feasibility of their replication. In addition to author preferences and prerogative, this variability is reflected in the type and number of tables included and their placement in the chapter, as well as in each chapter's narrative discussion.

Even though each chapter is distinctive, the uniqueness of several chapters is worth noting in terms of their departure from the proposed chapter structure. Specifically, the chapters on

childbirth practices, clinical interventions for preterm delivery, and screening and treatment of sexually transmitted infections and HIV, because of their focus on clinical guidelines and medical practice based on individual risk, posed challenges in terms of assessing and summarizing their relevance to population-based approaches to reducing disparities in reproductive outcomes. The chapter on family planning reviewed the evidence base for intervention strategies designed to increase access to family planning and safe abortion services (rather than reviewing the effectiveness of family planning services themselves, which is already well-established). Given the unique character of the evidence evaluated, results of this review were summarized in tables but not subjected to quality ratings. Another unique feature of some of the chapters in this book relates to those interventions (e.g., infertility treatments) which if made more available and accessible to women might potentially increase disparities in reproductive outcomes. Although the book editors were involved in extensive editing, each chapter ultimately reflects the perspective of the chapter author(s).

Overall, the chapters in this book highlight the dynamic relationship between politics and science and how social values are embedded in the scientific process of inquiry as well as in the application of "scientific" findings. Each chapter forces us to ask how and why it is that public health and medicine sometimes persist in pursuing practices and approaches that are in contradiction to solid evidence, or fail to universally adopt practices for which there is good evidence. The following chapters by Handler, and Aviles and Filc, highlight potential causes of these sometimes disconcerting approaches and the particular challenges of evidence-based public health.

Appendix A: Detailed Instructions to Substantive Chapter Authors

1) *Each chapter is expected to be no more than 25–30 pages double-spaced including the tables.* Authors will focus on a specific intervention that has been assumed to make a positive contribution to enhancing reproductive and perinatal health outcomes and examine the underlying theories and scientific basis of these assumptions. Chapters should address the following:

   - *Definition of the intervention*: Describe the selected intervention and provide a brief overview of its *theoretical or scientific* basis. Include a brief history and describe the current role of the intervention with respect to reducing racial/ethnic disparities in key reproductive/perinatal outcomes. If the studies to date have not focused on racial/ethnic disparities, state this.
   - *Outcomes affected by the intervention*: Provide a brief overview of the outcomes assumed to be affected by the intervention. Select no more than *two outcomes* which will be the focus of your review of the evidence. Typically, *these outcomes should be those considered to be the "main" outcomes related to the intervention.* However, if there has been a major review of the evidence of the intervention vis a vis a particular outcome, you might want to briefly summarize the findings of that review and provide readers with information about how to access that review. Then choose one of the "lesser" *outcomes* as one of your two outcomes for your review. *For each outcome chosen, very briefly describe the overall prevalence and trends over time for the major ethnic/racial disparities. Keep this brief as this information is likely to appear in more than one chapter.*

2) *Review of the evidence*

A. *Overall instructions*

Authors are requested to select research studies completed since *1985 or the last major review, if this is later*. To ensure consistency between chapters, we ask that authors use the following search engines: *MEDLINE, CINAHL (Cumulative Index to Nursing and Allied Health Literature), Popline, WHO Reproductive Health Library, Web of Science, Cochrane Library, OCLC First Search and Academic Search Elite.* It is assumed that all authors will have access to the proposed search engines through their institutional affiliations. Some engines might require access through a university's library portal. If problems arise in freely accessing any of the engines, please consult with your university librarian and advise the editors.

1. *Study designs for consideration include*: randomized controlled trials, observational studies (cohort and case-control, ecologic epidemiology studies, quasi-experimental designs including time series analyses), studies that have integrated qualitative and quantitative methods (if not already included in above), and expert reports. If a meta-analysis has been done, authors should include the results of the meta-analysis in the list of studies. Authors are requested to follow the paradigm for classifying study designs and *determining the suitability of a study design for assessing effectiveness* as presented in "Developing an Evidence-Based *Guide to Community Preventive Services* – Methods," by Briss et al. The paradigm figure and suitability table are included below.

   Given the hierarchy of study designs determining suitability for assessing intervention effectiveness, and to reduce author burden, it might be best to select studies hierarchically, with a focus on the methodologically strongest studies. However, if you find a series of weaker studies that tend to support the same conclusion, you will want to include these as well. In general, where there is an overwhelming amount of evidence, focus on the strongest evidence and comment on the amount of evidence available.

   Because the focus of the book is on reducing racial/ethnic disparities, authors should if possible select studies conducted within racial/ethnic minority groups or those that directly compare the outcomes of an intervention for one or more racial/ethnic minority groups with the outcomes for European-Americans/majority culture. If a study directly addresses disparities, to the extent possible, please describe how "disparity" was defined and what determinants of disparity were included in the study. If none of the studies for this intervention are focused on racial/ethnic disparities per se, you should review the evidence at hand, and provide your own insights with respect to the potential effectiveness of the intervention for reducing racial/ethnic disparities.

   Studies need not be limited to the U.S; however, for the most part studies are expected to be derived from the developed world. We are still considering devoting a separate chapter to the effectiveness of developing world interventions introduced in multiple locales in improving reproductive/perinatal outcomes.

B. *Specific Approach for Identified Studies: Reviewed studies are to be summarized in both tabular (see mock Tables 2.1 and 2.2 below) and narrative format.*

1. **Table 2.1**

   For each study related to each selected health status outcome, delineate the study design according to the algorithm and identify the study type. Study type refers to where the

findings and evidence were found, such as in a published article, technical report, abstract presentation, book or book chapter, unpublished manuscript, dissertation or thesis. Provide a description of the intervention (what was done, how, and where), denote the populations studied (ages, racial and ethnic categories included) and the sample size. Summarize key findings related to intervention effectiveness, list major caveats/biases, and note whether the study supports the effectiveness of the intervention and for which populations, if known.

2. **Table 2.2**

   For each study, complete a set of questions (approximately 25–30) based on the Quality Checklist for RCTs and Observational Studies of Treatment Studies (used in the AHRQ study of perinatal depression and in turn, based on the Methodological Quality checklist developed by Downs & Black, 1998). This checklist (included in Appendix B) has several domains: reporting, external validity, internal validity (bias), internal validity (confounding), and power. Each domain generates a score; the scores are then summed for a total quality score. In the proposed checklist (slightly revised by the editors to accommodate approaches used in most population based evaluations as opposed to clinical research studies) scores greater than or equal to 20 are considered good studies, scores between 15 and 19 are considered fair, and scores of 14 and below are considered poor. Report the scores for each study in Table 2.2. *For meta-analyses, leave columns 3–9 blank*.

   In Column 9, indicate the suitability of each study's design for assessing intervention effectiveness. As noted above, this classification is taken from the *Guide to Community Preventive Services*. Table 2.2 will help authors in preparing a narrative summary of the evidence.

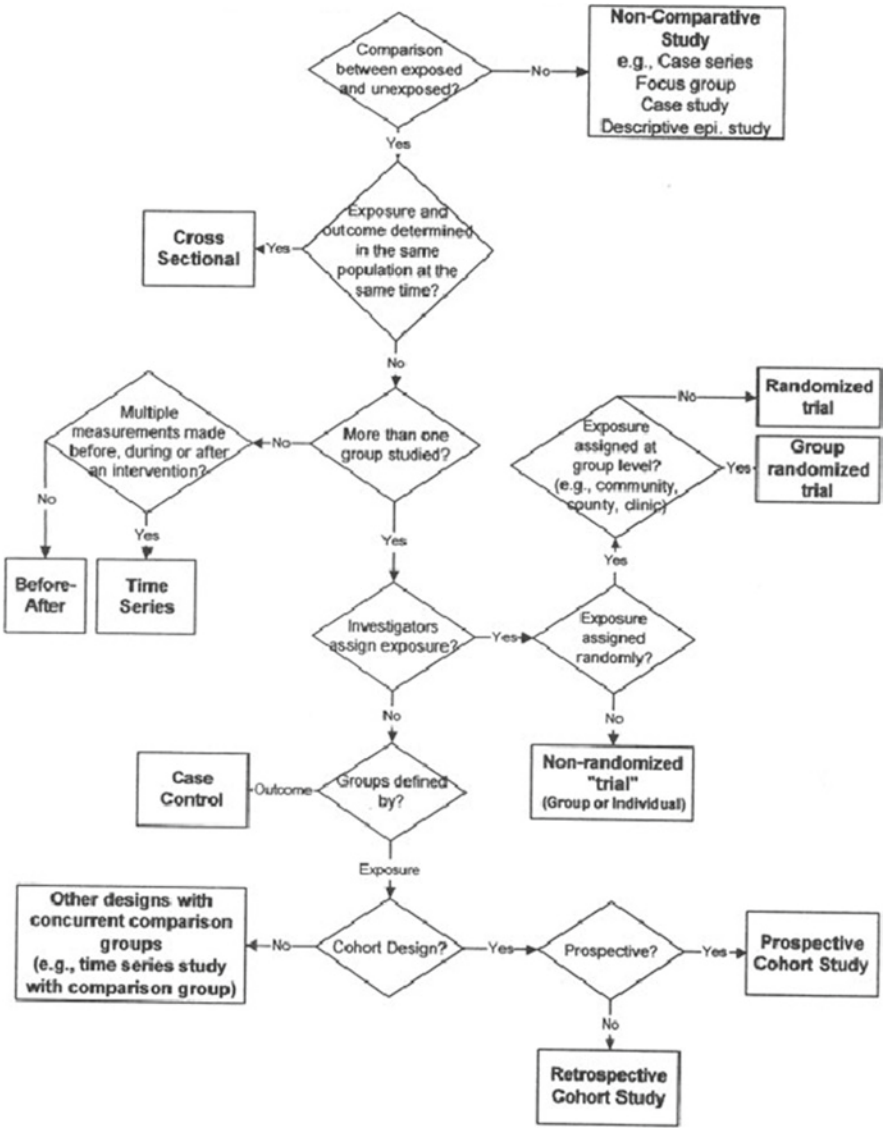
3) *Summary of the evidence and role or potential role of the intervention in reducing racial/ethnic disparities in repro/perinatal outcomes.*

Informed by the study designs, their suitability and quality, as well as the underlying theory and appropriateness of the intervention for the desired outcome, authors should use their judgment to describe and evaluate the overall state of the evidence reported. To the extent possible, authors should address: What are the demonstrated effects of the interventions with respect to reducing racial/ethnic disparities in reproductive/perinatal outcomes? Was there a great deal of variability in the implementation of the intervention? In the absence of any demonstrated effects, what might be reasons why these interventions continue to demand support and favor in public health practice? If positive effects of the intervention have been demonstrated but these effects have not been specific to reducing racial/ethnic disparities, consider the potential of this intervention for reducing racial/ethnic disparities. In doing so, be sure to consider whether (in your judgment), just simply "applying the evidence" to more populations will result in a reduction of racial or ethnic disparities, or whether other actions might need to be taken.

4) *Relevance of evidence for practitioners:*

Each chapter should provide commentary on whether the evidence to date has been well-translated into public health practice (e.g., how widespread is the intervention? where has it been implemented?). To the extent possible, discuss barriers, challenges, and solutions to translating the evidence into MCH public health practice. What can practitioners do to implement the evidence? What system/policy changes might be necessary to disseminate the evidence and to encourage its implementation?

Study Design Algorithm and Suitability Guidelines



Suitability of Study Design for Assessing Effectiveness in the Guide to Community Preventive Services

Suitability	Attributes
Greatest	Includes designs with concurrent comparison groups *and* prospective measurement of exposure and outcome
Moderate	Includes all retrospective designs *or* multiple pre or post measurement designs with no concurrent comparison group
Least	Includes single pre and post measurement designs and no concurrent comparison group designs *or* exposure and outcome measured in a single group at the same point in time

Appendix B: Quality Checklists

B1. Quality Checklist for RCTs and Observational Studies

(used in the AHRQ study of perinatal depression and based on a Methodological Quality checklist developed by Downs & Black, 1998).

Reviewer's initials _____
First Author _____ Journal: _____
Year published_____

	Reporting		Yes	No	U/D
1.	Is the hypothesis/aim/objective of the study clearly described?		1	0	0
2.	Is the underlying theory described?		1	0	0
3.	Are the main outcomes to be measured clearly described in the Introduction or Methods section?		1	0	0
4.	Are the characteristics of the study population included in the study clearly described?		1	0	0
5.	Are the interventions under study clearly described?		1	0	0
6.	Was exposure to the intervention measured?		1	0	0
		Yes	P*	No	U/D
7.	Are the distributions of principal confounders in each group of study participants to be compared clearly described?	2	1	0	0
			Yes	No	U/D
8.	Are the main findings of the study clearly described?		1	0	0
9.	Does the study provide estimates of the random variability (e.g., standard error, standard deviation, confidence intervals, inter-quartile range) in the data for the main outcomes?		1	0	0
10.	Have all important adverse events/negative outcomes that may be a consequence of the intervention been reported?		1	0	0
11.	Have the characteristics of study participants lost to follow up been described?		1	0	0
12.	Have actual probability values been reported (e.g., 0.035 rather than <0.05) for the main outcomes except where the probability value is less than 0.001?		1	0	0
	Total reporting score:_____				

**P* partially; *U/D* unable to determine

	External validity	Yes	No	U/D
13.	Were the study participants asked to participate representative of the entire population from which they were recruited?	1	0	0
14.	Were study participants who agreed to participate representative of the entire population from which they were recruited?	1	0	0
15.	Were the staff, places, and facilities where the study participants received the intervention representative of the intervention the majority of subjects receive?	1	0	0

External validity	Yes	No	U/D
16. Were the screening criteria for study eligibility specified?	1	0	0
Total external validity score:_____			
Internal validity – bias	**Yes**	**No**	**U/D**
Answer this 17 and 18 only if this was a randomized controlled trial:			
17. Was an attempt made to blind study participants to the intervention they received?	1	0	0
18. Was an attempt made to blind those measuring the main outcomes of the intervention?	1	0	0
Answer alternative 17 and 18 if this was not a randomized controlled trial:			
19. Were appropriate methods used to adjust for the differences between groups with and without the intervention (to control for selection bias)?	1	0	0
20. Were appropriate methods used to account for any biases related to differential ascertainment of the outcome in groups with or without the intervention?	1	0	0
21. If any of the results of the study were based on "data dredging," was this made clear?	1	0	0
22. In trials and cohort studies, do the analyses adjust for different lengths of follow-up of study participants, or in case-control studies, is the time period between the intervention and outcome the same for cases and controls?	1	0	0
23. Were the statistical tests used to assess the main outcomes appropriate?	1	0	0
24. Was compliance with the intervention reliable?	1	0	0
25. Were the main outcome measures used accurate (valid and reliable)?	1	0	0
Total bias score:_____			

**P* partially; *U/D* unable to determine

Internal validity – confounding	Yes	No	U/D
26. Were the study participants in the different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited from the same population?	1	0	0
27. Were study participants in the different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited over the same period of time?	1	0	0
28. Were study participants randomized to intervention groups?	1	0	0
29. *Answer this Q.27, if randomization occurred:* was the randomized intervention assignment concealed from both study participants and intervention staff until recruitment was complete and irrecoverable?	1	0	0
30. *Answer this Q.27, if randomization did not occur:* were study participants in the research or evaluation, unaware of the study hypotheses?	1	0	0
31. Was there adequate adjustment for confounding in the analyses from which the main findings were drawn?	1	0	0
32. Were losses of study participants to follow-up taken into account?	1	0	0
Total confounding score:_____			

Power

33. Did the study mention having conducted a power analysis to determine the sample size needed to detect a significant difference in effect size for one or more outcome measures?

No	0
Yes, one measure	1
Yes, two or more measures	2
Total Power Score	

Total quality score: _____
(sum of all domain scores)

**P* partially; *U/D* unable to determine

Instructions for select questions for the quality checklist for RCTs and observational studies

1. If the authors describe the formative research, theoretical basis(es) or constructs upon which the intervention was developed the question should be answered yes.
2. If the main outcomes are first mentioned in the Results section, the question should be answered no.
3. In cohort studies and trials, inclusion and/or exclusion criteria should be given. In case control studies, a case-definition and the source for controls should be given.
4. Interventions and placebo (where relevant) that are to be compared should be clearly described.
5. Give one point if some confounders are described and two only if most of these principal confounders are described.
6. Simple outcome data (including denominators and numerators) should be reported for all major findings so that the reader can check the major analyses and conclusions. (This question does not cover statistical tests that are considered below).
7. In non-normally distributed data the inter-quartile range of results should be reported. In normally distributed data the standard error, standard deviation or confidence intervals should be reported. If the distribution of the data is not described, it must be assumed that the estimates used were appropriate and the question should be answered yes.
8. This should be answered yes if the study demonstrates that there was a comprehensive attempt to measure adverse events/negative outcomes of the intervention.
9. This should be answered yes where there were no losses to follow-up or where losses to follow-up were so small that findings would be unaffected by their inclusion. This should be answered no where a study does not report the number of patients lost to follow-up.
10. The study must identify the source population for study participants and describe how the study participants were selected. Study participants would be representative if they comprised the entire source population, an unselected sample of consecutive participants, or a random sample. Random sampling is only feasible where a list of all members of the relevant population exists. Where a study does not report the proportion of the source population from which the study participants are derived, the question should be answered as unable to determine.
11. The proportion of those asked who agreed should be stated. Validation that the sample was representative would include demonstrating that the distribution of the main confounding factors was the same in the study sample and the source population.
12. For the question to be answered yes, the study should demonstrate that the intervention was representative of that in use in the source population. The question should be answered no if, for example, the intervention was undertaken in a clinically located site in which only subjects participating in clinical care might have participated in the intervention. *For randomized studies* where the subjects would have no way of knowing which intervention they received, this should be answered yes.
13. *For randomized studies* where the researchers would have no way of knowing which intervention subjects received, this should be answered yes.
14. *For non-randomized studies*, if methods were used to adjust for initial differences between groups, the answer should be yes.

15. *For non-randomized studies*, if the same methods were used for ascertainment of the outcome in both groups, the answer should be yes.
16. Any analyses that had not been planned at the outset of the study should be clearly indicated. If no retrospective unplanned subgroup analyses were reported, then answer yes.
17. Where follow-up was the same for all study subjects the answer should be yes. If different lengths of follow-up were adjusted for by, for example, survival analysis the answer should be yes. Studies where differences in follow-up are ignored should be answered no.
18. The statistical techniques used must be appropriate to the data. For example, nonparametric methods should be used for small sample sizes. Where little statistical analysis has been undertaken but where there is no evidence of bias, the question should be answered yes. If the distribution of the data (normal or not) is not described it must be assumed that the estimates used were appropriate and the question should be answered yes.
19. Where there was non-compliance with the allocated treatment or where there was contamination of one group, the question should be answered no. For studies where the effect of any misclassification was likely to bias any association to the null, the question should be answered yes.
20. For studies where the outcome measures are clearly described, the question should be answered yes. For studies which refer to other work or that demonstrates the outcome measures are accurate, the question should be answered as yes.
21. For example, subjects for all comparison groups should be selected from the same population. The question should be answered unable to determine for cohort and case control studies where there is no information concerning the source of subjects s included in the study.
22. For a study which does not specify the time period over which subjects were recruited, the question should be answered as unable to determine.
23. Studies which state that subjects were randomized should be answered as yes except where method of randomization would not ensure random allocation. For example, alternate allocation would score no because it is predictable.
24. *If randomization occurred*, and assignment was concealed from subjects but not from staff, it should be answered no.
25. *If randomization did not occur* and if methods used ensure that those in the intervention group and those in the comparison group were unaware of the study hypotheses, then the answer should be yes.
26. This question should be answered no for trials if: the main conclusions of the study were based on analyses of treatment rather than intention to treat; the distribution of known confounders in the different treatment groups was not described; or the distribution of known confounders differed between the treatment groups but was not taken into account in the analyses. In non-randomized studies if the effect of the main confounders was not investigated or confounding was demonstrated but no adjustment was made in the final analyses the question should be answered as no.
27. If the numbers of patients lost to follow-up are not reported, the question should be answered as unable to determine. If the proportion lost to follow-up was too small to affect the main findings, the question should be answered yes.

Source: Based on a modified version of the form from Downs & Black (1998)

B2. Guidelines to Evaluate the "Quality and Evidence" of Qualitative Studies

The proposed questions consider study design, study quality and consistency and address issues related to internal and external validity and reliability.

Research design	Yes	No	U/D
• The study's purpose and research aims are clearly stated.	1	0	0
• Qualitative methods of inquiry are appropriate for the study aims. (The research sought to understand, illuminate, or explain the subjective experience or views of those being researched in a defined context or setting.)	0	0	
• The authors discussed why they decided to use qualitative methods.	1	0	0
Total research design score:_____			
Sampling			
• Participant selection is clearly described and appropriate	1	0	0
• The sample size is discussed and justified.	1	0	0
Total sampling Score:_____			
Data collection			
• Data collection methods are clearly described and justified.	1	0	0
• The methods are appropriate given the study aims and research questions.	1	0	0
Total data collection score:_____			
Data analysis:			
• The analytic process is clearly described.	1	0	0
• All relevant data were taken into account.	1	0	0
• The authors considered/discussed contradictory evidence and data.	1	0	0
• The study included triangulationm (namely, comparison of different sources of data re: the same issue).	1	0	0
• Triangulation produced convergent conclusions.	1	0	0
• If "no," was this adequately explained?	1	0	0
• Study findings were generated by more than one analyst.	1	0	0
Total data analysis score:_____			
Findings/Results:	Yes	No	U/D
• There is a clear statement of the findings.	1	0	0
• The study findings are discussed in terms of their relation to the research questions posed.		0	0
• The findings appear credible.	1	0	0
• Sufficient data are presented to support findings.			
• Potential researcher biases are taken into account.			
• Conclusions are explicitly linked with exhibits of data.	1	0	0
Total findings/results score:_____			
Research value:			
• Study findings contribute to the current knowledge base.	1	0	0
• Findings can reasonably be expected to inform current practices or policies.	1	0	0
• These contributions are discussed by the authors.	1	0	0
• The authors identified new research areas.		1	0
• The authors discussed how the research findings could be used and for what populations.	1	0	0

Research design	Yes	No	U/D
• Enough descriptive detail was included to allow readers to make their own judgments about potential transferability to other settings.	1	0	0
Total research value score:			
Total score: (sum of all domains)			

U/D unable to determine

Please identify the suitability of the qualitative study to generate knowledge, facilitate interpretation of relevant quantitative studies, and/or illuminate critical factors anticipated to influence the effectiveness of an intervention.

High value: The qualitative study addresses important research questions about the intervention and outcomes of interest (a minimal criterion for even considering it in the review) and the study design is appropriate for addressing those questions [implying that it is well-documented in the paper(s)] and the findings are credible and make a contribution to our understanding of the relationship between the intervention and the outcome that we otherwise would not have based on the quantitative studies alone.

Fair value: The study addresses important questions, is well designed, and adds support for other findings but does not contribute substantial new knowledge.

Low value: The study addresses important questions, but its contribution to our understanding of the issue is not apparent, due to lack of rigor in the study, inadequate documentation of the study design and/or findings.

Suitability: _____

Adapted from: *B*eck (1993); Critical Appraisal Skills Program (CASP) (2002); Rychetnik and Frommer (2002); Miles & Huberman (2002).

References

Beck, C. T. (1993). Qualitative research: the evaluation of its credibility, fittingness, and auditability. *Western Journal of Nursing Research*, *15*(2), 263–266.

Briss, P. A., Zaza, S., Pappaioanou, M., Fielding, J., Wright-De Agüero, L., Truman, B. I., et al. (2000). Developing an evidence-based guide to community preventive services-methods. The Task Force on Community Preventive Services. *American Journal of Preventive Medicine*, *18*(Suppl. 1), 35–43.

Critical Appraisal Skills Program (CASP). (2002). Milton Keynes Primary Care Trust.

Downs, S. H., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomized and non-randomized studies of health care interventions. *Journal of Epidemiology and Community Health*, 52, 377–387.

Miles, M. B., & Huberman, A. M. (2002). *Qualitative data analysis*, Thousand Oaks, CA: SAGE.

Rychetnik, L., & Frommer, M. (2002). *A schema for evaluating evidence on public health interventions V.4*. National Public Health Partnership, University of Sydney, April 2002

Chapter 3
Evidence-Based Public Health: Origins, Assumptions, and Cautions[1]

Luis A. Avilés and Dani Filc

Evidence-Based Public Health: Origins, Assumptions, and Cautions

The adoption of evidence-based approaches to medicine has been rapid and pervasive, with books, journals and websites devoted to everything from evidence-based radiation oncology to evidence-based complementary and alternative medicine. While adherents were initially professionals from the health sciences, the overt focus on evidence-guided practice has moved beyond the health sciences to be embraced within such disparate fields as environmental management, social work, and library sciences. There is a popular book on evidence-based medicine that provides guidelines on how to read a scientific paper, and there is even a mystery novel in which the main character is a physician whose evidence-based medicine skills allow him to solve puzzling murders (Godwin & Hodgetts, 2003). However, it is important to keep in mind that an emphasis on evidence-based practice was initially introduced in the field of medicine, and defined as "the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients" (Sackett, Rosenberg, Gray, Haynes, & Richardson, 1996). Evidence-based medicine (EBM) attempts to ground clinical practice not on clinical intuition or idiosyncratic judgment but on the best existing scientific evidence. The initial propositions and methods of EBM evolved to a different level with the creation of the Cochrane Collaboration and its Cochrane Database of Systematic Reviews, which are now considered the most ambitious and rigorous source of EBM.

Similarly, evidence-based public health (EBPH) is defined as "a public health endeavor in which there is an informed, explicit, and judicious use of evidence, that has been derived from any of a variety of science and social science research and evaluation methods" (Rychetnik, Hawe, Waters, Barratt, & Frommer, 2004). The motivation to adopt this normative framework likewise responds to the need for health policies and population-based interventions to be grounded in "sound facts."

The disciplines of medicine and public health should in principle have no objections to the implementation of evidence-based methods, as nothing is more essential to scientific inquiry than the

[1] This chapter is based on a paper presented at the American Public Health Association Annual Meeting, 2005, sponsored by the Spirit of 1848 Caucus Session, Evidence-based Public Health: Critical Histories and Contemporary Critiques.

L.A. Avilés (✉)
Department of Social Sciences, University of Puerto Rico, Mayagüez, PR, USA
e-mail: luis.aviles3@upr.edu

A. Handler et al. (eds.), *Reducing Racial/Ethnic Disparities in Reproductive and Perinatal Outcomes: The Evidence from Population-Based Interventions*, DOI 10.1007/978-1-4419-1499-6_3,

production of empirical evidence. However, the naïve form of empiricism that often lies beneath the surface of some EBM approaches has generated antagonism from those who argue that EBM promotes a uniformity of thinking and intolerance – or "microfascism," to use their term – as well as methodological narrowness that preclude the necessary scientific debate for the advancement of science (Holmes, Murray, Perron, & Rail, 2006). The enthusiasm generated by evidence-based approaches should be accompanied by a comparable degree of skepticism, which, after all, is a key ingredient of all scientific endeavors. Accordingly, this chapter examines the basic premises of EBM and EBPH, identifying the theoretical foundations and specific features of their respective practices. Our examination intends to encourage dialogue and advance relevant theories that might lead to further evolution of these fields. Before dissecting the specific theoretical assumptions of evidence based approaches, we begin with a brief description of current perspectives on scientific thinking.

A Perspective on Science Discourse

Science is the use of systematic study and methods to acquire knowledge to describe and theoretically explain natural and social phenomena. The image of scientists as individuals working in isolation, passionately pursuing a discovery which may eventually contribute to society and bring them fame and recognition, is common in popular culture. In contrast, most scholars of science understand that scientists do not work in isolation and that the advancement of science requires a community (Porter, 1995). Scientific activity has been described as representing a…

> "particular moment in the division of labor, in which resources, people, and institutions, are set aside in a specific way to organize experience for the purpose of discovery. In this tradition a self-conscious effort has been made to identify sources and kinds of errors and to correct for capricious biases" (Lewontin & Levins, 2007, p. 87).

Scientific progress does not result from the ongoing accumulation of new knowledge under unchanging assumptions; rather science regularly progresses by jumps and discontinuities, while preserving or discarding old ideas, as well as proposing new ones that may represent a radical departure from previous conceptualizations. Landmarks in the history of science, such as Newton's laws of motion, the theory of relativity and the discovery of the DNA macromolecule, represent paradigmatic shifts that changed the way science is practiced and the ways in which scientists view the world. Whether we embrace Kuhn's model of paradigm shifts (Kuhn, 1970) or a model of coexisting and competing paradigms (Fuller, 2000), interactions generated within a scientific community are crucial for the routine conduct of research. In contemporary science, journals have become an essential instrument for orienting new generations of scientists and for reproducing existing paradigms. This chapter will use illustrative examples from "specialized libraries," such as the Cochrane Library, and one selected public health journal to demonstrate their roles in promoting paradigms, as well as creating and strengthening scientific communities.

The Positivist Approach to Science

Emerging in the nineteenth century, a philosophy of science known as positivism incorporated many of the principal ideas of the time and eventually became the prevailing scientific paradigm. Positivism, also known as logical positivism or logical empiricism, continues to play a dominant role in our scientific inquiry today. Initially, positivism represented a progressive and systematic response in the struggle against dogmatism and authority-based knowledge of the preceding epochs. The term 'positivism' was coined by Auguste Comte, who emphasized the prestige and elegance

of the field of physics, which he considered as a model that all other fields should emulate. Consequently, the characteristics of physics, and its quintessential object of study, celestial mechanics (astronomy), emerged as the standard of "true science." Physics developed as a quantitative discipline that studied highly regular, universally observable phenomena; as a consequence positivism has privileged quantitative methods as the most robust framework for investigation of complex natural processes. Quantification aspired to assure objectivity by separating observable facts from opinions; accordingly, practitioners were expected to be experts in numbers and measurements.

In addition to privileging quantitative methods, the focus on physics as a model further promoted the notion of an indisputable neutrality of the scientist who engages in the process of discovering laws that govern natural phenomena. The belief that scientific activities are inherently free from the influence of the personal interests, values or morals of individual scientists, became a basic premise of the research enterprise. Dismissing the social nature of the research process itself created a gap between theory and practice, and precluded opportunities to transform that reality through social action. Developed by natural scientists in the beginning, positivism was later incorporated into the various fields of social science. As a philosophy of science, positivism places emphasis on rationality and objectivity, in addition to the prediction and control of events under study. From the perspective of Joyce Nielsen, positivism is determined by a number of assumptions that shape the scientific study of various phenomena. These assumptions relate to: the knowability and objective reality of the natural and social world; the relationship between subjectivity and objective truth; the universal meaning of research findings and evidence; the cause and effect patterning of social life; and, the primacy of deductive reasoning over potentially valid but less acceptable inductive approaches (Nielsen, 1990). Generating knowledge and evidence by applying these assumptions has generally been understood to bring us closer and closer to reality and objective truths.

Although positivism has maintained a persistent hold on the scientific process, opposition from a number of disciplines including feminism, critical psychology, anthropology, ethnography, and social epidemiology, as well as developments in qualitative research, have generated new views of science that are a significant shift away from the central tenets of positivism, into the realm of 'post-positivism'. Where positivists believe that science is all about uncovering 'truths', for a post-positivist, science is about meaning and the creation of new knowledge. The post-positivist also believes that all observations are theory-laden and that scientists, like everyone else are inherently biased by their cultural experiences and worldviews. The assertion that evidence-based approaches are grounded in positivist philosophy is not universally accepted in the public health literature and community (Holmes et al., 2006; Rychetnik et al., 2004). Nonetheless, many evidence-based efforts in public health continue to be influenced by the assumptions of positivism, remaining indifferent to the criticisms of this philosophy with respect to the generation and application of knowledge. While scientific theory evolves and positivism is increasingly challenged, the objective and mechanistic view of positivist philosophy continues to play a dominant role in our scientific practice.

Modern science has successfully replaced the claim of absolute objectivity with that of mechanical objectivity, which can be achieved if scientists follow certain rules and procedures in order to assure a trustworthy production of a measurement (Porter, 1995). Mechanical objectivity assumes that scientists are interchangeable observers, and when using the same methods in conducting scientific research, anyone regardless of their social position (class, gender, race, etc.) will arrive at the same conclusions. The emphasis on quantification has served to promote the belief that science is above and distinct from the personal interests of scientists, implying that adherence to a particular version of the scientific method makes science immune to the influence of politics and ideology. The use of qualitative research methods in confronting 'objectivity' in the scientific process is not necessarily the antithesis of positivism. While qualitative methods offer the researcher the opportunity to embrace subjectivity and to acknowledge the researcher's social position, they may simultaneously embrace other positivist assumptions. The positivist tenet of universal validity places less emphasis on the importance of time, place, and culture in mediating natural and social

phenomena, despite their potential as determining factors in both the biological and social sciences. Taken to the extreme, positivism holds that true science is trans-historical and trans-cultural. Recent research emphasizing ecological and life span frameworks illuminates the limitations of such concepts in scientific inquiry (Collins, Wambach, David, & Rankin, 2009; Halfon & Hockstein, 2002; Lu & Halfon, 2003; Watt, Carson, Lawlor, Patel, & Ebrahim, 2009).

The positivist assumptions regarding the neutrality of science, consider science and policy as two distinct realms of human endeavors. However, over time, the scientific community has recognized the need to move away from this absolute claim of neutrality. For example, many nuclear physicists needed to come to terms with the social consequences of the massive release of energy that they made possible; more recently scientists have been required to publicly acknowledge potential and real conflicts of interests that may influence their research and create bias, such as receiving research funding from a pharmaceutical company. The current standard that demands a relative disinterest of the scientist constitutes by itself a questioning of the assumption of the neutrality of science. Nevertheless, the separation of scientific inquiry from its social consequences maintains and promotes the premise of scientific neutrality. Any valid analysis of science must acknowledge its dual nature. While "it enlightens us about our interactions with the rest of the world, producing understanding and guiding our actions…as a product of human activity, science reflects the conditions of its production and the viewpoint of its producers or owners" (Lewontin & Levins, 2007, p. 90).

A Critique of Positivist Science as a Foundation for Evidence-Based Medicine

As noted above, challenges to positivist science are rooted in numerous disciplines and traditions, from Marxism, feminism, and post-modernism, to chaos, complexity, and critical theory. Leading criticisms have centered on the constructs of empiricism, exclusivity, autonomy, neutrality, and objectivity (Johnston, Gregory, Pratt, & Watts, 2000). Potential consequences of empiricism in the health sciences include the proliferation of atheoretical and ahistorical research. Notably, the validity of scientific statements is usually mediated and conditioned by the particularities of history and culture. Ample evidence demonstrates that adherence to context-stripping theories with universal validity are sometimes more difficult to sustain than context-sensitive historically and culturally located theories (Briggs, 2002; Haraway, 1991; Harding, 2006; Terry & Urla, 1995). It is now widely accepted that the validity of scientific assertions is – to varying degrees – dependent on context. Modern physics even asserts that the motion of objects varies according to their context; thus, Newton's universal law of gravity does not hold at the scale of subatomic or galactic distances. Feminist critique argues through the concept of positionality, that in gender-, class-, and race-stratified societies it is not possible to have a disinterested, impartial, value-free, or detached scientific perspective (Harding, 1991). Science and scientific inquiry are definitively normative endeavors which influence the ways in which society is conceptualized and organized.

The origins of EBM were highly influenced by the assumptions and tenets of positivist philosophy and current medical practice is still often burdened by their limitations. To illustrate the influence of positivism on the practice of EBM, this section examines a series of reproductive and perinatal health studies from the Cochrane Library. Recognized as the world's leading authority on EBM, the library aims to facilitate decision-making related to clinical care, health services and programs, as well as population based interventions. The Cochrane Library was created by the Cochrane Collaboration, an international not-for-profit organization founded in 1993. The Collaboration establishes Cochrane review groups which generate systematic reviews on the state of our knowledge and evidence related to specific health and related issues.

The systematic reviews of the Cochrane Library strive to attain a high degree of mechanical objectivity, through the use of precise inclusion and exclusion criteria in assessing quality and synthesizing the evidence of selected studies, a task difficult to attain through an unstructured literature review (Rychetnik et al., 2004). Seven reviews related to the subject of reproductive and perinatal health outcomes have been selected and are presented in Table 3.1. While these are not a representative sample of systematic reviews, our purpose is to demonstrate that randomized control trials (RCTs) and quasi-randomized control trials (quasi-RCTs) studies were the only designs selected to answer a broad range of research questions, and to illustrate the positivist foundation and underlying assumptions of EBM in such reviews. When a range of scientific methods and approaches are excluded from the possibility of contributing evidence, then, EBM is likely not what the name suggests.

Empiricism and Exclusivity: What Type of Evidence Counts as Evidence?

EBM classifies evidence according to a hierarchy of study designs assumed to provide varying degrees of validity. In this hierarchy, a RCT is considered the gold standard method generating the most robust evidence on the efficacy of interventions. The RCT is followed by other trials, observational studies, comparison of descriptive studies, and finally expert opinions and case studies. This hierarchy of designs implies that there are scientific approaches which provide evidence with varying degrees of legitimacy, regardless of the levels of complexity for the problems/issues being addressed. This approach marginalizes and to some extent dismisses complex problems and interventions by reducing them to their simple component parts. Such simplification distorts the reality and meaning of the complex issue or system, which is more than the sum of its parts. When diverse research methods are accepted as valid, but their relative importance is ranked on a linear scale, exclusivity can limit the scope of a systematic review and evaluation process. While acknowledging that RCTs offer a unique advantage in studies of the efficacy of therapeutic interventions (Last, 2001), clearly the biggest challenge is to select a design that best fits the research question and adequately represents the populations of interest with maximum validity and reliability. An overemphasis and priority on randomized trials undoubtedly influences the types of studies that receive funding, with the potential to impact progress in the identification and full implementation of new and improved interventions. Uncritical acceptance of evidence has also led to the introduction of ineffective and dangerous practices on numerous occasions (Anderson et al., 2004; Barrett-Connor, 2007; Dalen & Bone, 1996; Rossouw et al., 2002) The important issue therefore, is the extent to which EBM constrains the range of sources for information that could inform structured reviews and strengthen our collective knowledge base.

A summary of the systematic review of the efficacy of fertility awareness methods by Grimes and colleagues (Table 3.1, Study #1) illustrates the limitations of exclusivity in EBM. The objective of this study was to assess the relative efficacy of the rhythm, “natural family planning” method and periodic abstinence methods of contraception. The review was conducted by exclusively selecting RCTs, published in any language that compared any fertility awareness-based contraceptive methods with another method or a placebo. The researchers conclude that the comparative efficacy of these methods could not be assessed, due to the lack of quality RCT (Grimes, Gallo, Halpern, Nanda, & Schultz, 2004). However, the scientists in the review group did not attempt to include other types of studies. The claim that there is no evidence on which to substantiate the efficacy of the fertility awareness methods in the absence of an RCT or a quasi-RCT seems difficult to support.

Table 3.1 A sample of Cochrane Library reviews related to reproductive and perinatal health

Study		Objectives	Selection criteria for studies	Results
1	Fertility awareness-based methods for contraception (Grimes et al., 2004)	To determine the efficacy of fertility awareness-based methods used for contraception ("rhythm," "natural family planning" and "periodic abstinence")	Randomized controlled trials in any language that compared any fertility awareness-based methods for contraception with a placebo or another method	The comparative efficacy of fertility awareness-based methods of contraception remains unknown, because of poor methodological quality
2	Hormonal versus non-hormonal contraceptives in women with diabetes mellitus type 1 and 2 (Visser et al., 2006)	To investigate whether progestogen-only, combined estrogen/progestogen or non-hormonal contraceptives differ in terms of effectiveness in preventing pregnancy, side effects, or long-term complications when used in women with diabetes mellitus	Randomized and quasi-randomized controlled trials that studied women with diabetes mellitus comparing hormonal and non-hormonal contraceptives. Three randomized controlled trials were included. Only one was of good methodological quality	The three included randomized controlled trials in this systematic review provided insufficient evidence to assess whether progestogen-only and combined contraceptives differ from non-hormonal contraceptives in diabetes control, lipid metabolism and complications
3	Surgical methods for first trimester termination of pregnancy (Kulier, Fekih, Hofmeyr, & Campana, 2001)	To compare the safety and efficacy of different surgical methods for first trimester abortion	Three trials were included, resulting in two comparisons: vacuum aspiration versus dilatation and curettage and flexible versus rigid vacuum aspiration cannula	There were no statistically significant differences for excessive blood loss, blood transfusion, febrile morbidity, incomplete or repeat uterine evacuation procedures, re-hospitalization, post operative abdominal pain, or therapeutic antibiotic use
4	Medical methods for first trimester abortion (Kulier et al., 2004)	To compare different medical methods for first trimester abortion	Randomized controlled trials comparing different medical methods (e.g. single drug, combination), ways of application, or different dose regimens, single or combined, for medical abortion, were considered. Thirty-nine trials were included in the review	Safe and effective medical abortion methods are available. Combined regimens are more effective than single agents. Misoprostol vaginally is more effective than orally. Some of the results are based on small studies only and therefore carry some uncertainty. Almost all trials were conducted in hospital settings with good access to support and emergency services. It is therefore not clear if the results are readily applicable to under-resourced settings where such services are lacking even if the agents used are available

5	Medical versus surgical methods for first trimester termination of pregnancy (Say et al., 2002)	To evaluate medical methods in comparison to surgical methods for first-trimester abortion with respect to efficacy, side effects and acceptability	Randomized trials of any surgical abortion method compared with any medical abortion method in the first trimester. Six studies mostly with small sample sizes, were include	Medical methods for early termination of pregnancy can be safe and effective. Prostaglandins used alone seem to be less effective and more painful compared to surgical first-trimester abortion. However, there is inadequate evidence to comment on the acceptability and side effects of medical compared to surgical first-trimester abortions
6	Interventions for emergency contraception (Cheng et al., 2004)	To determine the effectiveness of different emergency contraceptive methods – drugs or intrauterine devices (IUD) – in preventing pregnancy after unprotected sex	Randomized controlled trials and controlled clinical trials including women attending services for emergency contraception following a single act of unprotected intercourse were eligible Forty-eight trials with 33,110 women were included. Most trials were conducted in China (37/48)	Levonorgestrel and mifepristone are very effective with few adverse effects, and are preferred to oestrogen and progestogen combined. Levonorgestrel could be used in a single dose (1.5 mg) instead of two split doses (0.75 mg) 12 h apart. Mifepristone might delay the following menstruation. Women need to be informed about this to avoid anxiety. Another effective method for emergency contraception is the IUD and it can be kept for ongoing contraception
7	Regional versus general anaesthesia for caesarean section (Afolabi, Lesi, & Merah, 2006)	To compare the effects of regional anaesthesia (RA) with those of GA on the outcomes of CS	Randomized and quasi-randomized controlled trials evaluating the use of RA and GA in women who had CS for any indication Sixteen studies (1,586 women) were included in this review	There is no evidence from this review to show that RA is superior to GA in terms of major maternal or neonatal outcomes. Further research to evaluate neonatal morbidity and maternal outcomes, such as satisfaction with technique, will be useful

Universality: What is the Foundation of the Validity of the Evidence?

Accepting the positivist assumption that science produces universally valid results, the aim of the Cochrane Collaboration is to produce and disseminate conclusions that are "evidence-based across all areas of health care, providing health care decision-makers around the world with high-quality, timely research evidence" (Cochrane Library, 2006). By relying exclusively on the RCT and quasi-RCT study designs, EBM assumes that biological responses to medical interventions are overwhelmingly consistent across population strata (Victora, Habicht, & Bryce, 2004). While the practitioners of EBM argue that it "is not restricted to randomized trials and meta-analyses" (Sackett et al., 1996), the construction of evidence summarized for the Cochrane Collaboration relies almost exclusively on RCTs. While the social and historical location or context of any one research sample could potentially lead to effect modification, EBM assumes that conclusions based on subjects included in a limited set of studies conducted in a specific way can be generalized to most populations. Yet, there are limited data to justify this conclusion.

The work of Chronbach and colleagues in 1972 demonstrated that there is no ontological basis for defining a reference universe for a collection of objects and they noted that inclusion and exclusion criteria with respect to belonging to a particular 'reference universe' is empirical, relying on the principle of similarity (Chronbach, Gleser, Nanda, & Rajaratnam, 1972). While the definition of a reference universe is a necessity for the appropriate and accurate applications of research findings and decision-making, the complexity of this principle is often overlooked. As noted by Potvin, defining a reference universe is more than a technical process and requires three normative judgments. First, is the evaluation of similarity of objects/subjects based on their properties; second, is the measurement of these properties; and lastly, is the estimation and determination of variation thresholds beyond which two objects can no longer be considered part of the same reference universe (Potvin, 2006). Interestingly, with the advent of pharmacogenomics, evidence of varying drug responses within and across populations is beginning to emerge. Some countries, such as Japan, have required local trials demonstrating efficacy before new drugs can be marketed (http://www.ich.org). It is becoming increasingly recognized that it is essential for scientific inquiry and study design to consider and account for the social, cultural, and historical construction of responses to treatment and interventions according to strata such as social class, gender, race, and ethnicity.

A review of the effectiveness of hormonal vs. non-hormonal contraceptives in women with diabetes mellitus type 1 and 2 by Visser and colleagues (Table 3.1, Study #2) illustrates the limits of reliance on the positivist assumption of universality within EBM. The researchers concluded, without considering issues of race, ethnicity or class, that there is insufficient evidence to assess whether hormonal or non-hormonal contraception was more effective for women with type 1 and type 2 diabetes mellitus (Visser, Snel, & Van Vliet, 2006). However, the heterogeneity of women in the United States, a population divided along class, race, and ethnic lines, calls into question the universal validity of any medical 'one size fits all' statement. Racial and ethnic minorities in the United States, with the exception of Alaskan natives, have different levels and patterns of risk factor exposures for diabetes and an established prevalence of non-insulin-dependent diabetes mellitus that is two to six times greater than that of whites (Carter, Pugh, & Monterrosa, 1996); likewise, in situations where African-American women receive similar care to that of white women, they have a lower level of diabetic control (Bonds et al., 2003). Cultural differences within the United States influence communication between patients and their professional providers which can significantly affect patients' understandings of the causes, symptoms, progression, and treatment of diabetes (Drozd, 2000). Sexual orientation may also act as an effect modifier, particularly for women, putting some at higher risk of illness (Mays, Yancey, Cochran, Weber, & Fielding, 2002).

RCTs are considered to have methodological advantages in generating new knowledge. Yet, the assumed advantage of investigator control over experimental conditions may result in an accompanying threat to the validity of the results. The controlled conditions of an RCT most often cannot be replicated in ordinary practice, as recognized in the conceptual difference between effectiveness and efficacy. Early on, Cochrane articulated a standard for evaluating interventions, defining the concept of effectiveness as "a measure of the extent to which a specific intervention…deployed in the field in routine circumstances does what it is intended to do for a specified population" (Last, 2001, p. 57). In contrast, efficacy is a measure of "the extent to which a specific intervention produces a beneficial result under ideal conditions," presumably determined by a randomized control trial (Last, p. 58). When the ideal conditions of an experiment are difficult to replicate, a particular intervention can be very efficacious but ineffective. Such is the case of abstinence as a method for avoiding teenage pregnancy, which is highly efficacious when practiced under ideal conditions, but as any high school teacher can confirm, has a very poor record of effectiveness in a population of 'typical' users. Additionally, despite standardized evidence-based guidelines of care, uniform implementation remains a challenge. It may be difficult for providers to follow them in some clinical settings due to time constraints, practice economics, uneven geographic distribution of resources, and low reimbursement rates, among other factors (Larhme & Pugh, 2001).

A study by Kulier and colleagues on *Medical Methods for First Trimester Abortions* (Table 3.1, Study #4). reports that the methods used to induce a medical, non-surgical, abortion, are safe and effective (though the term that should be used is efficacious.) The review group was prudent in asserting that, since all trials were conducted in hospitals it was not clear if the results would be applicable to non-hospital settings, thus suggesting the need for caution with respect to the external validity of the results (Kulier, Gülmezoglu, Hofmeyr, Cheng, & Campana, 2004). A quantitative measure of an intervention under the ideal conditions of an RCT may be limited when those conditions are replaced by the actual circumstances of different social groups or cultures. Thus, universal application of study findings is not always appropriate.

Autonomy: Does Evidence-Based Medicine Guarantee the Neutrality of Science?

When characterized as a source of "high-quality, independent evidence for health care decision making" (Cochrane Library, 2006), the Cochrane Library appears to affirm the independence of evidence, which corresponds to the tenet of positivist neutrality in scientific inquiry. While it is asserted that RCTs produce objective knowledge, there are multiple aspects of the research undertaking that precede the conduct of the trial itself, most prominently the selection of the research question. Identifying an issue to investigate depends on the established priorities of funding agencies (government, private foundations, and corporations), researchers' personal interests (which include personal rewards such as tenure track jobs, funded research, and prestige), and the interests of organizations that train future researchers, all of which are significantly influenced by the values, biases, and power dynamics of the society at large. The lack of autonomy and neutrality in generating EBM is also manifest in the normative implications of the observed research results, as demonstrated by the research on medical abortion and emergency contraception listed in Table 3.1 (Studies #4, #5, and #6). (Chenget al., 2004; Kulier et al., 2004; Say, Kulier, Gülmezoglu, & Campana, 2002). Within the United States, society is polarized around issues of medical abortion and emergency contraception. Research in this area is not a matter of minor consequence and reflects a particular political perspective. Deciding to conduct research on the safety and efficacy of medical abortions is far from a neutral undertaking. This research agenda is supported by many

feminist groups who have been advocating for more research on alternatives for fertility regulation methods that respond to the diverse needs of women of different classes, races, and cultural groups, as well as those who declare themselves 'pro-life' and strongly oppose abortion and the right of women to control their own reproduction.

While Euro- and androcentric biases in research have steadily decreased, scientists continue to play key roles in the reproduction of gender, class and race hierarchies both through the issues they choose to study and the meanings and policy implications ascribed to their research findings. These findings tend to identify, describe and emphasize differences that reinforce existing social biases and power differentials in our social structure, rather than contribute to a critical analysis of existing inequalities (Weber, 2006).

Many supporters of EBM acknowledge the need for a broader perspective toward evidence rather than the use of the stringent criteria characteristic of the Cochrane Library. Some believe that clinical practice should integrate EBM with individual clinical expertise; while neither is sufficient for quality care, as noted by Sackett and colleagues, there is a risk of practice becoming tyrannized by excellent evidence that may be inappropriate or unsuitable for an individual patient (Sackett et al., 1996). Scientific efforts to promote the value and utility of knowledge resulting from the application of theory in a given situation along with knowledge resulting from the accumulation of empirical observations are critical (Potvin, 2006). As we strive for excellence in research methodology and design, it is critical to remember that what should guide the design of research is the significance of the question; important questions that do not fit into an RCT format should not be ignored. There are a number of alternatives to the RCT study design, along with statistical strategies appropriate for generating valid public health evidence and advancing knowledge (West et al., 2008).

Is Evidence-Based Medicine a Template for Public Health?

The separation of the fields of medicine and public health, at times called a schism (White, 1991), rests on the idea that there is an artificial distinction between these fields produced by particular historical circumstances. More appropriately, clinical medicine should be conceived of as a division within the overarching field of public health. While both contribute to the goal of enhancing health, they are driven by distinctly different theories and principles of practice. The philosophical basis of public health as social justice along with its community and population-based approach necessitates completely different concepts, constructs, and methodologies to that of clinical medicine.

We have attempted to provide a framework for understanding the historical underpinnings and philosophical basis of EBM, while advising caution in its application to public health. We have encouraged skepticism where appropriate and tried to promote critical thinking regarding the relevance of positivist-based science for both EBM and EBPH. Constraints of EBM and implications for public health have been articulated by the British social epidemiologist, George Davey-Smith and colleagues:

> "The sort of evidence gathered on the benefits of interventions aimed at individuals may not help in guiding policies directed towards reducing health inequalities…High variance apparently 'explained' by individual-level risk indicators (or markers manipulable in a discrete way within populations) does not mean that they are important determinants of the population level of any outcome. These are, however, precisely the factors that evidence-based research focuses on." (Davey-Smith, Ebrahim, & Frankel, 2001, pp. 184–185)

In the following section we present specific examples of EBPH that do not include a number of positivist assumptions yet demonstrate rigor and quality in their inquiry.

Evidence-Based Public Health: Moving Beyond Positivist Philosophy

The field of public health is characterized by complexity, given the inter-connectedness and continual state of change of the issues addressed, the diversity of study populations, the dynamics of social environments, and the effect of all of these on population health. Recognition of this complexity would seem to dictate that EBPH be based on a variety of study designs both quantitative and qualitative, without hierarchal priority, yet driven by methodological appropriateness and soundness for the questions being addressed (Jackson & Waters, 2005; Popay, 2001). To illustrate the extent to which such an approach characterizes EBPH and distinguishes it from EBM, we explored the *Journal of Epidemiology and Community Health* (JECH) as a source of evidence-based public health studies related to reproductive health.

Beginning in August of 2003, JECH was the first journal to initiate a section exclusively devoted to "*Evidence-Based Public Health Policy and Practice*." From August 2003 through December 2006, JECH published 542 articles in the areas of research reports, theory and methods, and evidence based public health policy and practice. A total of 145 articles were published under the heading of EBPH practice. Within the EBPH category, there were seven articles related to reproductive and perinatal health. These studies, presented in Table 3.2, are used to illustrate the non-positivist assumptions and principles of EBPH and demonstrate the range of study designs and diversity of research methods including: prospective analysis, ecological analysis, systematic reviews, toxicological analysis, analysis of hospital records, and qualitative analysis of in-depth interviews from which public health evidence can be generated.

Pilkington and colleagues (Table 3.2, Study #5) studied mothers' belief systems and the use of health care for their children in a malaria-endemic rural area of Gabon, illustrating that EBPH benefits from the transfer of knowledge that is qualitatively generated. In contrast to the assumed 'ideal' of mechanical objectivity, this in-depth interview-based study shows that trust plays a significant role in research by promoting and encouraging participation as well as generating meaningful communication. Such traits are not likely generated through standardized instruments and 'mechanical objectivity' (Pilkington, Mayombo, Aubouy, & Deloron, 2004).

EBPH depends on study designs that ensure validity and the generalizability of research findings to populations and contexts of interest. It embraces local research for local purposes; thus, the diversity of methods that can be used for EBPH allows for the possibility of research questions that make no claim of universality. The needs of a community, either a rural village in central Africa, or a war-torn country in Asia are legitimate subjects of study with no expectation of universality. Prasad's study of children in Kabul had the modest goal of generating data to understand the specific heath needs of children in this city (Table 3.2, Study #4) (Prasad, 2006). No doubt that the knowledge generated might be relevant for other regions of the country or for other countries facing war, but there was no intention for the research results to be generalized to other populations beyond the one that is targeted. With careful investigation of health phenomena within their real life context, EBPH promotes the framing of research within particular historical circumstances, including the history of the community/society where the research takes place, as well as the accumulated experience and understanding of the subject being examined. Locally contextualized research tests the robustness of a theory and hypothesis within diverse geographical, cultural, and historical circumstances, generating evidence and providing significant contributions to our scientific knowledge base for public health policy and practice.

When EBPH deals with issues that respond to local needs, and the policy implications at stake are sufficiently obvious, it becomes more apparent how the interests of researchers and society may be compatible. Take for example, the study of Mercury Contamination in Reproductive Age Women in Vieques, Puerto Rico (Table 3.2, Study #3) (Ortiz-Roque & López-Rivera, 2004). This research was conducted at a time when massive civil disobedience acts were taking place in Vieques,

Table 3.2 Articles related to Reproductive Health and Perinatal Health published in the Evidence Based Public Health Policy and Practice section from the Journal of Epidemiology and Community Health, 2003–2006

Study		Objectives	Methods	Results
1	Does an increase of low income families affect child health inequalities? A Swedish case study (Bremberg, 2003)	To determine the effect of the increase in the proportion of low income families on child health inequalities in the high income welfare state of Sweden	Social health inequalities were analyzed ecologically at the level of the primary health care district and were statistically related to a series of health indicators, which included low birth weight, abortions, etc.	During the study period there was a weak connection between family income and child health inequalities
2	Sudden unexpected death in infancy and socioeconomic status: A Systematic Review (Spencer & Logan, 2004)	To systematically review observational studies documenting the relation of sudden unexpected death in infancy (SUDI) and socioeconomic status	A review of 52 articles of SUDI indexed in Medline or Embase databases (case-control and cohort studies) that related the condition with some measure of socioeconomic status	An association low socioeconomic status was found with SUDI. A series of studies documented this relation even after controlling for factors such as maternal smoking and sleeping position
3	Mercury contamination in reproductive age women in a Caribbean island: Vieques (Ortiz-Roque & López-Rivera, 2004)	To assess levels of mercury contamination among reproductive age women in heavily industrialized areas of Puerto Rico	Analysis of hair mercury concentration levels among a sample of women of reproductive age in two Puerto Rico sites: northeastern Puerto Rico and the island of Vieques	Women of reproductive age in the Vieques site were exposed to mercury concentration levels that are unsafe and teratogenic
4	Disease profile for children in Kabul: The unmet need for health care (Prasad, 2006)	To generate data to understand health care needs in Afghanistan to plan accordingly	A retrospective analysis of 17,580 hospital records of children (less than 12 years old) visiting a hospital in Kabul, Afghanistan over a 12 month period ending on September, 2003	The most important causes of morbidity in children, acute respiratory infection and diarrhea (36% of all cases). Among neonates, there were a considerable proportion of preterm neonates (20%). A high incidence of Down's syndrome suggests women have children until an elderly age. The lasting impact of war on the health care system and the health status of children can well be seen here

5	Malaria from natural to supernatural: A qualitative study of mothers' reaction to fever in Dienga, Gabon (Pilkington et al., 2004)	To determine how mothers react when facing a child with fever	A qualitative study based on in-depth semi structured interviews of mothers and non structured interviews with traditional healers in the village of Dienga, a malaria endemic rural area of Gabon	Mothers' perceived malaria as a febrile illness. Mothers and healers use disease severity to determine disease etiology (natural or supernatural) and act accordingly to seek appropriate care in the health system or though traditional healers
6	Evaluation of the impact of the Family Health Program on infant mortality in Brazil, 1990–2002 (Macinko et al., 2006)	To assess the impact of Brazil's Family Health Program on state level infant mortality rates	An ecological study was conducted, using the state (province) as a unit of analysis, to relate infant mortality and the increase in the Family Health Program coverage, after controlling for a series of variables such as access to clean water, average income, and women's literacy	The Family Health Care Program is associated with a partial reduction of the infant mortality rate
7	Child care and social support modify the association between maternal depressive symptoms and early childhood behavior problems: A US national study (Lee, Halpern, Hertz-Picciotto, Martin, & Suchindran, 2006)	To investigate the presence and patterns of modification effects of the sex of the child, social support, and childcare on the relation between maternal depressive symptoms and child behavior problems	Prospective longitudinal study of 1,216 families from 10 locations across the USA	The association between child externalizing behavior problems and maternal depressive symptoms varied according to the social support received by the mother

demanding the termination of bombing practices of the U.S. military in that island of the Puerto Rican archipelago; the research results were another piece of evidence to support antimilitary activism. Similarly, the Children of Kabul study (Table 3.2, Study #4) (Prasad, 2006) is an authoritative document against war. Of particular relevance is the study of Brazil's Family Health Program by Macinko and colleagues (Table 3.2, Study #6), which demonstrates how an evaluation of a specific government project can provide evidence and illuminates policy implications that promote advocacy to support, improve, or eliminate a program (Macinko, Guanais, & Marinho de Souza, 2006).

Public health scientists are not mere messengers or commentators with 'accurate' data; in moving past positivism, most public health researchers conduct socially responsible science motivated by the growing unacceptable inequalities in society signifying that positivist assumptions of neutrality are not the foundation of public health inquiry.

The Contribution of Evidence-Based Public Health

As discussed earlier in this chapter, a methodological advantage of the evidence-based approach to scientific inquiry is the adoption of a systematic review process summarizing and integrating evidence that has been published. A major advantage over the traditional literature review, systematic reviews include specific inclusion and exclusion criteria for studies of interest. The evidence-based movement succeeded in reinvigorating the focus on rigorous research across diverse professional activities where the quality of evidence was weak, precisely at a time when there was an external environment that promoted anti-science attitudes in U.S. society. An unprecedented attack on science in government has been publicly documented (Glanz, 2004; The Economist, 2004), and the field of public health is one of the key targets of anti-science attitudes. We observed the removal of information about condoms from the Centers for Disease Control website, the distortion of facts regarding the safety of abortion, and neglect of the FDA's scientific advisory committee recommendations for providing over-the-counter status to emergency contraception (Chavkin, 2004).

Additionally, the discipline of public health has witnessed a series of sophisticated tactics used by interest groups that, in the name of science, distort scientific processes. Some examples include the economic manipulation of industry sponsored research, delaying practices to obstruct the release of scientific information, hiding their identity to conceal their blatant activities, and harassing organizations and individual scientists for publishing inconvenient research results that disturb particular interest groups (Rosenstock & Lee, 2002). Within a context of forces opposing the scientific use of evidence, either because of traditional naïve ignorance, idiosyncratic judgment, politically motivated obscurantism, allegiance to new age holistic health gurus' advice, or the particular motivations of interest groups, an evidence-based approach is required to enhance and promote the public's health.

Public health researchers are increasingly addressing questions for which the RCT is not a practical or ethical option, or for which an RCT should be complemented by alternative designs to enhance generalizability to populations and contexts of interest. Our challenge is to deepen understanding through the appropriate use of all the tools at our disposal to not only describe, but to reveal, explicate, and intervene with the processes which shape the health of communities and individuals. The complexities of our public health problems require complex thinking and responses. These responses and associated interventions need to be based on relevant and meaningful theory that maintain integrity while including context adaptation in order to strengthen and maximize the intended effects (Hawe, Shiell, & Riley, 2004).

Although traditional public health and clinical research has not strayed far from the historical roots of positivist philosophy embraced by the biomedical paradigm and the psychosocial and bio-behavioral approaches that follow it, these models have increased our knowledge and contributed

to improvements in the health of populations and many individuals. The promise and potential to adequately address and eliminate health disparities and inequities however, necessitates a very different paradigm. The emerging field of evidenced based public health demonstrates the potential for such a paradigm as it strives to successfully dissociate from the tenets of positivism and establish an alternative norm of inquiry that gives priority to social contexts and structure beyond the individual. EBPH is challenged to conceptualize inequality and inequity as privilege and power differentials across multiple dimensions (gender, race, ethnicity, class, nation, and sexual orientation), and not primarily as differences in the accumulation of resources and risk factors.

We must and can generate knowledge directly from the lived experiences of those burdened by the conditions we seek to improve. The success of EBPH will depend to a great extent on its ability to make transparent and assert the values, passion and politics of its research. It will require vigilance in assessing assumptions – those of the society at large, as well as those of the individual researcher. As EBPH evolves, its integrity and successful contribution will require vision, patience, openness, persistence, and creativity, as well as dialogue and debate, as researchers learn along with and from the populations and communities they study.

References

Afolabi, B. B., Lesi, F. E. A., & Merah, N. A. (2006). Regional versus general anesthesia for caesarean section. *Cochrane Database of Systematic Reviews*, (4), CD004350. doi:10.1002/14651858.CD004350.pub2.

Anderson, G. L., Limacher, M., Assaf, A. R., Bassford, T., Beresford, S. A., Black, H., et al. (2004). Effects of conjugated equine estrogen in postmenopausal women with hysterectomy: The Women's Health Initiative randomized controlled trial. *JAMA: The Journal of the American Medical Association*, *291*(14), 1701–1712.

Barrett-Connor, E. (2007). Hormones and heart disease in women: The timing hypothesis. *American Journal of Epidemiology*, *166*(5), 506–510.

Bonds, D., Zaccaro, D., Karter, A., Selby, J., Saad, M., & Goff, D. (2003). Ethnic and racial differences in diabetes care: The Insulin Resistance Atherosclerosis Study. *Diabetes Care*, *26*, 1040–1046.

Bremberg, S. (2003). Does an increase of low income families affect child health inequalities? A Swedish case study.. *Journal of Epidemiology and Community Health*, *57*, 584–588.

Briggs, L. (2002). *Reproducing empire: Race, sex, science, and U.S. imperialism in Puerto Rico*. Berkeley: University of California Press.

Carter, J., Pugh, J., & Monterrosa, A. (1996). Non-insulin-dependent diabetes mellitus in minorities in the United States. *Annals of Internal Medicine*, *125*(3), 221–232.

Chavkin, W. (2004). Access denied, science denied. *American Journal of Public Health*, *94*(8), 1298–1299.

Cheng, L., Gülmezoglu, A. M., Van Oel, C. J., Piaggio, G., Ezcurra, E., & Van Look, P. F. A. (2004). Interventions for emergency contraception. *Cochrane Database of Systematic Reviews*, (3), CD001324. doi:10.1002/14651858.CD001324.pub2.

Chronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioural measurements: Theory of generalizability of score and profiles*. New York: Wiley.

Cochrane Library. (2006). About the Cochrane Library. Retrieved December, 2006, from http://www3.interscience.wiley.com/cgi-bin/mrwhome/106568753/AboutCochrane.html.

Collins, J. W. Jr., Wambach, J., David, R. J., & Rankin, K. M. (2009). Women's lifelong exposure to neighborhood poverty and low birth weight: A population-based study. *Maternal and Child Health Journal*, *13*(3), 326–333.

Dalen, J. E., & Bone, R. C. (1996). Is it time to pull the pulmonary artery catheter? *JAMA: The Journal of the American Medical Association*, *276*(11), 916–918.

Davey-Smith, J., Ebrahim, S., & Frankel, S. (2001). How policy informs evidence: "Evidence based" thinking can lead to debased policy making. *British Medical Journal*, *322*, 184–185.

Drozd, M. (2000). Cultural sensitivity in diabetic care. *Home Health Care Management and Practice*, *12*(1), 1–6.

Fuller, S. (2000). *Thomas Kuhn: A philosophical history for our times*. Chicago: University of Chicago Press.

Glanz, J. (2004, February 19). Scientists say administration distorts facts. *The New York Times*.

Godwin, M., & Hodgetts, G. (2003). *The Bedford murder: An evidence-based clinical mystery*. Philadelphia: Hanley and Belfus.

Grimes, D. A., Gallo, M. F., Halpern, V., Nanda, K., & Schultz, K. F. (2004). Fertility awareness-based methods for contraception. *Cochrane Database of Systematic Reviews*, (4), CD004860. doi:10.1002/14651858.CD004860.pub2.

Halfon, N., & Hockstein, M. (2002). Life course health development: An integrated framework for developing health, policy and research. *The Milbank Quarterly*, *80*(3), 433–479.
Haraway, D. J. (1991). *Simians, cyborgs, and women: The reinvention of nature*. New York: Routledge.
Harding, S. (1991). *Whose science? Whose knowledge? Thinking from women's lives*. Ithaca, New York: Cornell University Press.
Harding, S. (2006). *Science and social inequality: Feminist and postcolonial issues*. Urbana: University of Illinois Press.
Hawe, P., Shiell, A., & Riley, T. (2004). Complex interventions: How 'out of control' can a randomized controlled trial be? *BMJ (Clinical Research Ed.)*, *328*, 1561–1563.
Holmes, D., Murray, S. J., Perron, A., & Rail, G. (2006). Deconstructing the evidence-based discourse in health sciences: Truth, power and fascism. *International Journal of Evidence-Based Health Care*, *4*, 180–186.
ICH harmonised tripartite guideline: E5: Ethnic factors in the acceptability of foreign clinical data E5 [ICH Web site]. Accessed December 16, 2009, from http://www.ich.org.
Jackson, N., & Waters, E. (2005). Criteria for the systematic review of health promotion and public health interventions. Health Promotion International, 20, 367–374.
Johnston, R., Gregory, D., Pratt, G., & Watts, M. (2000). *The dictionary of human geography*. Malden, MA: Blackwell.
Kuhn, T. (1970). *The structure of scientific revolutions*. Chicago: The University of Chicago Press.
Kulier, R., Fekih, A., Hofmeyr, G. J., & Campana, A. (2001). Surgical methods for first trimester termination of pregnancy. *Cochrane Database of Systematic Reviews*, (4), CD002900. doi:10.1002/14651858.CD002900.
Kulier, R., Gülmezoglu, A. M., Hofmeyr, G. J., Cheng, L. N., & Campana, A. (2004). Medical methods for first trimester abortion. *Cochrane Database of Systematic Reviews*, (1), CD002855. doi:10.1002/14651858.CD002855.pub3.
Larhme, A. C., & Pugh, J. A. (2001). Evidence-based guidelines meet the real world: The case of diabetes care. *Diabetes Care*, *24*, 1728–1733.
Last, J. M. (2001). *A dictionary of epidemiology*. New York: Oxford University Press.
Lee, L. C., Halpern, C. T., Hertz-Picciotto, I., Martin, S. L., & Suchindran, S. M. (2006). Child care and social support modify the association between maternal depressive symptoms and early childhood behaviour problems: A US national study. *Journal of Epidemiology and Community Health*, *60*, 305–310.
Lewontin, R., & Levins, R. (2007). *Biology under the influence: Dialectical essays on ecology, agriculture and health*. New York: Monthly Review Press.
Lu, M. C., & Halfon, N. (2003). Racial and ethnic disparities in birth outcomes: A life-course perspective. *Maternal and Child Health Journal*, *7*, 13–30.
Macinko, J., Guanais, F. C., & Marinho de Souza, M. F. (2006). Evaluation of the impact of the Family Health Program on infant mortality in Brazil, 1990–2002. *Journal of Epidemiology and Community Health*, *60*, 13–19.
Mays, V., Yancey, A. K., Cochran, S. D., Weber, M., & Fielding, J. E. (2002). Heterogeneity of health disparities among African American, Hispanic, and Asian American women: Unrecognized influences of sexual orientation. *American Journal of Public Health*, *92*, 632–639.
Nielsen, J. M. (1990). Introduction. In J. Nielsen (Ed.), *Feminist research methods: Exemplary readings in the social sciences* (pp. 1–37). Boulder, CO: Westview Press.
Ortiz-Roque, C., & López-Rivera, Y. (2004). Mercury contamination in reproductive age women in a Caribbean island: Vieques. *Journal of Epidemiology and Community Health*, *58*, 756–757.
Pilkington, H., Mayombo, J., Aubouy, N., & Deloron, P. (2004). Malaria, from natural to supernatural: A qualitative study of mothers' reactions to fever (Dienga, Gabon). *Journal of Epidemiology and Community Health*, *58*, 826–830.
Popay, J. (2001). What should count as evidence for Social Development Interventions – The case of Regeneration Initiatives. In J. Popay (Ed.), *Regeneration and health: A selected review of the literature*. London: Kings Fund & Nuffield Institute for Health.
Porter, T. (1995). *Trust in numbers: The pursuit of objectivity in science and public life*. Princeton: Princeton University Press.
Potvin, L. (2006). Should we worry about the enthusiasm toward evidence-based health promotion practices? *Health Promotion and Education*, *13*, 228–229.
Prasad, A. N. (2006). Disease profile for children in Kabul: The unmet need for health care. *Journal of Epidemiology and Community Health*, *60*, 20–23.
Rosenstock, L., & Lee, L. J. (2002). Attacks on science: The risks to evidence-based policy. *American Journal of Public Health*, *92*(1), 14–18.
Rossouw, J. E., Anderson, G. L., Prentice, R. L., LaCroix, A. Z., Kooperberg, C., Stefanick, M. L., et al. (2002). Risks and benefits of estrogen plus progestin in healthy postmenopausal women: Principal results from the Women's Health Initiative randomized controlled trial. *JAMA: The Journal of the American Medical Association*, *288*(3), 321–333.
Rychetnik, L., Hawe, P., Waters, E., Barrat, A., & Frommer, M. (2004). A glossary for evidence-based public health. *Journal of Epidemiology and Community Health*, *58*, 538–545.
Sackett, D., Rosenberg, W. M. C., Gray, J. A. M., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: What it is and what it isn't. *British Medical Journal*, *312*, 71–72.

Say, L., Kulier, R., Gülmezoglu, M., & Campana, A. (2002). Medical versus surgical methods for first trimester termination of pregnancy. *Cochrane Database of Systematic Reviews*, (4), CD003037. doi:10.1002/14651858.CD003037.pub2.

Spencer, N., & Logan, S. (2004). Sudden unexpected death in infancy and socioeconomic status: A systematic review. *Journal of Epidemiology and Community Health*, *58*, 366–373.

Terry, J., & Urla, J. (1995). *Deviant bodies: Critical perspectives on difference in science and popular culture*. Bloomington: Indiana University Press.

The Economist. (2004, April 7). Cheating nature? The Bush administration has been accused of manipulating science. *The Economist*. Digital version. Retrieved December 2006.

Victora, C. G., Habicht, J. P., & Bryce, J. (2004). Evidence-based public health: Moving beyond randomized trials. *American Journal of Public Health*, *94*(3), 400–405.

Visser, J., Snel, M., & Van Vliet, H. A. A. M. (2006). Hormonal versus non-hormonal contraceptives in women with diabetes mellitus type 1 and 2. *Cochrane Database of Systematic Reviews*, (4), CD003990. doi:10.1002/14651858.CD003990.pub3.

Watt, H. C., Carson, C., Lawlor, D. A., Patel, R., & Ebrahim, S. (2009). Influence of life course socioeconomic position on older women's health behaviors: Findings from the British Women's Heart and Health Study. *American Journal of Public Health*, *99*(2), 320–327.

Weber, L. (2006). Reconstructing the landscape of health disparities research: Promoting dialogue and collaboration between feminist intersectional and biomedical paradigms. In A. J. Schulz & L. Mullings (Eds.), *Gender, race, class, and health: Intersectional approaches* (pp. 21–59). Josey-Bass: San Francisco.

West, S. G., Duan, N., Pequegnat, W., Gaist, P., Des Jarlais, D. C., Holtgrave, D., et al. (2008). Alternatives to the randomized controlled trial. *American Journal of Public Health*, *98*, 1359–1366.

White, K. L. (1991). *Healing the schism: Epidemiology, medicine and public's health*. New York: Springer.

Chapter 4
Access to Family Planning and Induced Abortion

Julie Chor, Ashley Dyer, and Bryna Harwood

Introduction

Providing women with access to family planning – including contraception and abortion – is essential for primary prevention of maternal and infant morbidity and mortality. Given the relationship between unintended and unwanted pregnancies and adverse pregnancy outcomes, when women can safely avoid unintended pregnancies and unwanted births, maternal as well as infant morbidity and mortality is reduced (Mohllajee, Curtis, Morrow, & Marchbanks, 2007). Furthermore, because women of color bear a disproportionate burden of unintended pregnancy and maternal and infant morbidity and mortality, reducing rates of unintended pregnancy and increasing the proportion of pregnancies that are planned and wanted has potential to play a significant role in reducing racial/ethnic disparities in reproductive and perinatal outcomes (Berg, Chang, Callaghan, & Whitehead, 2003; Finer & Henshaw, 2006; Geller, Cox, Callaghan, & Berg, 2006; Jones, Darroch, & Henshaw, 2002).

Role of Family Planning in Maternal and Infant Morbidity and Mortality

There are few proven secondary prevention methods for the most common causes of maternal mortality or severe morbidity such as pregnancy-related hypertensive events and hemorrhage (Berg et al., 2003; Geller et al., 2006). Given that the rate of unintended pregnancy remains at approximately 50% and that unintended births comprise 1/3 of all births in the U.S., family planning interventions have the potential to prevent maternal morbidity and mortality to an equal or greater degree than any secondary prevention strategy such as improved obstetric care (Geller et al., 2004). Importantly, while the endemically high rate of unintended pregnancy impacts every racial and ethnic group in the United States to some extent, notable disparities exist in the rates of unintended pregnancy and use of abortion as a response to unintended pregnancy, based on education level, marital status, age, and race/ethnicity. Data from the 1982, 1988, 1995, and 2002 National Survey of Family Growth show higher rates of unintended pregnancy among Hispanic and Black women, low-income women, unmarried women, women 18–24 years of age, and women of low education level (Finer & Henshaw, 2006). Many of these disparities occur because of differential access to health care; despite the availability of highly effective methods of contraception and safe abortion, many women in need of family planning and abortion services in the U.S. are unable to access such services.

B. Harwood (✉)
University of Illinois, College of Medicine, Department of Obstetrics and Gynecology, Chicago, IL, USA
e-mail: brynah@uic.edu

A. Handler et al. (eds.), *Reducing Racial/Ethnic Disparities in Reproductive and Perinatal Outcomes: The Evidence from Population-Based Interventions*, DOI 10.1007/978-1-4419-1499-6_4,

In contrast to *maternal* outcomes, establishing the role of family planning services in reducing morbidity and mortality among *infants* is a much more complex endeavor given the more distal link between women's control of their reproduction and infant outcomes. Nonetheless, the effectiveness of family planning services and abortion in improving infant health through preventing unplanned and/or unwanted pregnancies is in fact, well documented globally and in the U.S. (Singh, Darroch, Vlassoff, Nadeau, 2009; Sonnenberg, Burkman, Hagerty, Speroff, & Speroff, 2004). Because there is a clear association between planned and wanted pregnancies and improved infant outcomes as well as a clear link between maternal morbidity and adverse infant outcomes, a strong case can be made that access to and use of family planning services is also a key primary prevention strategy related to improving infant outcomes (Atrash et al. 2008; Gruber, Levine, & Staiger, 1999; Meier & Mcfarlane, 1994).

Family planning services also lead to improved maternal and infant health outcomes by contributing to the optimization of maternal health conditions prior to conception, so that when pregnancies do occur they are healthier (Misra & Grason, 2000). Effective family planning, including access to contraception and safe abortion, is an important aspect of preconception care, but is often overlooked or neglected as an intervention to improve maternal and child health (Atrash et al., 2008).

Given the importance of access to family planning services including induced abortion in improving pregnancy outcomes, increasing such access may have a critical role to play in reducing disparities in reproductive and perinatal outcomes. This chapter reviews the evidence base for several intervention strategies designed to increase access to family planning and abortion services in the U.S. When examining the evidence base for strategies to improve access to family planning, the outcomes of interest are increased availability and use of family planning services, cost-savings related to maternal and infant healthcare, prevention of unintended pregnancies and unintended births, and reduction in birthrates. With respect to abortion services, this chapter reviews the evidence base examining the relationship between public funding of induced abortion and access to abortion as well as studies that examine the relationship between access to abortion services and maternal and infant outcomes. Finally, the chapter examines whether increased training of providers in comprehensive reproductive health care improves access to family planning and abortion services.

Methods

This review of the evidence of interventions associated with access to family planning and induced abortion emphasizes interventions employed in the United States, and includes: public financing and support for family planning services, private insurance coverage of contraceptive services, public funding of induced abortion services, and strategies to increase the number of trained comprehensive reproductive health providers. We conducted our literature search from January 2008 through November 2009 using Medline/Ovid and Web of Science databases to search for articles going back to the earliest years of each database and corresponding to the aforementioned interventions, cross-referencing induced abortion and family planning with the following keywords and phrases: access, Medicaid waiver, Title X, training, maternal mortality, maternal outcomes, maternal morbidity, infant outcomes, public financing, unsafe abortion practices, cost-effectiveness, and effectiveness. Once relevant manuscripts were identified, a further search of the authors, bibliographies and cited references was conducted. In addition, we searched online publications of the Alan Guttmacher Institute, the World Health Organization (WHO), and the United Nations Population Fund (UNFPA) for relevant manuscripts and collaborated with experts in the field to identify additional published literature for review. Finally, we reviewed the publications identified with respect to the year each intervention was initiated. For example, federal public financing of contraceptive services was established in 1970, federal legalization of induced abortion was in 1973, Medicaid expansion of

family planning services was begun in 1993 and requirements for induced abortion training in residency training programs were established in 1996.

Improving Access to Family Planning Services

Background: Insurance Coverage Does Make a Difference

Programs designed to provide and/or expand public and private family planning coverage are instrumental interventions to increase access to family planning services. Studies have repeatedly demonstrated that access and uptake of family planning services are typically improved by coverage for those services. A study by Wu and colleagues (2008) used data from the 2002 National Survey of Family Growth (NSFG) to compare risk factors for contraceptive nonuse and found that women who were contraceptive nonusers were more likely to be uninsured (OR 1.6, 95% CI 1.1–2.4), but also found that Medicaid recipients were at increased risk in comparison to women with private insurance (OR 1.9, 95% CI 1.2–2.9) (Wu, Meldrum, Dozier, Stanwood, & Fiscella, 2008). Another analysis of the 2002 NSFG found that young women (18–24 years) at risk for unintended pregnancy were three times more likely to use prescription contraceptives if they had private insurance or Medicaid than if they were uninsured (Nearns, 2009). Likewise, analysis of data on reproductive-age women from the 2002 Behavioral Risk Factor Surveillance System (BRFSS) compared the likelihood of prescription contraceptive use between insured and uninsured women in a nationally representative sample of over 25,000 women. Controlling for socioeconomic characteristics and self-reported overall health assessments, the authors found that a statistically significantly higher percentage of insured women reported using prescription contraceptives as compared to uninsured women (54 vs. 45%). Using multiple regression analysis, the authors found that women who were uninsured were 30% less likely to use prescription contraceptives than were women with either public or private insurance (Culwell & Feinglass, 2007). Although contraceptive methods were not evaluated individually, the analyses suggest that being uninsured poses a substantial barrier preventing some women from accessing the most effective methods to prevent unintended pregnancy.

The provision of family planning services in the United States began in the early twentieth century as a philanthropic effort. Today, government financing of family planning services is a non-cohesive fragmented system consisting primarily of public funding of services (the Title X program, described below) and reimbursement to providers of family services for low-income women (Medicaid, described below). In addition, there are public initiatives such as mandates to ensure coverage of family planning services for those women who have private insurance. An examination of the evolution of funding mechanisms for family planning in the US and their impact on access reveals that adequate and sustained funding is an essential strategy for increasing women's access to family planning services.

Public Provision of Family Planning Services: Title X

The socio-political environment of the 1950s and 1960s in the U.S. was conducive to the development of federal funding for family planning services (Goldberg 2009). Politicians and policymakers increased their attention to population control in response to concerns regarding global overpopulation and social unrest both at home and abroad among populations with high fertility rates. Research in the mid twentieth century from the emerging field of sociology demonstrated that unintended pregnancies resulted in increased rates of poverty and increased use of public resources.

In 1965, The National Academy of Sciences issued a report entitled "The Growth of U.S. Population" which stressed that medical and public health organizations should provide family planning services in order to enhance public health and to decrease poverty in the United States ("The growth of U.S. population," 1965). This report helped lay the groundwork for incorporating family planning into federal public health and anti-poverty programs.

In 1965, as part of President Lyndon Johnson's War on Poverty, the Office of Economic Opportunity (OEO) established the first Federal grants for family planning. President Richard Nixon furthered the federal role in the funding of family planning services by declaring to Congress in 1969 that "no American woman should be denied access to family planning assistance because of her economic condition" ("Message to Congress from President Richard M. Nixon, July 18, 1969," 1970). Shortly afterward, Congress enacted Title X of the Public Health Service Act of 1970. This established what is to this day the only federal program solely dedicated to the provision of funding of family planning services (Gold, 2001).

An Institute of Medicine review (2009) of the Title X program found that it provided funding for family planning services for almost five million clients. Title X is not an insurance program but rather provides fiscal resources each year to partially support the delivery of family planning services in family planning clinics and other sites across the U.S. While these sites receive public funding from several sources including Medicaid, state appropriations, the Maternal and Child Health Block Grant, and the Social Services Block Grant (Sonfield, 2003), Title X is the backbone of the family planning clinic system in the U.S. Based on surveillance data over time, the evidence suggests that the direct provision of funds through Title X increases women's access to contraceptive services (Sonfield, Alrich, & Gold, 2008). An analysis of U.S. trend data on publicly funded family planning services found that in 2001, Title X funding was associated with increased contraceptive access for low-income women in need of family planning services. Nationally, this represented an increase of 11% since 1994, meeting 28% of the national need for publicly-funded family planning services (Frost, Frohwirth, & Purcell, 2004). The number of clinics receiving Title X funding increased by 5% between 1994 and 2001, and the number of women receiving services in these clinics increased by 10%. Furthermore, in states that had Medicaid waivers (see below) in place, the capacity of Title X clinics to serve women in need of publicly funded care increased by 33% between 1994 and 2001. More recent data indicate that in 2006, Title X-supported clinics provided care to 66% of the 7.2 million women who received contraception from clinics receiving some form of public funding (Gold, Sonfield, Richards, & Frost, 2009). By providing infrastructure support, Title X helps to augment the benefits of other funding streams, including Medicaid (Frost et al., 2004).

The direct provision of contraception by publicly funded family planning clinics helps ensure that both health and cost benefits to the public are realized. A cost-benefit analysis of National Survey of Family Growth (NSFG) data including 7,643 women ages 15–44 years between 2002 and 2003, examined the impact of public funding of family planning clinics on unintended pregnancies and government cost savings (Frost, Finer, & Tapales, 2008). An estimated 1.4 million unintended pregnancies were averted, and there were savings of over four dollars for every public dollar spent (Frost et al., 2008). This is consistent with the known health and cost benefits of contraception; cost-savings increase as the efficacy of the contraceptive method increases (Sonnenberg et al., 2004; Trussell, 2007; Trussell et al., 1995). Similarly, a study conducted in four states compared family planning facilities that receive Title X funding to similar clinics that do not receive such funding and found that the Title X-funded clinics were more likely to prescribe emergency contraception, injectable contraception and other methods of contraception (Klerman, Johnson, Chang, Wright-Slaughter, & Goodman, 2007). These studies demonstrate that an investment in publicly-funded family planning clinics results in increased provision of services, and increased numbers of clients served and pregnancies averted, along with cost savings.

Public Reimbursement of Family Planning Services: Medicaid and Medicaid Family Planning Waiver Programs

Upon its creation in 1965, Medicaid allowed states to claim reimbursement for family planning services. Today, Medicaid coverage accounts for 61% of public funding of family planning services. The first expansions of Medicaid services were initiated in the 1980s; in 1989 Congress required states to expand Medicaid health care coverage to young children and pregnant women with family incomes up to 133% of the federal poverty level, and in 1990 mandated coverage of children under 100% of the federal poverty level. Beginning in 1993, states were permitted to apply for waivers to provide expanded family planning services to women who would not otherwise be eligible for Medicaid (Gold, Richard, Ranji, & Salganicoff, 2007). As demonstrated below, this has resulted in reduced rates of unintended pregnancy and related health costs for these states.

There are three main models of waiver programs: extending Medicaid post-partum coverage beyond 60 days to cover family planning services for 1–5 years (post-partum waiver), continuing Medicaid coverage of family planning services for women who leave the Medicaid system for other reasons (loss-of-coverage waiver), and extending Medicaid coverage of family planning services based on income alone (income-based waiver). At this time (August 2009), 27 states have obtained waivers for Medicaid coverage of family planning services using one of these three models. Twenty-one of these states use the income-based waiver approach (Guttmacher Institute, 2009a).

A review of the literature focused on evaluations of the impact of the Medicaid Family Planning waiver programs on access to family planning services yielded six studies (Table 4.1). Markers for increased access to family planning services included: enrollment in the waiver programs, increased availability of providers, use of family planning services, pregnancies averted, and state health care cost savings. In the first of these studies, Edwards and colleagues conducted a national evaluation of the 14 states that had Medicaid waivers between 1996 and 2001 (Edwards, Bronstein, & Adams, 2003). The authors evaluated the utilization of Medicaid family planning services in the six states with income-based waivers that had adequate resources to collect data for the evaluation. Between 1996 and 2001, the Medicaid income-based waiver program in these six states was associated with both an increase in the proportion of eligible women who enrolled in the program and in the proportion of enrolled women who used family planning services. In addition, five of the six states had an increase in the numbers of private physicians and non-Title X clinics providing family planning services. Furthermore, by looking at the availability of family planning service providers by population density, the evaluators found an increase in geographic availability of providers, especially in areas of the lowest population density. However, this increase in geographic availability was not directly associated with an increase in use of family planning services in all states, reflecting the fact that increased availability is important, but not necessarily sufficient to ensure increased uptake of services.

In a subsequent evaluation, Frost and colleagues (2004) used trend data on publicly funded clinics providing family planning services between 1994 and 2001 to examine the effect of the Medicaid family planning waiver (all types) on uptake of family planning services (Frost et al., 2004). The authors compared states with and without family planning waiver programs and also compared states with income-based waiver programs to states that had either post-partum or loss-of-coverage waiver programs. In 2001, there were 7,683 publicly funded family planning clinics in the United States serving 6.7 million women, which represented an 8% increase in the number of clinics and a 2% increase in the number of clients served since 1994. Between 1994 and 2001, states with income-based waiver programs had an overall 24% increase in the number of contraceptive clients served compared to a 2% decrease in states without waiver programs and an 8% decrease in states with the other types of family planning waiver programs. Frost and colleagues (2004) also assessed

Table 4.1 Evaluations of medicaid Waiver programs for family planning

Study Authors (years)	Population studied/sample size	Adjustment for potential confounders	Outcome measures	Key findings
Edwards, Bronstein and Adams (2003)	• States with medicaid family planning demonstration waiver programs • *N* = 14 states	The analyses were adjusted for the following variables: • Age • Race • Marital status • Family income • Recipient status regarding public aid • Parity • Tobacco use • Insurance pre-pregnancy • Alcohol use • Number of stressors • Abuse reporting	• Budget neutrality • Low-income women's access to family planning services • Unintended/mis-timed pregnancy rates • Changes in use of other Federal funding of family planning services	• All states were budget neutral • Use of family planning services increased in some, but not all states • Family planning waiver programs do seem to be associated with decreased rates of unintended pregnancy • Medicaid waiver programs were used by most states to partially substitute for other Federal family planning funding sources
Frost, Frohwirth and Purcell (2004)	• U.S. agencies and clinics providing publicly funded family planning services • *N* = 7,683	Analysis of data was adjusted for missing data, different reporting periods, and number of visits when number of clients served was unavailable	• Number of agencies providing publicly funded family planning services • Contraceptive clients served by publicly funded family planning clinics • Met contraceptive need • Clinic accessibility	Between 1994 and 2001 • 8% increase in number of publicly funded family planning clinics • 2% increase in number of clients served by publicly funded family planning clinics • States with income-based Medicaid family planning waiver programs had a 24% increase in number of clients served • States with post-partum or loss of coverage based Medicaid waiver plans had an 8% decrease in the number of contraceptive clients served • States without Medicaid family planning waiver programs had a 2% decrease in the number of contraceptive clients served

Foster, Biggs, Amaral, Brindis, Navarro, Bradsberry and Stewart (2006)	• Female Family PACT clients in 2002 • $N = 926,218$ women	Markov model was used to estimate pregnancies averted each month of contraceptive coverage, using assumptions of contraceptive use, contraceptive efficacy, risk of pregnancy by age, and pregnancy outcome. Sensitivity analyses examined effect of changes in above assumptions on estimates	• Pregnancies averted	Model estimated number averted in 2002 • 205,000 pregnancies • 79,000 induced abortions • 94,000 births • 21,400 adolescent pregnancies
Bronstein, Vosel, George, Freeman and Payne (2007)	• Alabama Title X and Medicaid Family Planning expansion program clients • $N = 170,122$ clients served (2000–2004)	Likelihood of returning for subsequent family planning visit adjusted by: • Care coordination • Performance of risk assessment tool • Type of provider • Post-partum status • Age • Race • Area of residence within the state	Female clients served by both programs over time • Total number • Demographic characteristics • Services used • Return visits for family planning services	• 30% increase in number of women receiving public funding for family planning between 2000–2004 • Greatest expansion in clients served was in Medicaid family planning clients using non-Title X providers
Lindrooth and McCullough (2007)	• State level data • $N = 50$ states	Trends in fertility rates by states were controlled for: • Age • Gender • Unemployment • Employment status of women • Race/Ethnicity • Annual population growth rates States with expansion programs were pooled with comparison states within the region and dummy variables were used to control for regional and national variation of trends in fertility rates	• Births per 1,000 women of child-bearing age • Program net costs to states	At state level • Income-based and postpartum-based family planning expansion programs were either budget neutral or cost-effective At national level • Family planning expansion programs are budget neutral or cost-effective except for California • Income-based family planning expansion programs were more effective than postpartum-based programs at reducing birth rates

(continued)

Table 4.1 (continued)

Study Authors (years)	Population studied/sample size	Adjustment for potential confounders	Outcome measures	Key findings
Kearney and Levine (2009)	State level data aimed at women eligible for Medicaid family planning waiver programs	Regression models controlled for: • Population subgroup prevalence by state • Changes to state abortion restrictions • Welfare benefit changes • Medicaid policy changes • State mandates requiring insurance coverage of contraception • State unemployment rates The estimates were modeled by: • Age group (teen versus adult) • Race and ethnicity • Education level	• Rates of sexual activity • Rates of contraceptive use • Overall birth rates • Overall induced abortion rates • Birth rates of program eligible women • Adolescent birth rates • Adolescent induced abortion rates	Income-based Medicaid waiver programs • Reduce overall births by 2% • Reduce adolescent births by >4% • Provided family planning services for an additional 22.5% of eligible women • Decrease birth rates due to increased contraceptive use

unmet need for publicly-funded contraceptive services (i.e., the proportion of low-income women who are at risk of unintended pregnancy and not using contraception) in states with income-based waiver programs, other types of family planning waiver programs, and no waiver programs. Prior to the introduction of waiver programs, the unmet contraceptive need was similar among states that did and did not adopt waiver programs (when this became an option for states). States that adopted income-based waiver programs had a 27% increase in met contraceptive need from 1994 to 2001, rising from 39 to 50%. In states with the other types of waiver programs, the met contraceptive need decreased from 38 to 34% due to a decrease in the number of women served in these states during that time. In the same time period, states without waiver programs had a constant unmet contraceptive need of 40%. Thus, the income-based Medicaid family planning waiver programs proved to be the most successful intervention, increasing both access and uptake of family planning services.

Foster and colleagues used data from California to evaluate the impact of the state's Medicaid waiver program on pregnancy rates and state cost-savings (Foster et al., 2006). The California Family Planning, Access, Care and Treatment program (Family PACT) is an income-based Medicaid waiver program with eligibility extended to both men and women up to 200% of the federal poverty level. In its first year (1997), Family PACT served more than 750,000 clients. By 2003, the number of clients enrolled in the program increased to more than 1.5 million California residents. In that year, nearly one million women received contraception and the number of participating providers reached almost 3,000. Foster and colleagues (2006) estimated the number of unintended pregnancies averted with Family PACT by comparing participants' contraceptive methods before and after entry into the program. They found that 6.4 million woman-months of contraception were provided in 2002 through the Family PACT program, resulting in 205,000 unintended pregnancies averted. Assuming that women would not use any contraception at all if they were not enrolled in Family PACT triples the estimate of the number of pregnancies averted. Based on the cost of contraception versus the cost of covering pregnancy healthcare services, it is estimated that this program saved California $1.1 billion dollars over 2 years. The researchers make the important point that the "ability of a program to prevent pregnancy lies primarily in its provision of contraceptive methods to women who would otherwise not use them. Enabling women to switch from methods with relatively high failure rates (such as condoms) to more effective methods is also an important, albeit less powerful, factor" (Foster et al., 2006). In short, by allowing women and couples to have access to their preferred method of contraception and by removing the barrier of cost, Medicaid family planning expansion programs, like all coverage efforts, enable individuals and couples to determine which method will be most effective for them and provide the means to procure this method.

Foster and colleagues (2006) also used Family PACT data to determine the cost-effectiveness of and number of pregnancies averted for each contraceptive method covered by Family PACT in 2003. The cost-savings per pregnancy averted were calculated by estimating the cost to the public of an unintended pregnancy for 2 years after birth (Foster et al., 2006). The authors estimated that 205,000 pregnancies were averted overall due to the contraceptive provision by Family PACT. All methods were found to be cost-effective, and the overall average cost savings per dollar expenditure was $3.52. Oral contraceptives and injectables accounted for the greatest number of pregnancies averted (91,000 and 39,000 respectively), and the implant and the intrauterine device provided the most cost savings per dollar expenditure ($15.90 and $7.24, respectively) (Foster, Rostovtseva, Brindis, Biggs, Hulett, & Darney, 2009).

In 2007, Lindrooth and McCullough performed a subsequent assessment of the cost-effectiveness of Medicaid waiver programs by comparing birth rate data (births per 1,000 women of reproductive age) in states with and without Medicaid family planning waiver programs. Between 1991 and 2001, birth rates decreased on average by 1.95% points in states with income-based waivers, and by 0.87% points in states with postpartum-based waivers. Because the federal government match rate for Medicaid family planning is 90% (compared to 50–83% for infant and maternal health services),

the financial benefit of family planning waiver programs to the states is significant. According to Lindrooth and McCullough, the savings in Medicaid maternal and infant healthcare costs per avoided unintended birth exceeded the total family planning waiver program costs in the majority of states with Medicaid waivers. This finding was most striking in states with income-based family planning waiver programs (Lindrooth & McCullough, 2007).

Bronstein and colleagues (2007) evaluated the impact of the first 4 years of Alabama's income-based Medicaid Family Planning waiver program. The authors compared Medicaid claims and enrollment data pre and post-expansion as well as Title X data in order to determine whether the Medicaid Family Planning expansion program served a different population than Title X. The authors found a 30% increase in the number of women receiving public funding for family planning services as a result of the Medicaid Family Planning program. The greatest expansion in new enrollees was in women who used non-Title X providers. Thus, the Medicaid Family Planning program in Alabama was found to expand the population of women receiving public funding for family planning services in that state (Bronstein, Vosel, George, Freeman, & Payne, 2007).

In the most recent evaluation of Medicaid Family Planning waiver programs found through this review, Kearney and Levine (2009) evaluated the impact of these programs on birth rates and contraceptive use. The authors compared state birth rates in the 4 years prior and 2 years following the waiver implementation. They found that income-based waiver programs were the most effective, resulting in a 2% decline in births to adult women and a 4% decline in births to teenage women. They also demonstrated that the income-based waivers resulted in the provision of family planning services to an additional 22.5% of women eligible for the program. By using the baseline birth rates in states prior to the implementation of the family planning waiver programs and the estimated birth reduction due to the implementation of the programs, Kearney and Levine concluded that one unintended birth was avoided for every additional 36 Medicaid family planning waiver recipients. Therefore, Medicaid family planning waiver programs, especially income-based programs, were associated with decreased unintended birth rates and with healthcare cost-savings to the states.

In summary, six studies to date have evaluated the impact of Medicaid Family Planning waiver programs on women's access to and use of family planning services. The authors of these studies used several different outcomes as markers for access to family planning services: enrollment in waiver programs, use of contraception, pregnancy rates, unintended pregnancy rates, and cost-savings. Across these different markers, the authors consistently found that Family Planning waiver programs improve access to family planning services, improve women's use of contraception and decrease unintended births.

Contraceptive Insurance Mandates and Other Approaches to Increasing Private Coverage for Family Planning

Ten percent of women ages 18–64 years have Medicaid coverage and 18% are uninsured, meaning that the majority of reproductive age women are privately insured (Kaiser Family Foundation. *Women's Health Insurance Coverage. Women's Health Policy Facts*, October 2009). While most private insurance plans include prescription drug coverage, these plans may not cover contraceptive drugs or devices ("Insurance Coverage of Contraceptives," 2009). Lack of private insurance coverage for contraception has historically resulted in decreased family planning access for a large number of reproductive aged women. As such, contraceptive coverage mandates are an important intervention to increase women's access to family planning services. The first enacted contraceptive coverage mandate was in Maryland in 1998 (Dailard, 2004). Over the next 10 years, the number of states with contraceptive coverage mandates increased to 27, but provisions for "conscience clauses" have allowed

many employers and insurance companies to be excused from compliance with state mandates (Dailard, 2004). There are two other types of mandates aimed at increasing private insurance coverage of contraceptive services. In 1998, Congress passed a contraceptive coverage mandate for the Federal Employees Health Benefits Program which covers 1.2 million women of reproductive age who work for the federal government ("Contraceptive Coverage in the Federal Employees Health Benefits Program," 2003). The judicial system has also played an important role. In 2000, the Equal Employment Opportunity Commission and a Seattle federal court both ruled that employer-based insurance plans that excluded contraception coverage violated Title VII of the Civil Rights Act of 1964. As a result, employers who employ 15 or more persons and do not cover contraception in their insurance plans are vulnerable to litigation (Dailard, 2004).

Our review of the literature found only one study that evaluated the impact of contraceptive mandates on women's access to contraception. Sonfield and colleagues (2004) analyzed data from a national sample of insurance companies surveyed about coverage of reproductive health care services in employment-based managed care plans in 1993 and 2001–2002. According to the authors, between 1993 and 2002, the percentage of private insurance companies surveyed that provided coverage for the full range of reversible prescription methods of contraception increased dramatically from 28 to 86%. Though several factors may have contributed to this change, contraceptive mandates likely played an important role (Sonfield, Gold, Frost, & Darroch, 2004). The influence of contraceptive coverage mandates at the state level extends to states without mandates because insurance companies must provide contraceptive coverage consistent with the mandate of the state in which the insured participate in the plan, even if that is not the state in which the company is based (Sonfield et al., 2004).

Sonfield et al. found that the proportion of insurance plans covering each contraceptive method was higher in 2002 (78–97%) than in 1993 (32–59%). Insurance plans in states with mandates had a significantly higher rate of coverage of the five leading contraceptive methods than did plans limited to states without contraceptive mandates (87–92% vs. 47–61%). Additionally, the authors determined that between 1993 and 2002, state mandates were responsible for 30% of the national increase in coverage of oral contraceptives and 40% of the increase in coverage of the 3-month injectable contraceptive. Importantly, in addition to increasing overall coverage, state mandates increased the number of contraceptive methods available to women and their partners. By making more methods available, these mandates likely increase access of women and couples to the more effective contraceptive methods. Given these findings, the authors argue for going beyond state mandates towards a federal contraceptive coverage mandate to increase access to and effectiveness of family planning services across the U.S. (Sonfield et al., 2004).

Access to Safe and Legal Abortion Services

Background

Access to comprehensive family planning services – including induced abortion – is advantageous for maternal and child health outcomes if those services are legal, affordable, and available to the population at large. In the U.S., access to safe abortion (interpreted as part of a woman's right to privacy) has been legal in all states since the landmark Roe v. Wade decision in 1973. However, there are significant disparities in access to contraception, rates of unwanted pregnancy, insurance coverage for induced abortion and state differences in legal access, all of which lead to differential access to safe abortion care (Joyce & Kaestner, 1996).

Medicaid coverage for induced abortion is restricted by the Hyde Amendment, first passed in 1976, which banned the use of federal funding for induced abortion. Currently, the amendment allows federal funding for induced abortion in specific cases of pregnancy resulting from rape or incest or in the case of life endangerment for the pregnant woman. As a result of these restrictions, responsibility for public funding has fallen to the states; 17 currently use state funding to pay for all or most medically-indicated induced abortions for Medicaid recipients (Boonstra, 2007). Thirty-two states and the District of Columbia provide Medicaid funding in cases of life endangerment, rape, and incest – consistent with Federal guidelines ("State Funding of Abortion Under Medicaid," 2009). South Dakota appears to violate the federal guidelines by providing Medicaid funding for abortion only in cases of life endangerment.

As a result of Medicaid abortion funding restrictions, the federal government funds fewer than 1% of publicly funded abortions (Sonfield, Alrich, & Gold, 2008). This rate of federal funding has been consistent since the institution of the Hyde Amendment (Nestor & Gold, 1984; Gold & Nestor, 1985; Gold & Macias, 1986; Gold & Guardado, 1988; Gold & Daley, 1991; Daley & Gold, 1993; Sollom, Gold & Saul, 1996).

The Effect of Public Funding on Access to Safe Abortion

Because women who are eligible for Medicaid are among the poorest of American women, they are less able to pay for abortion on their own. In the U.S., restrictions on public funding of abortion have important implications for both induced abortion rates and birth rates. A lack of resources can result in women either delaying an induced abortion or having an unwanted birth (Finer et al., 2006). A previous review of the literature on this topic by Henshaw et al. identified 38 studies evaluating the impact of Medicaid funding restrictions on a range of outcomes. (Henshaw, Joyce, Dennis, Finer, & Blanchard, 2009). Our review of the literature focused specifically on studies evaluating the impact of public funding status on access to abortion as measured by: abortion rate, gestational age at the time of abortion, unmet need for abortion, and availability of providers. Of the 22 studies included in this review, 16 studies found that Medicaid funding is clearly associated with access to abortion and six studies found no significant relationship between funding restrictions and abortion access/uptake.

Seven studies evaluated the impact of Medicaid restrictions and availability of abortion close to the implementation of the Hyde Amendment. A 1979 study by Cates and colleagues used prospectively collected CDC surveillance data from October 1977 to June 1978 to evaluate the health impact of Medicaid abortion funding restrictions. By comparing data in states with and without Medicaid funding of abortions, the authors found that women who were eligible for Medicaid had their abortions 2.4 weeks later in gestational age than women of higher socioeconomic status in states that did not provide Medicaid funding for abortion. In contrast, states that did provide Medicaid funding for abortion did not have a significant difference in gestational age between Medicaid eligible and non-Medicaid eligible women (Cates et al., 1979). Trussel and colleagues (1980) examined abortion rates in Ohio and Georgia before and after the implementation of Medicaid abortion funding restrictions in those states. The authors found that 18 and 25% of women in Ohio and Georgia respectively did not have abortions as a result of the funding restrictions. Furthermore, the authors used survey data in Ohio to demonstrate that funding restrictions resulted in a 3 day delay in obtaining an abortion (Trussell, Menken, Lindheim, & Vaughan, 1980). Chrissman and colleagues (1980) looked at state level data in Texas in 1980, in order to evaluate the impact of Medicaid funding restrictions that were implemented in 1977. They found that 35% of Medicaid-eligible women who would have used Medicaid funding prior to the restrictions continued their unintended pregnancies. Additionally,

the fertility rate amongst Medicaid eligible women increased 4.2% from 1976 to 1977 and 12% from 1977 to 1978, which was largely attributed by the authors to Texas abortion funding restrictions (Chrissman et al., 1980).

Gold (1980) found the unmet need for abortion rose from 31% in 1977 to 55% in 1978 after implementation of the Hyde Amendment; in 1977 294,600 indigent women received Medicaid funding for abortion compared to 193,800 women in 1978 (Gold, 1980). In 1984, Henshaw and Wallisch compared data from abortion clinics prior to 1977 to data in 1982 and found an average delay in obtaining an abortion of 0.43 weeks after introduction of Medicaid funding restrictions for Medicaid-eligible women. In 1982, 50% of Medicaid-eligible women had their terminations after 10 weeks gestational age, compared to 37% prior to 1977 (Henshaw & Wallisch, 1984). Two studies evaluated Medicaid funding for induced abortion. Singh used 1980 state level data to evaluate factors impacting adolescent pregnancy rates, and found that the availability of Medicaid funding for induced abortion was associated with decreased birth rates in black adolescents, and with an increase in induced abortion rates and ratios in both black and white adolescents (Singh, 1986). Medoff also used 1980 state level data in an economic model which analyzed factors influencing the demand for abortions and found that state Medicaid funding for abortion was associated with 44 additional abortions per 1,000 pregnancies compared to states without such funding (Medoff, 1988).

Nine subsequent studies using data from the mid- 1980s to the present demonstrated consistent findings regarding the impact of Medicaid funding restrictions on abortion access and uptake. Korenbrot and colleagues used state level data to compare abortion rates in Colorado, North Carolina, and Pennsylvania before and after the states introduced Medicaid restrictions on abortion funding in 1985; they found these restrictions to be associated with a 1.2% increase in pregnancies resulting in live births (compared to a 0.4% nationally) (Korenbrot, Brindis, & Priddy, 1990). In 1990, Lundberg and Plotnick evaluated the impact of welfare, abortion, and family planning policies on the decision-making of adolescent white females using data from the National Longitudinal Survey of Youth from 1979 to 1986 and found that policies restricting public funding of abortion resulted in decreased rates of abortion in this population (Lundberg & Plotnick, 1990).

Also in 1993, Haas-Wilson evaluated the impact of policies restricting abortion funding and parental involvement laws on abortion rates and the availability of abortion providers. They found that states with Medicaid funding for abortion had higher abortion rates in general, higher abortion rates among minors, and increased availability of abortion providers compared to states with Medicaid funding restrictions (Haas-Wilson, 1993). Meier and McFarlane (1994) conducted a pooled time series looking at the impact of family planning expenditures and funding on several health outcomes including abortion rates, using data from all U.S. states from 1982 to 1988 and found that public funding of abortion resulted in increased numbers of induced abortions, along with decreases in rates of births to teenaged women (Meier & McFarlane, 1994). In 1996, Blank and colleagues looked at the impact of policies including Medicaid funding restrictions on abortion rates in the 50 states and the District of Columbia between 1974 and 1988. They found Medicaid restrictions to be associated with a 13% decline in state abortion rates. The authors also found that abortion rates of Medicaid-eligible women decreased 19–25% (Blank, George, & London, 1996).

Cook and colleagues (1999) evaluated the effects of changes in public funding of abortion on both induced abortion rates and birth rates in North Carolina. At that time, North Carolina used a separate fund for public financing of abortion which fluctuated with legislative appropriations (Cook, Parnell, Moore, & Pagnini, 1999). In an analysis of the month-to-month induced abortion and birth rates, the authors determined that insufficient public funding for abortion resulted in a situation in which 37% of women who would have otherwise obtained an abortion continued the unwanted pregnancy to term. Women who were most affected were African-American, aged 18 years and older, and women with less than a high school education. The authors further noted

that the women who continued their pregnancies as a result of these decreases in funding were women eligible to receive public assistance for food, housing, and medical assistance as a result of the unintended birth. In other words, cuts in public funding for abortion in North Carolina were associated with decreased access to reproductive health services; this in turn was reflected in a decrease in the number of women accessing induced abortion care, an increase in unintended births, and an overall increase in public spending (Cook et al., 1999). In 2002, Morgan and Parnell confirmed Cook et al.'s findings in a subsequent evaluation of the effect of limits in the North Carolina State Abortion Fund on abortion rates and birth rates between 1988 and 1995 (Morgan & Parnell, 2002).

A 2007 study conducted by New for the Heritage Foundation looking at the effect of state level legislation on the incidence of teenage abortion rates found that restrictions on Medicaid abortion funding were associated with a decline in the abortion rates among girls 13–17 years old (New, 2007). Finally, in 2008, Foster and colleagues analyzed data from 398 women seeking first and second trimester induced abortions in an urban hospital in San Francisco between September 2001 and March 2002 (Foster, Jackson, Cosby, Weitz, Darney, & Drey, 2008). The authors found that difficulty in obtaining coverage from MediCal (California's Medicaid program) was associated with a delay between diagnosis of pregnancy and first contact with an abortion clinic. Additionally, difficulty financing an abortion was associated with a delay between the first contact with an abortion clinic and obtaining an abortion procedure (Foster et al., 2008).

In contrast to the above studies, six other evaluations of the effect of restrictions in Medicaid funding found no significant impact of these restrictions on abortion access and uptake. In 1979, Rubin and colleagues. conducted a survey of abortion providers in a large metropolitan area in Texas before and after Texas withdrew Medicaid funding for abortion in 1977. The authors found that the number of abortions performed in this metropolitan area actually increased in the year after the Medicaid abortion funding restriction was implemented. In examining the largest public hospital in the area, the authors found no change in birth patterns during this time period. However, they note that they were unable to determine the rates of unintended births or illegal terminations during this period (Rubin, Gold, & Cates, 1979). A 1995 study by Wetstein using a time-series model to evaluate the impact of national policy changes on abortion rates found that the passing of the Hyde Amendment was not associated with a significant change in national abortion rate trends (Wetstein, 1995). In 1996, Kane and Staiger used national data at the county-level to conduct an empirical analysis of the impact of factors affecting abortion access on adolescent birthrates between 1973 and 1988. The authors found that Medicaid abortion restrictions were associated with a decrease in adolescent birth rates. However, the authors highlighted that the decline in adolescent birth rates had started prior to the funding restrictions (Kane & Staiger, 1996). Haas-Wilson (1997) using data from 1978 to 1992 and controlling for demographic characteristics and the availability of abortion providers, found that Medicaid funding restrictions did not have significant impacts on abortion or birthrates in states with the restrictions (Haas-Wilson, 1997).

A 1996 study of state level data from 1977 to 1988 by Levine et al. found that Medicaid restrictions resulted in a decrease of 2 births per 1,000 women aged 15–44 in states with restrictions. However, when controlling for state-level trends, this difference was no longer present (Levine, Trainor, & Zimmerman, 1996). Finally, in 2001, Bitler and Zavodny used CDC data to compare gestational age at the time of abortion in states with and without Medicaid abortion restrictions and parental involvement laws. The authors found no significant difference in the gestational age at the time of abortion in states with and without Medicaid funding restrictions (Bitler & Zavodny, 2001).

Access to Abortion and Effect on Maternal and Infant Outcomes

State funding of abortion has been associated with improved maternal and perinatal outcomes. First and foremost, there is a direct relationship between access to legal induced abortion and reduced maternal mortality rates due to a decrease in the numbers of abortion-related deaths that accompanies legalization ("Trends in Abortion in the United States, 1973–2005" 2008). Abortion-related maternal mortality began improving dramatically pre-Roe v. Wade as 15 states had already liberalized abortion laws by 1970 and continued to improve post-Roe v. Wade until today (Strauss et al., 2007; Tietze, 1975; "Vital Statistics of the United States," 1969; "Vital Statistics of the United States, 1965," 1967; "Vital Statistics of the United States, 1966," 1968). While positive maternal mortality outcomes are readily associated with improved access to abortion services, the legalization of abortion has also been demonstrated to lead to better infant and child health outcomes.

We evaluated three population level studies that examine the effect of abortion on maternal and child outcomes (Table 4.2). Gruber and colleagues conducted a national ecologic cohort study examining the effect of change in legal access to induced abortion by state on the living standards of children by cohort. They found that after induced abortion was legally available, children were less likely to live in single-parent households and thus less likely to live in poverty, less likely to receive welfare, and experienced lower infant mortality rates (Gruber, Levine, & Staiger, 1999). Meier and McFarlane (1994) found that public funding of abortion was associated with improvements in prenatal care and decreases in low-birth weight and premature births. In contrast, Currie et al. found little evidence for a direct association between state restrictions on public funding of abortion and birth weight or incidence of low birth weight (Currie, Nixon, & Cole, 1996).

Findings of a positive association between abortion access and improved maternal and child outcomes at the *population* level can seem difficult to reconcile with studies (including two meta-analyses) showing an association at the *individual* level between a history of prior induced abortion and an increased risk of preterm birth in subsequent pregnancies (Ancel, Lelong, Papiernik, Saurel-Cubizolles, & Kaminski, 2004; Brown, Adera, & Masho, 2008; Moreau et al., 2005; Shah & Zao, 2009; Swingle, Colaizy, Zimmerman, & Morriss, 2009). Proposed mechanisms for the relationship include infection, mechanical trauma to the cervix, and uterine scarring following curettage. It is important to note that when the breadth of literature on this issue is examined, one sees conflicting and inconsistent findings, with some studies detecting a significant association for any prior abortion, others showing an association only for multiple abortions, and others finding no significant association between prior induced abortion and the risk of preterm delivery and low birth weight.

A number of explanations for inconsistencies between studies have been proposed. It has been pointed out that several studies that demonstrated an association between prior induced abortion and adverse pregnancy outcomes were carried out in Europe (Ancel et al., 2004; Moreau et al., 2005), whereas studies carried out in the U.S. are less likely to find a significant association (Mandelson, Maden, & Daling, 1992; El-Bastawissi, Sorensen, Akafomo, Frederick, Xiao, & Williams, 2003; Chasen, Kalish, Gupta, Kaufman, & Chervenak, 2005). Study location may be relevant because procedures involving curettage are more common in Europe, while procedures involving vacuum aspiration are more common in the U.S., and the latter may carry less risk of trauma to the uterine lining (Ancel et al., 2004). In some studies (Brown et al., 2008) induced abortions are not distinguished from spontaneous abortion. Another proposed explanation for conflicting results is variation in controlling for potential confounders. For example, Mandelson et al. (1992) found a strong association between history of multiple abortions and risk factors for low birth weight, such as age and smoking. Once confounders were controlled for, the most significant associations between abortion and preterm birth disappeared. Chasen et al. (2005) found that women who experienced spontaneous preterm birth after surgical abortion at or beyond 20 weeks of gestation tended to

be women who underwent abortion because of premature cervical dilation and/or rupture of membranes, or women with a multiple gestation in the post-abortion pregnancy. Although not adjusted for, these three conditions act as confounders as they are recurrent risk factors that place women at an increased risk of preterm delivery regardless of having had a surgical abortion.

Virtually all the studies examining the association between abortion and preterm birth/low birth weight at the individual level have been retrospective case-control designs or cross-sectional designs with retrospective ascertainment of prior induced abortion. Given the threats to validity inherent in such studies, a definitive answer to the question of whether induced abortion puts women at risk for future adverse pregnancy outcomes will likely have to await the fielding of prospective cohort studies that include medical as well as surgical abortion, and that collect accurate information on such factors as the indication for the induced abortions and a variety of potential confounders.

Training Providers to Deliver Comprehensive Reproductive Health Care as a Strategy to Increase Access to Family Planning and Abortion Services

Training providers in family planning and abortion provision is essential to ensuring women's access to comprehensive family planning. Over the course of the past 40 years, trends in the number of clinicians trained in comprehensive family planning provision have shifted as federal, state, and local governments' and accreditation organizations' policies have changed. These policies – from the legalization of induced abortion in 1973 to the inclusion of induced abortion training as a requirement within U.S. Obstetrics and Gynecology (OB-GYN) residency training programs in 1996 – have changed the opportunities for clinical training in the provision of comprehensive family planning services.

A review of the evidence revealed no studies which directly examine whether training providers in comprehensive family planning provision translates into better access for women seeking family planning services. However, the need for training future providers remains evident. In particular, a major concern in the U.S. with respect to the ability to offer comprehensive family planning services to women in need is that the number of abortion providers has been declining. Data illustrate substantial drops in the number of abortion providers between 1982 and 1984 (7%) and between 1992 and 1995 (12%); they also show slowing, but nevertheless continuous drops between 2000 and 2005 (2%) (Jones, Khost, Singh, Henshaw, & Finer, 2009). Additionally, when looking at the number of trained abortion providers from a geographical perspective, 87% of U.S. counties in 2005 lacked an abortion provider (Guttmacher Institute, 2009b).

Historically, training opportunities in induced abortion provision were not readily incorporated into OB-GYN residency programs until the legalization of abortion in 1973 (Lindheim & Cotterill, 1978). While there was an initial increase in the number of physicians trained to perform induced abortions after legalization, declining trends in abortion training opportunities were detected as early as 1995 (Mackay & Mackay, 1995). Over the course of merely 18 years, the percentage of Obstetrics and Gynecology (OB-GYN) residency programs requiring exposure and training in first-trimester induced abortions – the majority of gynecological procedures performed – fell by more than half, from 26% in 1978 to only 12% of residency programs providing regular abortion care in 1995 (Lindheim et al., 1978; Mackay et al., 1995). In response to this decline in abortion training opportunities, the American College of Graduate Medical Education (ACGME) mandated the inclusion of induced abortion training into standard OB-GYN residency education beginning January 1996 (Almeling, Tews, & Dudley, 2000).

The most recent survey of U.S. OB-GYN residency program directors (conducted in 2004) found two models of abortion training: "opt out" programs (51%) required routine instruction in current induced abortion procedures unless an individual resident expressed moral obligations to learning induced abortion techniques, while "opt in" programs (39%) made abortion training optional.

Table 4.2 The association between induced abortion and maternal and infant outcomes at the population level

Study authors (years)	Population studied/ sample size	Adjustment for potential confounders	Outcome measures	Key findings
Meier and McFarlane (1994)	– U.S. national population	The effect of funding of abortion services on state abortion rates was controlled for the following variables: – Black and Latino population percentages – Percentage of women in labor force – Per capita income – Proportion of Catholics in each state – Percentage of state population living in counties with large volume abortion providers	– Abortion rates – Teenage birthrates – Low birth weight – Premature births – Late prenatal care – Infant mortality – Neonatal mortality	States that funded abortion had significantly: – Higher abortion rates – Lower teenage pregnancy rates – Lower rates of low birth weight babies – Lower premature birth rates – Lower rates of births with late or no prenatal care States with higher spending on family planning had significantly: – Fewer abortions – Lower rates of low birth weight babies – Lower rates of births with late or no prenatal care – Lower infant mortality rates – Lower neonatal mortality rates
Currie, Nixon and Cole (1996)	– Women enrolled in the National Longitudinal Survey of Youth. – $N = 6,283$	Model controlled for: – Abortion laws – Provider availability – Rate of unmarried births – Prevalence of subsidized care clinics – Birth rate – Birth weight – Low birth weight – Income – Age – Education level – Obstetric history – Marital status – Height – Primary family size – Primary language – Employment status – Single parent households	– Probability of birth – Birth weight – Low birth weight	– Restrictions on Medicaid funding of abortion increase the probability that pregnancy is carried to term. This increase is greater among African-American women than white women, and among low-income women compared with wealthier women. – No significant effect of restrictions on Medicaid funding on birth weight or rate of low birth weight. – Observed relationships between restrictive abortion laws and birth weight likely reflect state characteristics that are correlated with passage of such laws, rather than true effects of restrictions.

(continued)

Table 4.2 (continued)

Study authors (years)	Population studied/ sample size	Adjustment for potential confounders	Outcome measures	Key findings
		– Education level of respondent's mother – Religious affiliation and attendance – Geographic location – Urban vs. rural residence – Gender – Alcohol use – Tobacco use		
Gruber, Levine and Staiger (1999)	– All non-institutionalized children in the 1980 census born between 1965 and 1979. – $N = 2.4$ million	The effect of abortion access on the living circumstances of children was controlled for the following variables: – Insured unemployment rate – Per capita income – Crime rates – Percent non-white population The effect of abortion access on the living circumstances of children was controlled using the following strategies: – State and year dummies – State-specific quadratic trends	– Poverty – Single-parent household – Household receiving welfare benefits – Birth weight – Infant mortality	Legalization of abortion reduced the percentage of children living in: – Poverty – Single-parent household – Household receiving welfare benefits No significant reduction in: – Birth rate – Low birth weight – Infant mortality rate

Ten percent of programs offered no induced abortion training. "Opt out" programs were more likely than "opt in" programs to train greater than 50% of their residents in the following procedures: first trimester surgical abortion (85% vs. 21%, $p < 0.001$); first trimester medication abortions (59% vs. 28%, $p < 0.001$); second trimester inductions (51% vs. 31%, $p = 0.001$); and second trimester surgical procedures or "D&E" (40% vs. 14%, $p < 0.001$) (Eastwood, Kacmar, Steinauer, Weitzen, & Boardman, 2006).

Exposure to induced abortion training during residency is not the only predictor of whether an OB-GYN will provide induced abortion services. In a survey of OB-GYN graduates of five residency programs from 1989 to 1998, current and past provision of induced abortion care was examined with respect to the demographic, practice and training characteristics of the clinician (Steinauer, Landy, Jackson, & Darney, 2003). Of the 161 respondents, 83% participated in induced abortion training in their residency. Although the majority (58%) had performed induced abortions at some time after they completed their training, fewer than half (47%) had performed a first trimester abortion procedure in the preceding year, and even fewer (26%) had performed a second trimester abortion procedure in the same time period. In addition to offering abortion training in residency, providing abortion training in the hospital setting and providing experience with a high volume of procedures were all independent predictors of current provision of induced abortion care. In other words, training clinicians in how to perform abortions as well as the setting and the clinical volume of the training experience are all important predictors of future abortion provision.

Physicians trained in specialties other than OB-GYN, and other clinicians who provide reproductive healthcare for women can also provide induced abortion care, but training in abortion is less common in fields outside of OB-GYN. Family Medicine is a specialty that does routinely train physicians in other aspects of reproductive health beyond family planning services. In a 1995 survey of Family Medicine program directors and chief residents in 422 training programs, both the interest and need for training in comprehensive family planning care including abortion provision was queried (Steinauer, Depineres, Robert, Westfall, & Darney, 1997). Approximately half (244) of the programs responded, with 197 program directors and 152 chief residents completing surveys. The majority of respondents reported no training in contraceptive procedures such as diaphragm fitting or intrauterine device placement. Not surprisingly, 85% of program directors and chief residents together reported no clinical experience in first trimester induced abortion in their training program. Although only 37% of chief residents felt that first trimester induced abortions are appropriate in family practice, the majority (64%) agreed that training in these procedures should be optional in Family Medicine training programs. Although these data are almost a decade old, the disparate responses may represent some of the complexity of practitioners' personal beliefs about induced abortion provision. In one Family Medicine training program in which a biopsychosocial model was applied to induced abortion care training, the training included individual consideration of personal and professional aspects of abortion provision (Gawinski, Bennett, Rousseau, & Schaff, 2002). The authors concluded that this model of training might increase the number of trained abortion providers by allowing trainees to thoughtfully consider their decision to become an abortion provider and once the decision is made, to fully commit to their training and later provision.

There is evidence that other aspects of medical training and practice may contribute to the availability of induced abortion care. In a survey of graduates of five OB-GYN residency programs between 1989 and 1998, Obstetrician-gynecologists who currently provided induced abortion services were more likely to practice in an urban setting and less likely to practice in a group with restrictions against induced abortion provision (Steinauer et al., 2003). In their earlier survey, Steinhauer et al. (1997) found that chief residents' responses regarding the appropriateness of induced abortion training or plans to provide induced abortion care were correlated with the region of the country in which the training program was situated. In fact, the number of trained and practicing abortion providers does vary by state.

Training clinicians in the delivery of comprehensive reproductive health services is an essential step in ensuring that there are a sufficient number of providers to deliver such services but as demonstrated above, this is a necessary but not sufficient approach to increasing access. As such, privately funded programs such as the Kenneth J. Ryan Residency Training Program in Abortion and Family Planning have been initiated to expand training opportunities to physicians in training. Ensuring that an adequate comprehensive family planning delivery system is in place will require considering a broader approach to training; increasing the number of residency programs in which comprehensive family planning training is incorporated, expanding the types of medical providers who receive such training, and expanding the role of non-physicians such as nurse practitioners who already provide family planning services to enable these individuals to provide some abortion services.

Conclusion

Most women expect and desire pregnancy in their lifetime (Abma, Chandra, Mosher, Peterson, & Piccinino, 1997). Thus the goal of fertility regulation is not simply to prevent pregnancy, but to prevent pregnancies that women desire to avoid. It is well-established that family planning and safe abortion services prevent maternal and infant morbidity and mortality. Simply put, when comprehensive family planning needs are met, the health and welfare of women and their children are improved. This chapter reviewed several interventions aimed at improving access to family planning and abortion services and identified several essential factors necessary for effective and comprehensive family planning delivery including: the availability of public and private funding of family planning and abortion services, and the availability of trained providers.

Unintended pregnancy affects all women of reproductive potential. U.S. unintended pregnancy rates have remained at endemically high rates for women regardless of age, race, ethnicity and socio-economic status, and overall rates of contraceptive use, unintended pregnancies and unintended births among groups are more similar than they are different (Abma et al., 1997; Henshaw, 1998; Ranjit, Bankole, Darroch, & Singh, 2001; Mosher, Martinez, Chandra, Abma, & Willson, 2004). However, women disadvantaged by poverty or racial/ethnic disparity face special health, social and economic risks when they experience an unintended pregnancy. The interventions evaluated in this chapter rarely focused on any particular population group other than women disadvantaged by poverty.

Despite the similarities among groups regarding contraceptive use and unintended pregnancies and births, the demonstrated disparity is widening between groups who do and don't have access or coverage for their care. From 2000 to 2001, poor women (<100% of the federal poverty level) had higher rates of induced abortion compared to wealthier women (44 vs. 10 per 1,000 women) (Jones et al., 2002). The prevalence of induced abortion is explained in part by the higher rate of pregnancy among poor women compared to wealthier women (133 vs. 66 per 1,000 women) (Jones et al., 2002). But in the same analysis, the disparity in induced abortion rates increased over time. Between 1994 and 2000, the rate of induced abortion for women with Medicaid coverage increased and the rate decreased for women without Medicaid (Jones et al., 2002).

Thus, this chapter reveals that while several interventions exist to improve access to comprehensive family planning services, the current system of family planning delivery in the United States is comprised of piecemeal interventions rather than a cohesive system ensuring universal access for all women and families. This system creates gaps in access for vulnerable populations in particular, and only furthers disparities in reproductive and perinatal outcomes where health

inequities persist for women of lower socio-economic status and women of color. As such, the potential for comprehensive family planning services to contribute to a reduction in disparities in reproductive and perinatal outcomes holds great promise, but that promise remains substantially unfulfilled.

References

Abma, J. C., Chandra, A., Mosher, W. D., Peterson, L. S., & Piccinino, L. J. (1997). Fertility, family planning, and women's health: New data from the 1995 National Survey of Family Growth. *Vital and Health Statistics*, *23*(19), 1–114.

Almeling, R., Tews, L., & Dudley, S. (2000). Abortion training in U.S. obstetrics and gynecology residency programs, 1998. *Family Planning Perspectives*, *32*(6), 268–320.

Ancel, P. Y., Lelong, N., Papiernik, E., Saurel-Cubizolles, M. J., & Kaminiski, M. (2004). History of induced abortion as a risk factor for preterm birth in European countries: Results of the EUROPOP survey. *Human Reproduction*, *19*(3), 734–740.

Atrash, H., Jack, B. W., Johnson, K., Coonrod, D. V., Moos, M. K., & Stubblefield, P. G., et al. (2008). Where is the "W"oman in MCH? *American Journal of Obstetrics and Gynecology*, *199*(6), S259-S265.

Berg, C. J., Chang, J., Callaghan, W. M., & Whitehead, S. J. (2003). Pregnancy-related mortality in the United States, 1991–1997. *Obstetrics & Gynecology*, *101*(2), 289–296.

Bitler, M., & Zavodny, M. (2001). The effect of abortion restrictions on the timing of abortions. *Journal of Health Economics*, *20*, 1011–1032.

Blank, R. M., George, C. C., & London, R. A. (1996). State abortion rates the impact of policies, providers, politics, demographics, and economic environment. *Journal of Health Economics*, *15*, 513–553.

Boonstra, H. D. (2007) The heart of the matter: Public funding of abortion for poor women in the United States. *Guttmacher Policy Review, 10* (1). Retrieved on December 27, 2009 from http://www.guttmacher.org/pubs/gpr/10/1/gpr100112.html

Bronstein, J. M., Vosel, A., George, S. K., Freeman, C., & Payne, L. A. (2007). Extending Medicaid coverage for family planning services: Alabama's first four years. *Public Health Reports*, *122*, 190–197

Brown, J. S., Adera, T., Masho, S. W. (2008). Previous abortion and the risk of low birth weight and preterm births. *Journal of Epidemiology and Community Health*, *62*, 16–22.

Cates, W., Kimball, A. M., Gold, J., Rubin, G. L., Smith, J. C., & Rochat, R. W., et al. (1979). The health impact of restricting public funds for abortion October 10,1977-June 10, 1978. *Public Health Briefs*, *69* (9), 945–947.

Chasen, S. T., Kalish, R. B., Gupta, M., Kaufman, J., & Chervenak, F. A. (2005). Obstetric outcomes after surgical abortion at ≥20 weeks' gestation. *American Journal of Obstetrics and Gynecology*, *193*, 1161–4.

Chrissman, M., Moore, R., Mondy, L., Weber, B., Peter, W. P., Price, P. C., et al. (1980). Effects of restricting federal funds for abortion – Texas. *Morbidity and Mortality Weekly Report*, 29(2), 253–254.

Contraceptive Coverage in the Federal Employees Health Benefits Program. (2003). Contraception Fact Sheets. Retrieved from http://reproductiverights.org/en/document/contraceptive-coverage-in-the-federal- employees-health-benefits-program

Cook, P. J., Parnell, A. M., Moore, M. J., & Pagnini, D. (1999). The effects of short-term variation in abortion funding on pregnancy outcomes. *Journal of Health Economics*, *18*(2), 241–57.

Culwell, K. R., & Feinglass, J. (2007). The association of health insurance with use of prescription contraceptives. *Perspectives on Sexual & Reproductive Health*, *39*(4), 226–30.

Currie, J., Nixon, L., & Cole, N. (1996). Restrictions on Medicaid funding of abortion: Effects on birth weight and pregnancy resolutions. *Journal of Human Resources*, *31*(1), 159–188.

Dailard, C. (2004). Contraceptive coverage: A 10-year retrospective. *The Guttmacher Report on Public Policy, 7*, 6–9. Retrieved on December 27, 2009 from http://www.guttmacher.org/pubs/tgr/07/2/gr070206.html

Daley, D., & Gold, R. B. (1993). Public funding of contraceptive, sterilization and abortion services, fiscal year 1992. *Family Planning Perspectives*, *25*(6), 244–25.

Eastwood, K. L., Kacmar, J. E., Steinauer, J., Weitzen, S., & Boardman, L. (2006). Abortion training in United States obstetrics and gynecology residency programs. *Obstetrics & Gynecology*, *108*(2), 303–308.

Edwards, J., Bronstein, J., & Adams, K. (2003). Evaluation of Medicaid family planning demonstrations. Final report CNA evaluation of family planning demonstrations, Centers for Medicare and Medicaid Services.

El-Bastawissi, A. Y., Sorensen, T. K., Akafomo, C. K., Frederick, I. O., Xiao, R., & Williams, M. A. (2003). History of fetal loss and other adverse pregnancy outcomes in relation to subsequent risk of preterm delivery. *Maternal and Child Health Journal*, *7*(1) 53–58.

Finer, L. B., Frohwirth, L. F., Dauphinee, L. A., Singh, S., Moore, A.M. (2006). Timing of steps and reasons for delays in obtaining abortions in the United States. *Contraception*, *74*(4), 334–44.

Finer, L. B., & Henshaw, S. (2006). Disparities in rates of unintended pregnancy in the United States, 1994 and 2001. *Perspectives on Sexual and Reproductive Health*, *38*(2), 90–96.

Foster, D. G., Biggs, M. A., Amaral, G., Brindis, C., Navarro, S., & Bradsberry, M., et al. (2006). Estimates of pregnancies averted through California's family planning waiver program in 2002. *Perspectives on Sexual & Reproductive Health*, *38*(3), 126–31.

Foster, D. G., Jackson, R. A., Cosby, K., Weitz, T. A., Darney, P. D., & Drey, E. A. (2008). Predictors of delay in each step leading to an abortion. *Contraception*, *77*(4), 289–93.

Foster, D. G. Rostovtseva, D. P. Brindis, C. D., Biggs, M. A. Hulett, D., & Darney, P. (2009). Cost savings from the provision of specific methods of contraception in a publicly funded program. *American Journal of Public Health*, 99(3), 446–51.

Frost, J. J. Finer, L. B., & Tapales, A. (2008). The impact of publicly funded family planning clinic services on unintended pregnancies and government cost savings. *Journal of Health Care for the Poor and Underserved*, *19*(3), 778–796.

Frost, J. J., Frohwirth, L., & Purcell, A. (2004). The availability and use of publicly funded family planning clinics: U.S. trends, 1994–2001. *Perspectives on Sexual and Reproductive Health*, *36*(5), 206–15.

Gawinski, B. A., Bennett, P. A., Rousseau, S. J., & Schaff, E. (2002). A biopsychosocial model of training in abortion care. *Families, Systems, & Health*, *20*(4), 439–446.

Geller, S. E., Cox, S. M., Callaghan, W. M., & Berg, C. J. (2006). Morbidity and mortality in pregnancy: Laying the groundwork for safe motherhood. *Women's Health Issues*, *16*(4), 176–188.

Geller, S. E., Rosenberg, D., Cox, S. M., Brown, M. L., Simonson, L., & Driscoll, C. A., et al. (2004). The continuum of maternal morbidity and mortality: Factors associated with severity. *American Journal of Obstetrics and Gynecology*, *191*(3), 939- 944.

Gold, R. B. (1980). After the Hyde amendment: Public funding for abortion in FY 1978. *Family Planning Perspectives*, *12*(3), 131–134.

Gold, R. B. (2001). Title X: Three decades of accomplishment. *The Guttmacher Report on Public Policy*, 5–8.

Gold, R. B. (2008). Breaking new ground: Ingenuity and innovation in Medicaid family planning expansions. *Guttmacher Policy Review*, *11*(2).

Gold, R. B. & Daley, D. (1991). Public funding of contraceptive, sterilization and abortion services,1990. *Family Planning Perspectives*, *23*(5), 204–211.

Gold, R. B. & Guardado, S. (1988). Public funding of contraceptive, sterilization and abortion services, 1987. *Family Planning Perspectives*, *20*(5), 228–233.

Gold, R. B. & Macias, J. (1986). Public funding of contraceptive, sterilization and abortion services, 1985. *Family Planning Perspectives*, *18*(6), 259–264.

Gold, R. B. & Nestor, B. (1985). Public funding of contraceptive, sterilization and abortion services, 1983. *Family Planning Perspectives*, *17*(1), 25–30.

Gold, R. B., Richard, C. L. Ranji, U. R. & Salganicoff, A. (2007). Medicaid's role in family planning. Retrieved from http://www.kff.org/womenshealth/7064.cfm

Gold, R. B., Sonfield, A., Richards, C. L., & Frost, J. J. (2009). Next steps for America's family planning program: Leveraging the potential of Medicaid and Title X in an evolving health care system. Retrieved from http://www.guttmacher.org/pubs/summaries/NextStepsExec.pdf

Goldberg, M. (2009). The great population panic, or fighting communism with contraception. The means of reproduction. New York: The Penguin Press.

Guttmacher Institute (2009a). Insurance Coverage of Contraceptives. *State Policies in brief. retrieved from* http://www.guttmacher.org/statecenter/spibs/spib_ICC.pdf accessed October 15, 2009.

Guttmacher Institute, (2009b). Facts on Induced Abortion in the United States *In Brief*, http://www.guttmacher.org/pubs/fb_induced_abortion.html, accessed October 15, 2009.

Gruber, J., Levine, P., & Staiger, D. (1999). Abortion legalization and child living circumstances: Who is the "marginal child"? *Quarterly Journal of Economics*, *114*(1), 263–291.

Haas-Wilson, D. (1993). The economic impact of state restrictions on abortion: Parental consent and notification laws and Medicaid funding restrictions. *Journal of Policy Analysis and Management*, *12*(3), 498–511.

Haas-Wilson, D. (1997). Women's reproductive choices: The impact of Medicaid funding restrictions. *Family Planning Perspectives*, *29* (5), 228–233.

Henshaw, S. K. (1998). Unintended pregnancy in the United States. *Family Planning Perspectives*, *30*(1), 24–29, 46.

Henshaw, S. K., Joyce, T. J., Dennis, A., Finer, L. B., & Blanchard, K. (2009). Restrictions on Medicaid funding for abortions: A literature review, New York: Guttmacher Institute, Retrieved from http://www.guttmacher.org/pubs/MedicaidLitReview.pdf

Henshaw, S. K., & Wallisch, L. S. (1984). The Medicaid cutoff and abortion services for the poor. *Family Planning Perspectives*, *16*(4), 170–180.

Insurance Coverage of Contraceptives. (2009). State policies in brief, retrieved from http://www.guttmacher.org/statecenter/spibs/spib_ICC.pdf

IOM (Institute of Medicine). (2009). A review of the HHS family planning program: Mission, management, and measurement of results. Washington, DC: The National Academies Press.

Jones, R. K., Darroch, J. E., & Henshaw, S. K. (2002). Patterns in the socioeconomic characteristics of women obtaining abortions in 2000–2001. *Perspectives on Sexual & Reproductive Health*, *34*(5), 226–35.

Jones, R. B., Kost, K., Singh, S., Henshaw, S. K., & Finer, L. B. (2009) Trends in abortion in the United States. *Clinical Obstetrics and Gynecology*, *52*(2), 119–129.

Joyce, T., & Kaestner, R. (1996). State reproductive policies and adolescent pregnancy resolution: The case of parental involvement laws. *Journal of Health Economics*, *15*, 579– 607.

Kaiser Family Foundation. (2007). *Women's Health Insurance Coverage, Women's Health Policy Facts*. Retrieved from http://www.kff.org/womenshealth/upload/7064_03.pdf

Kane, T. J., & Staiger, D. (1996). Teen motherhood and abortion access. *Quarterly Journal of Economics*, *111*(2), 467–506.

Kearney, M. K., & Levine, P. B. (2009). Subsidized contraception, fertility, and sexual behavior. *Review of Economic & Statistics*, *91*(1), 1–26.

Klerman, L. V., Johnson, K. A., Chang, C. H., Wright-Slaughter, P., & Goodman,, D. C. (2007). Accessibility of family planning services: Impact of structural and organizational factors. *Maternal & Child Health Journal*, *11*(1), 19–26.

Korenbrot, C. C., Brindis, C., & Priddy, F. (1990). Trends in rates of live births and abortions following state restrictions on public funding of abortion. *Public Health Reports*, *105*(6), 555–562.

Levine, P. B., Trainor, A. B., & Zimmerman,, D. J. (1996). The effect of Medicaid abortion funding restrictions on abortions, pregnancies, and births. *Journal of Health Economics*, *15*, 555–578.

Lindheim, B. L., & Cotterill, M. A. (1978). Training in induced-abortion by obstetrics and gynecology residency programs. *Family Planning Perspectives*, *10*(1), 24–28.

Lindrooth, R. C., & McCullough, J. S. (2007). The effect of Medicaid family planning expansions on unplanned births. *Women's Health Issues*, *17*(2), 66–74.

Lundberg, S., & Plotnick, R. D. (1990). Effects of state welfare, abortion and family planning policies on premarital childbearing among white adolescents. *Family Planning Perspectives*, *22*(6), 246–275.

Mackay, H. T., & Mackay, A. P. (1995). Abortion training in obstetrics and gynecology residency programs in the United States, 1991–1992. Family *Planning Perspectives*, *27*(3), 112–115.

Mandelson, M. T., Maden, C. B., & Daling, J. R. (1992) Low birth weight in relation to multiple induced abortions. *American Journal of Public Health*, *82* (3), 391–394.

Medoff, M. H. (1988). An economic analysis of the demand for abortions. *Economic Inquiry*, *26*(2), 353–359.

Meier, K. J., & Mcfarlane, D. R. (1994). State family planning and abortion expenditures: Their effect on public health. *American Journal of Public Health*, *84*(9), 1468–1472.

Misra, D. P., & Grason, H. (2000). An intersection of women's and perinatal health: The role of chronic conditions. *Women's Health Issues*, *10*(5), 256–267.

Mohllajee, A. P., Curtis, K. M., Morrow, B., & Marchbanks, P. A. (2007). Pregnancy intention and its relationship to birth and maternal outcomes. *Obstetrics & Gynecology*, *109*(3), 678–86.

Moreau, C., Kaminiski, M., Ancel, P. Y., Bouyer, J., Escande, B., Thiriez, G., et al. (2005). Previous induced abortions and the risk of very preterm delivery: Results of the EPIPAGE study. *BJOG: An International Journal of Obstetrics and Gynaecology*, *112*, 430–437.

Morgan, S. P., & Parnell, A. M. (2002). Effects on pregnancy outcomes of changes in the North Carolina state abortion fund. *Population Research and Policy Review*, *21*, 319–338.

Mosher, W., Martinez, G., Chandra, A., Abma, J., & Willson, S. Use of contraception and use of family planning services in the United States: 1982–2002. Hyattsville, MD: National Center for Health Statistics; December 10 2004. Retrieved from http://www.cdc.gov/nchs/data/ad/ad350FactSheet.pdf

Nearns, J. (2009). Health insurance coverage and prescription contraceptive use among young women at risk for unintended pregnancy. *Contraception*, *79*(2), 105–10.

Nestor, B., & Gold, R. B. (1984). Public funding of contraceptive, sterilization and abortion services, 1982. *Family Planning Perspectives*, *16*(3), 128–133.

New, M. J. (2007). Analyzing the effect of state legislation on the incidence of abortion among minors. The Heritage Foundation: Washington, D.C. Retrieved from http://www.heritage.org/Research/Family/CDA07–01.cfm

Nixon, R. *Special Message to the Congress on Problems of Population Growth*. President Richard Nixon, July 18, 1969. The Nixon Foundation and Library. Retrieved from http://www.nixonlibraryfoundation.org/clientuploads/directory/archive/1969_pdf_files/1969_0271.pdf

Ranjit, N., Bankole, A., Darroch, J. E., & Singh, S. (2001). Contraceptive failure in the first two years of use: Differences across socioeconomic subgroups. *Family Planning Perspectives*, 33(1), 19–27.

Rubin,, G. L., Gold, J., & Cates,, W. (1979). Response of low income women and abortion facilities to restriction of public funds for abortion: A study of a large metropolitan area. *American Journal of Public Health*, *69*(9), 948–950.

Shah, P. S., & Zao, J. (2009) Induced termination of pregnancy and a low birth weight and preterm birth: A systemic review and meta-analyses. *British Journal of Obstetrics and Gynaecology*, *116*, 1425–1442.

Singh, S. (1986). Adolescent pregnancy in the United States: An interstate analysis. *Family Planning Perspectives*, *18*(5), 210–220.

Singh, S., Darroch, J. E., Ashford, L. S., & Vlassoff, M. (2009). *Adding it up: The costs and benefits of investing in family planning and maternal and newborn health.* New York: Guttmacher Institute and United Nations Population Fund. Retrieved from: http://www.unfpa.org/webdav/site/global/shared/documents/publications/2009/adding_it_up_report.pdf

Sollom, T., Gold, R. B., & Saul, R. (1996). Public funding of contraceptive, sterilization and abortion services, fiscal year 1994. *Family Planning Perspectives*, *28*(4), 166–173.

Sonfield, A. (2003). Preventing unintended pregnancy: The need and the means. The Guttmacher Report on Public Policy 6. Retrieved from http://www.guttmacher.org/pubs/tgr/06/5/gr060507.html

Sonfield, A., Alrich, C., & Gold, R. B. (2008). Public funding of contraceptive, sterilization and abortion services, FY 1980–2006. Occasional Report, New York: Guttmacher Institute, 38. Retrived from http://www.guttmacher.org/pubs/2008/01/28/or38.pdf

Sonfield, A., Gold, R. B., Frost, J. J., & Darroch, J. E. (2004). U.S. Insurance Coverage of Contraceptives and the Impact of Contraceptive Coverage Mandates, 2002. *Perspectives on Sexual & Reproductive Health*, *36*(2), 72–9.

Sonnenberg, F., Burkman, R., Hagerty, C., Speroff, L., & Speroff, T. (2004). Costs and net health effects of contraceptive methods. *Contraception*, *69,*447–459.

State Funding of Abortion Under Medicaid. (2009). Guttmacher Institute. Retrieved on July 1, 2009 from http://www.guttmacher.org/statecenter/spibs/spib_SFAM.pdf

Steinauer, J., Depineres, T., Robert, A., Westfall, J., & Darney, P. D. (1997). Training family practice residents in abortion and other reproductive health care: A nationwide survey. *Family Planning Perspectives*, *29*(5), 222–227.

Steinauer, J. E., Landy, U. P., Jackson, R. A., & Darney, P. D. (2003). The effect of training on the provision of elective abortion: A survey of five residency programs. *American Journal of Obstetrics & Gynecology*, *188*(5), 1161–1163.

Strauss, L., Gamble, S., Parker, W., Cook, D., Zane, S., & Hamdan, S. (2007). Abortion surveillance – United States, 2004. MMWR Surveillance Summaries *53*, 1–33. Retrieved on December 28, 2009 from http://www.cdc.gov/mmWR/preview/mmwrhtml/ss5609a1.htm

Swingle, H. M., Colaizy, T. T., Zimmerman, M. B., & Morriss, F. H. (2009) Abortion and the risk of subsequent preterm birth, a systematic review with meta-analyses. *Journal of Reproductive Medicine*, *54*, 95–108.

The growth of U.S. population: Analysis of the problems and recommendations for research, training, and service. (1965) National Academy of Sciences – National Research Council Committee on Population, Washington, DC. Retrieved from: http://books.google.com/books?id=cjErAAAAYAAJ&dq=The+growth+of+U.S.+population:+analysis+of+the+problems+and+recommendations+for+research,&printsec=frontcover&source=bl&ots=kJXFubsyqJ&sig=pQ9yQ_M3575iTGyMafXU97ojqDw&hl=en&ei=ZNBES7mPB438MeSAsYgH&sa=X&oi=book_result&ct=result&resnum=3&ved= 0CBEQ6AEwAg#v=onepage&q=&f=false

Tietze, C. (1975). Effect of legalization of abortion on population-growth and public-health. *Family Planning Perspectives*, *7*(3), 123–127.

Trends in Abortion in the United States, 1973–2005. (2008). Retrieved July 2, 2009, from http://www.alanguttmacherinstitute.org/presentations/trends.html

Trussell, J. (2007). The cost of unintended pregnancy in the United States. *Contraception*, *75*(3), 168–170.

Trussell, J., Leveque, J. A., Koenig, J. D., London, R., Borden, S., Henneberry, J., et al. (1995). The economic value of contraception: A comparison of 15 methods. *American Journal of Public Health*, *85*(4), 494–503.

Trussell, J., Menken, J., Lindheim, B. L., & Vaughan, B. (1980). The impact of restricting Medicaid financing for abortion. *Family Planning Perspectives*, *12*(3), 120–130.

Vital Statistics of the United States, 1965 (1967). National Center for Health Statistics, U.S. Government Printing Office. Vol II – Mortality, Part A.

Vital Statistics of the United States, 1966 (1968). National Center for Health Statistics, U.S. Government Printing Office. Vol II – Mortality, Part A.

Vital Statistics of the United States, 1967 (1969). National Center for Health Statistics, U.S. Government Printing Office. Vol II – Mortality, Part A.

Wetstein, M. E. (1995). The abortion rate paradox: The impact of national policy change on abortion rates. *Social Science Quarterly*, *76*(3), 607–618.

Wu, J., Meldrum, S., Dozier, A., Stanwood, N., & Fiscella, D. (2008). Contraceptive nonuse among US women at risk for unplanned pregnancy. *Contraception*, *78*, 284–289.

Chapter 5
Preconceptional Health Promotion

Merry-K. Moos and Amanda C. Bennett

More than two decades ago, a quiet but persistent energy to rethink and reframe the nation's traditional efforts to impact poor pregnancy outcomes began [Institute of Medicine (IOM), 1985; Moos & Cefalo, 1987; Thompson, Walsh, & Merkatz, 1990]. Built on the foundation that prenatal care starts too late to provide primary prevention of the two leading causes of infant morbidity and mortality, congenital malformation and low birth weight, the early effort, known as preconceptional health promotion, has grown to attract national attention.

The purpose of this chapter is to review the literature on preconceptional health promotion to determine whether interventions and services during this period are effective in improving pregnancy outcomes and reducing racial/ethnic disparities in these outcomes. First, the chapter provides an overview of the history of the preconceptional care movement and the evidence towards the areas in which preconceptional health promotion may be the most effective. The remainder of the chapter is then broken into two sections: one relating to preconceptional diabetes care, and one relating to preconceptional folic acid interventions. Each of these sections contains an overview of the methods used in the literature search, a summary of the evidence for the intervention, and a discussion of the implications of this evidence for racial/ethnic disparities. At the end of the chapter, there is discussion of overall preconceptional health promotion and a set of recommendations for future research in the area of racial/ethnic disparities.

Background

In early 2006, the Centers for Disease Control and Prevention (CDC) published a report of the CDC/ATSDR Preconception Care Work Group and the Select Panel on Preconception Care, entitled *Recommendations to Improve Preconception Health and Health Care – United States* (Johnson et al., 2006). With this publication, national energy to reframe the perinatal prevention paradigm gained momentum. The CDC's Select Panel on Preconceptional Health put forth four goals:

- Improve knowledge, attitudes, and behaviors of men and women related to preconception health.
- Assure that all women of childbearing age receive preconception care services (i.e., evidence-based risk screening, health promotion and interventions) that will enable them to enter pregnancy in optimal health.

M.-K. Moos (✉)
Maternal Fetal Medicine Division, Department of Obstetrics and Gynecology University of California, Campus Box #7516, Chapel Hill, NC 27599-7516, USA
e-mail: mkmoos@med.unc.edu

A. Handler et al. (eds.), *Reducing Racial/Ethnic Disparities in Reproductive and Perinatal Outcomes: The Evidence from Population-Based Interventions*, DOI 10.1007/978-1-4419-1499-6_5,

- Reduce risks identified by a previous poor pregnancy outcome through interventions during the interconception period, which can prevent or minimize health problems for a mother and her future children.
- Reduce disparities in adverse pregnancy outcomes.

The premise behind promoting preconception in addition to prenatal care is that for some maternal conditions and exposures, unalterable damage (such as abnormal placentation, spontaneous loss or congenital malformations) can occur before prenatal care begins. It is in the earliest days following fertilization that the placental foundations supporting the rest of the pregnancy are laid and it is days 17–56 following fertilization (commencing 3 days after the first missed menstrual period) that determines whether the fetus will be affected by major congenital anomalies. This window of development is referred to as the period of embryogenesis or organogenesis; its timing makes it virtually impossible to alter the occurrence of congenital anomalies through traditional prenatal care. Therefore, to promote normal placentation and reduce risks of birth defects, education and appropriate interventions must be identified and implemented prior to conception.

Little is known about successful strategies to achieve these goals or whether unintended consequences will attend efforts to reframe the perinatal prevention paradigm. Two potential negative consequences have been suggested (Moos, 2006). The first is that women will be held accountable for securing a new category of care, the preconception visit, and the second is that a focus on preconceptional health promotion could cast all women as potential mothers, irrespective of their personal reproductive goals or fecundity. Moos (2006) addresses both of these potential problems in an editorial. Another unintended consequence is that, rather than reducing disparities in adverse pregnancy outcomes, an emphasis on preconception health promotion may have an opposite impact in that disparities may, at least initially, widen. Exploration of this hypothesis follows.

The American College of Obstetricians and Gynecologists (ACOG) and the American Academy of Pediatrics (AAP) have grouped potential preconception interventions under four main categories (American Academy of Pediatrics & the American College of Obstetricians and Gynecologists, 2007): maternal assessment (e.g., family history, behaviors, obstetric history, general physical exam), vaccinations (e.g., rubella, varicella and hepatitis B), screening (e.g., HIV, STD, genetic disorders), and counseling (e.g., folic acid consumption, smoking and alcohol cessation, weight management). Atrash and colleagues (Atrash, Johnson, Adams, Cordero, & Howes, 2006) undertook a review of the literature and identified six categories of potential risk that deserved attention during the preconception period. The categories are:

- Chronic disease (e.g., diabetes, heart disease, hypertension, etc.).
- Infectious diseases (e.g., immunity to vaccine preventable diseases, risks related to sexually transmitted disease, etc.).
- Reproductive concerns (e.g., contraception, unplanned pregnancy, infertility, previous adverse pregnancy outcomes, etc.).
- Genetic/inherited conditions (e.g., sickle cell anemia, fragile X syndrome, cystic fibrosis, etc.).
- Medication and medical treatment (e.g., use of prescription medications known to cause abnormal embryogenesis).
- Personal behaviors and exposures (e.g., tobacco, alcohol and other exposures to mood altering substances, folic acid supplementation, domestic violence).

Atrash and colleagues (2006) also looked for clinical practice guidelines (CPGs) to guide preconceptional care and found 14 separate conditions or exposures for which CPGs exist. They are:

- Folic acid
- Rubella seronegativity
- Diabetes mellitus
- Hypothyroidism

- HIV/AIDs
- Maternal phenylketonurea (PKU)
- Oral anticoagulants
- Antiseizure drugs
- Isotretinoins
- Smoking
- Alcohol
- Obesity
- STDs
- Hepatitis B

While the theoretical advantages of preconceptional health promotion grow, the actual effectiveness of preconception care interventions is largely untested. Some exceptions, however, do exist. In 2002, Korenbrot and colleagues (Korenbrot, Steinberg, Bender, & Newberry, 2002) published a systematic review of the effectiveness of pre-pregnancy interventions for impacting pregnancy outcomes. Using criteria employed by the U.S. Clinical Preventive Services Task Force, their review presented the evidence for preconception care interventions according to two major criteria for effectiveness: preconception health activities resulted in the risk condition being detected or treated earlier than without the activity, and the likelihood that there would be a change in a measurable pregnancy process or outcome for mothers or infants as the result of the activity. A search was conducted for literature associated with more than 40 preconception risk conditions, and 470 articles were abstracted. The researchers found demonstrated effectiveness only for: screening women who present for family planning care with risk factors, encouraging sexually active women of reproductive age to take dietary folate supplements, and providing women affected by certain metabolic conditions (diabetes and hyperphenylalanemia) with clinical interventions.

The largest bodies of research on the effectiveness of preconceptional interventions are the studies supporting preconceptional glucose control and preconceptional folic acid supplementation. The remainder of this chapter will focus on the evidence base for these interventions and will address whether widespread adoption of the interventions will be associated with a narrowing or widening of health disparities in reproductive and perinatal outcomes.

Diabetes: Perinatal Risks and Opportunities for Preconceptional Interventions

A woman with poorly controlled diabetes at the time of conception is at increased risk for a number of adverse pregnancy outcomes, including spontaneous abortion and congenital malformations. Major congenital anomalies occur in 6–12% of women with preconceptional diabetes and are the leading cause of perinatal mortality for this population (ACOG, 2005). Hyperglycemia (high blood sugar) during the period of organogenesis (days 17–56 following conception), is believed to be the major cause of abnormal embryogenesis in pregnancies of women with diabetes.

There are two types of pregestational diabetes, both of which are associated with increased risk for congenital anomalies in offspring. Type I diabetes mellitus tends to occur early in life and is characterized, in part, by an autoimmune process that destroys the ability of the pancreas to produce insulin (ACOG, 2005). There are no proven prevention strategies against Type I diabetes. Type II diabetes is much more common and is characterized by resistance to insulin. It is the most common type of pregestational diabetes seen in pregnancy and its prevalence is rapidly increasing, in part because of the obesity epidemic in the United States (ACOG, 2005). Weight management through avoidance of overweight (BMI ≥25) and obesity (BMI ≥30.0) is a proven protection against Type II diabetes.

The goal of preconceptional care for women with diabetes is to help them achieve a normal blood sugar level, known as euglycemia, before pregnancy and to continue tight control of blood sugar levels throughout the first trimester of pregnancy. Most studies evaluating the impact of preconceptional care for women with diabetes correlate hemoglobin A1c levels with pregnancy outcomes (Ray, O'Brien, & Chan, 2001). Indication of ideal control, according to the American Diabetes Association (ADA, 2004) is a glycosolated hemoglobin A1c level that is less than 1% higher than reference levels. Achieving optimal hemoglobin A1c levels requires patient education about family planning and the interactions of diabetes and pregnancy, education in diabetes self-management skills including dietary intake, self monitoring and medication administration, and medical supervision of laboratory and other indicators of disease status.

Evidence for the Effectiveness of Preconceptional Diabetes Control

Searches of Medline, CINAHL and of the reference lists of identified articles were undertaken in the summer of 2006. Search strategies and key words were chosen to locate reports from 2000 to 2006 which could illuminate the impact of preconceptional diabetes control on pregnancy outcomes, how the impact might be differentially expressed among racial and ethnic subpopulations, and whether disparities exist in either women's abilities to achieve preconceptional diabetes control or on pregnancy outcomes related to poor control. Classic literature which predates 2000 was also used to provide background information.

Impact of Preconceptional Diabetes Control on Pregnancy Outcomes. A meta-analysis by Ray et al. (2001) underscored the protective benefits of preconceptional care on the incidence of congenital anomalies in diabetic women (see Table 5.1). The 16 studies included in the analysis were from a variety of countries including the United States, Germany, Denmark, the United Kingdom, Scotland and Spain. The reviewers calculated a pooled relative risk and found the incidence of major congenital anomalies in diabetic women exposed to preconceptional diabetes control interventions to be one-third that of diabetic women who did not receive the same exposure. Of the seven studies from the United States included in this meta-analysis (Cousins, 1991; Herman, Janz, Becker, & Charron-Prochownik, 1999; Janz et al., 1995; Kitzmiller et al., 1991; Rosenn, Miodovnik, Combs, Khoury, & Siddiqi, 1991; The Diabetes Control and Complications Trial Research Group, 1996; Willhoite et al., 1993), few included information on the racial or ethnic composition of the samples and none reported analyses based on these demographic variables. It was in light of this meta-analysis that ACOG recommended that preconceptional care be encouraged for all women with diabetes who have the potential to become pregnant (ACOG, 2005).

A 2000 study by McElvy et al., which was not included in the Ray et al. (2001) meta-analysis, aimed to evaluate the impact of a focused preconceptional and early pregnancy program for women with Type I diabetes mellitus on perinatal mortality and congenital malformations (see Tables 5.2, 5.3). The study, which utilized a retrospective comparison analysis, supports the findings of the meta-analysis for the perinatal benefits of preconception care. As with the studies included in the meta-analysis, no analysis regarding racial and ethnic subpopulations was offered.

Studies on Diabetes Control for Subpopulations. To understand the potential impact of preconceptional diabetes control on various ethnic and racial populations, we sought reports that investigated access to care and quality of services by subpopulations. No studies were found specific to women of childbearing age. Studies examining racial disparities in general diabetes care report a narrowing in the achievement of quality markers between non-Hispanic blacks and whites. Several descriptive reports (Heisler, Smith, Hayward, Krein, & Kerr, 2003; LeMaster, Chanetsa, Kapp, & Waterman, 2006; Sequist, Adams, Zhang, Ross-Degnan, & Ayanian, 2006) suggest that various aspects of

Table 5.1 Summary of a meta-analysis of preconceptional diabetes care

Source	Number of studies/N/% in "preconception care" (PCC) and "no preconception care" (no PCC) groups	Findings	Contextual factors	Disparities/comments
Ray et al. (2001)	Eight prospective and eight retrospective cohort studies Major and minor anomalies as outcomes: nine studies, N = 2,104 (43% in PCC group, 57% in no PCC group) Major anomalies as outcome: 14 studies, N = 2,651 (45% in PCC group, 55% in no PCC group)	Calculated summary statistics Favors PCC: reduced rate of major and minor congenital anomalies in offspring of PCC group compared to the no PCC group Major and minor anomalies: Pooled RR = 0.32 (0.17–0.59) Major anomalies: Pooled RR = 0.36 (0.22–0.59)	Settings of studies (# studies): – United Kingdom (3) – Europe (5) – United States (7) – Israel (1) Three studies include Types I and II diabetes – all other studies only include Type I diabetes patients Three studies included some in-patient PCC intervention activities Wide variation between studies in approaches to type, duration, and extent of PCC care, as well as assessment of glycemic control	Does not present data by race/ethnicity

diabetes care have improved but that disparities persist. In two of the studies (Heisler et al., 2003; Sequist et al., 2006), receipt of hemoglobin A1c testing was not significantly different between non-Hispanic blacks and whites. However, in both these studies, the mean age was beyond childbearing potential, making it difficult to extrapolate the findings to the preconceptional population. In a potentially more relevant study, LeMaster and colleagues (2006) used data from the Missouri Behavioral Risk Factor Surveillance System which allowed for a broader age distribution of respondents. They found that non-Hispanic blacks were less likely than whites to report hemoglobin A1c testing. Harris (2001) reported National Health and Nutrition Examination Survey III (NHANES) data on 1,480 people with Type II diabetes. The study goal was to evaluate health care access, utilization, health status and outcomes among non-Hispanic whites and blacks and Mexican-Americans. Again, the sample was beyond childbearing age and included a significant number of men. The author found some differences by race and ethnicity in health care access, utilization, health status and outcome for patients with Type II diabetes but "the magnitude of these differences pale in comparison with the suboptimal health status of all three race and ethnic groups relative to established treatment goals" (p. 454). The author noted that health status did not appear to be influenced by access to health care. These findings are especially concerning when considering the likely impact of new guidelines for the care of women with diabetes who are of childbearing potential. Given that adherence to long standing guidelines has been so poor, onr might well wonder what the likelihood is that dissemination of new evidence-based recommendations will change practice or health status.

Table 5.2 Major outcomes associated with an individual study of preconceptional diabetes care

Author (year)	Study design/ study type	Description of intervention what, how and where	Populations studied (ages included, race and ethnicity) and sample size	Address disparities (yes/no)	Key findings related to intervention effectiveness	Caveats and biases	Findings support the intervention? (yes/no); For which populations?
McElvy et al. (2000)	Series of different intervention groups with no controls: article examined the difference in outcomes over the three PPG[a] stages/ published article	PPG I (1978–1983): Enrolled women (see population studied) were seen by care team every 1–2 weeks to receive tailored intensive diabetes education, a dietary consultation, diabetes management instruction, and education about risks of poor glycemic control during pregnancy. Most deliveries were by Cesarean section PPG II (1983–1988): Same as above, with increased focus and advertising on preconception enrollment, visits changed to weekly in 1984 PPG III (1988–1993): Same as above, with vaginal delivery attempted in all women unless contraindicated	Women with diabetes who were planning a pregnancy within 12 months of enrollment or were already pregnant	No	PPG II and PPG III, which had higher preconceptional enrollment, were associated with reductions in major congenital malformations and perinatal mortality compared to PPG I However, there was no statistical analysis of differences between PPG stages	Different populations participated in each stage of the program, with no concurrent control groups. These groups had significant differences in age, duration of diabetes, chronic hypertension, and weight gain	Trend seems to indicate benefit of PCC[b] for diabetes, but needs more examination with control groups and controlling for confounding factors

[a] *PPG* Program Project Grant
[b] *PCC* Preconceptional Care

Table 5.3 Quality rating of an individual study of preconceptional diabetes care: association with major and minor congenital anomalies

Author (year)	Reporting	External validity	Internal validity – bias	Internal validity – confounding	Power	Total quality score ≤14 = poor 15–19 = fair ≥20 = good	Suitability of study to assess effectiveness: greatest, moderate, least
McElvy et al. (2000)	10	1	4	1	1	17	Moderate

Content of Preconceptional Care for Women with Diabetes. A study of women ages 18–45 years enrolled in diabetes care conducted in managed care programs in three states described the content of care for reproductive aged women with diabetes (Kim et al., 2005). Data were collected through a computer-assisted telephone interview or in writing; data collection was augmented by medical record review. The study found that 52% of women recalled being counseled about the importance of optimal glucose control before pregnancy and 37% reported discussions about using a contraceptive method until optimal glucose levels were achieved. The sample included non-Hispanic whites and blacks, Hispanics and Asian/Pacific Islanders. There was no difference in recall of these emphases in care by race or ethnicity. In a study of 55 diabetic women assessed at entry to prenatal care, women who reported that they had not received specific advice about target hemoglobin A1c levels were more likely to enter prenatal care with suboptimal glucose control than those who reported receiving specific advice ($p = 0.02$) (Casele & Laifer, 1998). In this same study, women who had experienced a previous pregnancy with complications were also more likely to begin prenatal care with suboptimal control than those without a previous pregnancy with complications ($p = 0.02$).

In a retrospective investigation of 85 women with Type I diabetes diagnosed prior to pregnancy, Holing, Brown, Beyer, and Connell (1998) uncovered a possible relationship between provider attitudes about a woman becoming pregnant and her disclosure of the pregnancy as being planned or unplanned. Women who recalled that their clinician was reassuring and encouraging about the prospects of a healthy pregnancy and child, were more likely to identify their conceptions as planned than were women who recalled the clinician as discouraging and negative.

Public Health Implications for Reducing Racial/Ethnic Disparities

Little is known about how to successfully translate the body of literature supporting the benefits of preconceptional care for diabetic women into interventions to decrease racial disparities in diabetes-related congenital anomalies and other poor pregnancy outcomes. No quality research was found that reported on the differential impact of preconceptional diabetes care in racial and ethnic minorities or that included representative samples of women of childbearing age affected by diabetes mellitus. This gap in the literature is significant because of clear disparities in the prevalence of diabetes by race/ethnicity; among individuals 20 years old and greater, diabetes affects 8.7% of non-Hispanic whites, 13.3% of non-Hispanic blacks, 9.5% of Hispanic/Latino Americans, and 15.1% of American Indians and Alaska Natives (National Diabetes Information Clearinghouse, 2005). Over time, these disparities are likely to be impacted by the nation's obesity epidemic, as there are important racial/ethnic disparities in obesity as well. Between the time periods 1976–1980 and 2003–2004, the prevalence of obesity (BMI ≥ 30.0) in females ages 20 years and older increased in blacks from 31.0 to 53.9% (a 73.8% increase), in Hispanics from 26.6 to 42.3% (a 59% increase), and in whites from 15.4 to 30.2% (a 96% increase) (National Center for Health Statistics, 2003; Ogden et al., 2006).

Differing prevalence rates of diabetes places racial and ethnic subgroups at unequal risk for associated congenital anomalies and related poor pregnancy outcomes. However, this review found no studies suggesting that the pregnancies of subgroups would be differentially responsive to the benefits of preconceptional disease control. Therefore, access to and utilization of informed care becomes the salient concern in examining how preconceptional services are likely to impact disparities.

Owens, Kieffer, and Chowdhury (2006) identified several barriers to receiving preconception care related to diabetes which could differentially impact the risks of affected pregnancies in racial and ethnic subgroups. Included in their list are a woman's awareness of her own disease, knowledge about the associated reproductive risks, the rate of unintended conceptions, the likelihood of health insurance or other payment options to afford preconceptional care, and access to health care providers who are aware of and adhere to clinical care guidelines related to preconceptional glucose control. Access to preconceptional care in general (not specific to diabetes) is more common among certain socio-demographic groups. Several studies have described women who received general preconception care as more likely to be non-Hispanic whites, married or in a stable relationship, comparatively older and better educated, nonsmokers, and having annual incomes above $20,000, private medical insurance, and a more positive relationship with a prepregnancy care provider then women who did not receive preconception care (Holing et al., 1998; Janz et al., 1995). Clearly, these characteristics do not favor minority populations' access to clinical services as they are currently provided.

In addition to the access issue, effective preconceptional health promotion for women with diabetes is dependent on the knowledge, attitudes, and skills of their medical providers. There is evidence from research of the general population of diabetic patients that access to services does not necessarily result in optimal health status, irrespective of ethnicity or race (Harris, 2001). Access must be coupled with provider commitment to adhering to evidence-based guidelines. To date, scant attention has been paid in the literature to identifying strategies aimed at improving clinical practice regarding preconceptional care for diabetes.

A 2004 report of a special steering committee to advise the CDC on directions for a national public health initiative on diabetes and women's health (CDC, 2004a) made ten recommendations; two of them are of particular relevance given the poor state of our knowledge about the likely impact of preconceptional care on disparities in birth outcomes for women with pregestational diabetes. The first of these is to expand population-based surveillance to monitor and understand variations in the distribution of diagnosed and undiagnosed diabetes and the factors – cultural, racial, ethnic, geographic, demographic, socioeconomic and genetic – that influence the risk for diabetes and complications among women at all life stages. The second is to conduct public health research to further our knowledge about the epidemiological, socio-environmental, behavioral, translation and biomedical factors that influence diabetes and women's health. Achievement of these two recommendations in tandem with an increased emphasis on the prevention of obesity and improved access to competent preconceptional care is needed to understand and address disparities in the prepregnancy health status of women with diabetes and the related pregnancy outcomes.

Neural Tube Defects: Perinatal Risks and Opportunities for Preconceptional Interventions

Neural tube defects (NTDs) are the second most common major congenital anomaly worldwide. NTDs include structural abnormalities of the cranium and vertebral column, which are formed by the end of the fourth week following conception, with the critical window being during days 17–28. Numerous abnormalities are included under the heading of neural tube defects and the severity of the abnormalities can vary (ACOG, 2003). Anencephaly, which represents absence of all or part of the brain, skull and skin, accounts for approximately 50% of all NTDs and is incompatible with

sustained life (ACOG, 2003). In contrast, 90% of children born with spina bifida of the sacral area will survive and be able to walk, but they could experience life long complications related to their birth defect, such as impairment of bowel and bladder function and severe allergy to latex (Bowman, McLone, Grant, Tomita, & Ito, 2001; McDonnell & McCann, 2000) (Table 5.3).

Prior to 1992, approximately 4,000 pregnancies resulting in 2,500–3,000 births in the United States were affected by spina bifida or anencephaly annually (Mathews, Honein, & Erickson, 2002). Of affected pregnancies, one-third end in spontaneous or elective abortion (Botto, Moore, Khoury, & Erickson, 1999). NTDs occur either as an isolated anomaly or as part of a syndrome in which many organs and functions are affected. Isolated NTDs are believed to be the result of a combination of genetic predisposition and environmental influences. Epidemiologic data indicates that geographic region, ethnicity and maternal teratogenic exposures during the first month of gestation can influence normal neural tube development. An important environmental influence is the maternal blood folate level (Czeizel & Dudas, 1992; Medical Research Council Vitamin Study Research Group, 1991; Mulinare, Cordero, Erickson, & Berry, 1988; Lumley, Watson, Watson, & Bower, 2001). Folate levels can be altered by consuming folate rich foods or by ingesting folic acid, a synthetic compound available through dietary supplements and fortified foods. Folic acid is approximately 50% more bioavailable than folate and therefore has a greater efficiency in impacting maternal folate levels (Neuhouser & Beresford, 2001).

Compelling evidence about the benefits of folic acid in the prevention of NTDs resulted in the 1992 U. S. Public Health Service recommendation that "all women of childbearing age in the United States who are capable of becoming pregnant should consume 0.4 mg of folic acid per day for the purpose of reducing their risk of having a pregnancy affected with spina bifida or other NTDs" (CDC, 1992, p. 1). Subsequently, the U.S. Food and Drug Administration (FDA) approved a population-based strategy to fortify all enriched grain products with folic acid, effective January 1998 (CDC, 2004b). The fortification levels were set relatively low but have resulted in the average woman who eats a standard diet ingesting an extra 190–240 μg of folic acid daily (CDC, 2004b; Williams et al., 2002). In 1998, the Institute of Medicine (IOM) affirmed the U.S. Public Health Service recommendation and added that women of childbearing years should take 400 μg of synthetic folic acid daily, obtained from fortified foods and/or supplements, as well as consume a balanced, healthy diet of folate rich food (IOM, 1998). The IOM recommendation combines the three approaches available to individual women to increase their exposure to folic acid: ingestion of fortified foods, supplementation, and ingestion of foods naturally rich in folate. The success of all these approaches, however, is dependent on knowledge of the protective choices and appropriate behaviors.

The incidence of neural tube defects has decreased since 1992 when the first preconception recommendations aimed at prevention were advanced (CDC, 2004b, 2008; Honein, Paulozzi, Mathews, Erickson, & Wong, 2001; Williams et al., 2002; Williams, Rasmussen, Flores, Kirby, & Edmonds, 2005), but a very recent report released by the CDC after the initial literature search for this chapter suggests progress has slowed (CDC, 2008). The spina bifida rate per 100,000 live births declined 25% from 1995 to 2000 and 13% from 2000 to 2005. The anencephaly rate declined 36% from 1991 to 1995 and was unchanged from 1995 to 2005 (CDC, 2008). Interpreting the data to determine what strategies may have resulted in these improvements or why the impact has been less than forecast is challenging.

Evidence for the Effectiveness of Preconceptional Folic Acid Interventions

Searches of Medline, CINAHL, and the reference lists of identified articles were undertaken in the summer of 2006. Search strategies and key words were chosen to locate reports which could illuminate the impact of preconceptional strategies on maternal folate levels, variations in risk

between subpopulations, and whether and how various strategies differentially impact racial and ethnic subgroups. Studies from 2000 to 2006 were considered with background literature from prior to 2000 also cited.

The three main population-based preconceptional strategies to increase folate levels of women prior to pregnancy with the ultimate aim of reducing NTD rates have been: fortification, supplementation, and changes in dietary choices. No studies were located that were able to assess the individual contribution of each of these strategies to overall reductions in neural tube defects. Only one study was found that reported simultaneous examination of dietary choices and intake of fortified foods while controlling for use of supplementation in women of various racial and ethnic backgrounds (Dietrich, Brown, & Block, 2005). Several studies examined one or two of the avenues for protection but all of the studies relied on survey or case-control methodology making generalizability difficult.

Determining the effectiveness of all preconceptional folic acid strategies for reducing the occurrence of NTDs is complicated by the fact that surveillance frequently relies on birth certificate data rather than prenatal ascertainment. Three limitations of birth certificate data are that they do not include spontaneous or elective abortions, fetal deaths are not recorded, and birth defects are generally under-reported. Many of these shortcomings can be addressed through prenatal ascertainment of anomalies but this is not consistently achieved across surveillance programs. In a study by Williams and colleagues (2005), only 9 of the 21 birth defect surveillance systems used to track prevalence of NTDs across time were able to ascertain prenatally diagnosed cases, leading them to conclude that some of the observed differences in rates of NTDs among racial/ethnic subgroups may be attributable to differences in the frequency of prenatal diagnosis and elective termination of fetuses with NTDs. A review of studies published in the US and Canada reported a strong correlation between the completeness of case ascertainment and the percentage decrease in NTD rates (Mills & Signore, 2004).

Despite these limitations, these data sources are currently the best available for examining changes in NTD prevalence over time. A review of the effectiveness of the three main strategies (fortification, supplementation recommendations, dietary choices) aimed at changing outcomes such as preconceptional folate levels, folic acid supplement use, folic acid awareness, and reducing NTDs is provided below (and in Tables 5.4 and 5.5).

Strategy #1: Impact of fortification on preconceptional folate levels and/or NTDs. The National Health and Nutrition Examination Surveys (NHANES) provide an avenue to explore the impact of fortification on the folate levels of adults in the United States. The Centers for Disease Control and Prevention periodically conducts NHANES to assess the health, dietary practices, and nutritional status of non-institutionalized civilians in the United States. Subjects are chosen using a stratified multistage probability design representing people from 2 months of age and older. NHANES III included nearly 17,000 people over the age of 18 and NHANES 1999–2000 included nearly 4,500 people over the age of 19 years. The surveys include physical examinations, laboratory assessments, and dietary interviews. By comparing the data from NHANES III which was conducted between 1988 and 1994 (representing the time before folic acid fortification which was mandated in 1998) and NHANES 1999–2000 (representing the time after fortification) some assessments of impact can be offered. Two studies that used these data are summarized below and further described in Table 5.4. In 2002, the CDC (2002) analyzed the NHANES data from the two time frames and found that the median serum and erythrocyte (red blood cell, or RBC) folate concentrations (a better measure of long term folate status) increased significantly for women ages 15–44 following fortification for all three racial/ethnic groups examined: non-Hispanic whites, non-Hispanic blacks, and Mexican-Americans. Despite significant increases for all groups, disparities in serum and erythrocyte folate levels persisted between the races, with non-Hispanic white women having the highest folate levels and non-Hispanic black women having the lowest. The CDC continues to conduct NHANES surveys annually, which provides a mechanism for monitoring trends in serum and erythrocyte folate levels.

Table 5.4 Outcomes associated with studies of three population-based folic acid intervention strategies

Author (year)	Study design/ study type	Description of outcome: what, how, where	Populations studied (ages included, race and ethnicity) and sample size	Address disparities (yes/no)	Key findings related to intervention effectiveness	Caveats and biases	Findings support the intervention? (yes/no); For which populations?
Strategy #1: Fortification of cereal grains[a,b]							
Honein et al. (2001)	Time trend – multiple points in time with no concurrent comparison group/ published article	NTD prevalence in United States from 1990 to 1999	All births reported on birth certificates in 45 states and Washington, DC between January 1990 and December 1999, n = 35,165,226	No	Prevalence of spina bifida significantly decreased 23% between periods of October 1995 to December 1996 and October 1998 to December 1999 Prevalence of total NTDs significantly decreased 19% during the same time frame	Questions about validity of birth certificate data, particularly with respect to NTD reporting	Yes – for the general US population, but same decrease did not occur for infants of women receiving only third-trimester or no prenatal care
CDC (2002)	Pre-/post-intervention population comparison – no concurrent comparison group/published article	Blood serum and RBC folate levels as determined by assays of blood samples during 1988–1994 (pre-fortification) or 1999–2000 (post-fortification)	Non-Hispanic white, non-Hispanic black, and Mexican-American women ages 15–44 who participated in NHANES 1988–1994 (n = 5,616) and NHANES 1999–2000 (n = 1,656)	Yes – stratifies by race/ ethnicity	Median serum and RBC folate levels (ng/mL) significantly increased post-fortification for all racial/ethnic groups examined White women have the highest serum and RBC folate levels, while black women have the lowest	NHANES 1999–2000 has smaller sample size	Yes – all racial/ ethnic groups have benefited through increased folate levels, but disparities in folate levels persist
Williams et al. (2002)	Time trend – three time points with no concurrent comparison group/published article	NTD prevalence in United States from 1995 to 1999	All births reported to Neural Tube Defect Surveillance/ Folic Acid Education Committee of the National Birth Defects Prevention Network from January 1995 to December 1999 – 24 states participated, n = 8,596,835	No	Prevalence of spina bifida significantly decreased 31% from pre- to mandatory-fortification periods, prevalence of anencephaly significantly decreased 16% from pre- to mandatory-fortification periods		Yes – but greater decreases were seen in surveillance systems with prenatal ascertainment compared to those without prenatal ascertainment of NTDs

(continued)

Table 5.4 (continued)

Author (year)	Study design/ study type	Description of outcome: what, how, where	Populations studied (ages included, race and ethnicity) and sample size	Address disparities (yes/no)	Key findings related to intervention effectiveness	Caveats and biases	Findings support the intervention? (yes/no); For which populations?
CDC (2004b)	Pre-/post-intervention population comparison – no concurrent comparison group/published article	Spina bifida and anencephaly prevalence and estimated numbers in United States for the years 1995–1996 (pre-fortification) and 1999–2000 (post-fortification)	Live births, stillborns, fetal deaths, and elective terminations in 22 population-based surveillance systems (8 with prenatal ascertainment and 15 without prenatal ascertainment)	No	Total numbers of spina bifida and anencephaly-affected pregnancies decreased 27% from the pre-fortification period to the post-fortification period Total numbers of spina bifida and anencephaly-affected births decreased 26% from the pre- to post-fortification period	No tests for statistical significance, no controlling for confounding	Yes
Dietrich et al. (2005)	Pre-/post-intervention population comparison – no concurrent comparison group/published article	Blood serum and RBC folate levels as determined by assays of blood samples during 1988–1994 (pre-fortification) or 1999–2000 (post-fortification)	Men and non-pregnant women ages 20+ who are not folic acid supplement users and participated in NHANES 1988–1994: Men n = 5,075, women n = 4,844 (women ages 20–39 n = 2,260) or NHANES 1999–2000: Men n = 1,146, women n = 975 (women ages 20–39 n = 356)	No	Both serum and RBC folate levels increased significantly for all sex and age groups and the overall prevalence of inadequate blood folate status decreased for the general US population post-fortification Serum folate increased 153% for women ages 20–39 and RBC folate increased 63% for women ages 20–39	Excludes women younger than age 20, does not divide subjects into groups conducive to looking at all women of childbearing age	Yes – fortification has improved median serum and RBC folate for the adult US population, yet less than 10% of women reached an RBC folate level associated with a decreased risk of NTDs

Williams et al. (2005)	Time trend – three time points with no concurrent comparison group/published article	NTD prevalence in United States from 1995 to 2002	All births reported to Neural Tube Defect Surveillance/ Folic Acid Education Committee of the National Birth Defects Prevention Network from January 1995 to December 2002 – 21 states participated, Hispanic n = 2,653,365, non-Hispanic white n = 6,694,699, non-Hispanic black n = 1,730,343	Yes	Spina bifida prevalence significantly decreased 36% for Hispanics and 34% for non-Hispanic whites from pre- to mandatory-fortification periods. 19% decrease for non-Hispanic blacks neared, but did not reach significance (RR = 0.81 [0.67–1.00]). Anencephaly prevalence significantly decreased 26% for Hispanics and 29% for non-Hispanic whites from pre- to mandatory-fortification periods. There was no significant decrease for non-Hispanic blacks.		Yes – for Hispanics and non-Hispanic whites only
Strategy #2: Folic acid supplementation recommendations							
Ahluwalia and Daniel (2001)	Time trend/ published article	Awareness of folic acid	PRAMS participants (women who had delivered a live infant in the last 2–6 months) from 13 states during the years 1995–1998, N = 58,625 births	Yes	Folic acid awareness generally increased from 1995 to 1998 with the largest increase between 1996 and 1997. Hispanic women experienced the greatest increase in awareness during the time period Compared to white and non-Hispanic women, women of other races and ethnicities were more likely to lack awareness of folic acid Black (vs. white): OR = 1.54 [1.29, 1.84] Other non-white (vs. white): OR = 1.50 [1.10, 2.05] Hispanic (vs. non-Hispanic): OR = 1.29 [1.01, 1.65]	Data collected from women after they have a live birth – does not examine preconceptional knowledge	Yes – all populations have experienced increases in folic acid awareness, though disparities in level of awareness still exist between racial/ethnic groups

(continued)

Table 5.4 (continued)

Author (year)	Study design/ study type	Description of outcome: what, how, where	Populations studied (ages included, race and ethnicity) and sample size	Address disparities (yes/no)	Key findings related to intervention effectiveness	Caveats and biases	Findings support the intervention? (yes/no); For which populations?
Chacko et al. (2003)	Cross sectional/ published article	Awareness of folic acid, supplement use	Black non-Hispanic and Hispanic women ages 13–22 who sought services between 1999 and 2000 at one of three reproductive health clinics in Texas	Yes	There was no statistically significant difference between the racial/ ethnic groups in awareness of folic acid (Blacks: 52%, Hispanics: 52%) or awareness that birth defects can be prevented by multivitamins (Blacks: 49%, Hispanics: 55%) Only 9% of the young women reported taking a multivitamin daily (results not presented by race/ethnicity)	No time trend to assess changes in awareness due to recommendations	No – awareness of folic acid among young minority women is similar to 1998 Gallup poll of women of childbearing ages, but supplement use is very low in this population
de Jong-van den Berg et al. (2005)	Time trend – multiple points in time with no concurrent comparison group/ published article	Awareness of folic acid, Supplement use	All mothers who delivered infants and participated in the Slone Epidemiology Center Birth Defects Study, a case-control surveillance program with hospital centers in Boston, Philadelphia, and Toronto, n = 16,555 from 1988 to 2002; Malformed infants n = 12,370, not malformed infants = 4,185	Yes	Women aware of folic acid increased from 0% in 1988 to ~50% in 1996, with a sharp increase from 1992 to 1993	Most subjects may not represent general population because they were the mothers of malformed infants; responses collected after birth, so reflected knowledge may have been gained subsequent to the pregnancy	Yes – there appears to be an increase in folic acid awareness and use over time

					Compared to white women, all other racial/ethnic groups were significantly less likely to be aware of folic acid, after adjusting for confounding Black: OR = 0.38 [0.31, 0.48] Asian: OR = 0.24 [0.19, 0.31] Hispanic: OR = 0.41 [0.34, 0.51] Other: OR = 0.34 [0.24, 0.47] Use of folic acid in the periconceptional period increased from 15% in 1988 to ~40% in 2002, with a steep rise beginning in 1994 Compared to white women, all other racial/ethnic groups were significantly less likely to use folic acid, after adjusting for confounding Black: OR = 0.62 [0.48, 0.81] Asian: OR = 0.43 [0.33, 0.57] Hispanic: OR = 0.57 [0.45, 0.72] Other: OR = 0.52 [0.36, 0.76]	Disparities still remain between racial/ethnic groups in folic acid awareness and use. Non-Hispanic whites have the highest rates of folic acid awareness and use

(continued)

Table 5.4 (continued)

Author (year)	Study design/ study type	Description of outcome: what, how, where	Populations studied (ages included, race and ethnicity) and sample size	Address disparities (yes/no)	Key findings related to intervention effectiveness	Caveats and biases	Findings support the intervention? (yes/no); For which populations?
Green-Raleigh et al. (2006)	Time trend – multiple points in time with no concurrent comparison group/ published article	Awareness of folic acid, Supplement use	National sample of approximately 2,000 women aged 18–45 years from 1995 to 2005	Yes	Between 1995 and 2001, women reporting having heard or read about folic acid increased from 52 to 79% and, between 1995 and 2005, women reporting taking a daily folic acid supplement increased from 28 to 33% Those more likely to be aware of folic acid and to daily take a folic acid supplement were white, Non-Hispanic, older, and had higher education levels and household incomes	No tests for statistical significance, no control for confounding factors	Yes – awareness and use of folic acid increased over time corresponding to the IOM folic acid consumption recommendation. However, racial/ ethnic disparities in folic acid awareness and vitamin use persist
Strategy #3: Consumption of folate-rich foods							
Suarez et al. (2000)	Case-control/ published article	NTD risk (anencephaly, spina bifida, or encephalocele)	Cases were Mexican-American mothers of infants or fetuses with an NTD identified at birth or prenatally between January 1995 and February 1999 in Texas, n = 132	No	After adjustment for age, education, obesity, and previous stillbirth or miscarriage, cases were less likely than controls to consume levels of dietary folate in the upper two quartiles First quartile: referent Second quartile: OR = 1.06 (0.54, 1.06)	Recall bias – interviews were conducted after birth or termination of pregnancy	Yes – trends hint that dietary intake of folic acid is important, but results are limited by low power to detect significant differences

	Controls were Mexican-American women with normal births during the same time period who were residents in the study area, n = 150	Third quartile: OR = 0.53 (0.25, 1.09) Fourth quartile: OR = 0.97 (0.49, 1.91) Cases were less likely to consume above 400 μg folate (from supplements and diet combined) than controls <400 μg: referent 400–999 μg: OR = 0.96 (0.54, 1.71) >1,000 μg: OR = 0.73 (0.31, 1.72)

[a]Since this chapter was initially written, new NHANES data were released which suggest that median RBC folate levels of the US population increased between 2003–2004 and 2005–2006 after declining between 2001–2002 and 2003–2004. However, the median RBC folate level of non-Hispanic black women was significantly lower than that of non-Hispanic white and Mexican-American women; the latter two groups had similar levels. McDowell et al. (2008)

[b]Since the chapter was prepared, a new report was released by the CDC which shows that from the earliest post-fortification period 1999–2000, to the most recent surveillance period, 2003–2005, the prevalence of births complicated by spina bifida in the United States decreased only 6.9%. Among non-Hispanic black mothers, the prevalence decreased 19.8%; no significant decrease was observed for infants with non-Hispanic white and Hispanic mothers. Contrary to previous reports, the prevalence was similar for infants born to Hispanic and non-Hispanic white mothers. Centers for Disease Control (2009)

Table 5.5 Quality rating of studies associated with population-based folic acid interventions

Outcome of interest	Author (year)	Reporting	External validity	Internal validity – bias	Internal validity – confounding	Power	Total quality score ≤14 = poor 15–19 = fair ≥20 = good	Suitability of study to assess effectiveness: greatest, moderate, least
Strategy #1: Fortification of cereal grains								
Serum and RBC folate levels	CDC (2002)	6	2	4	2	0	14	Least
Serum and RBC folate levels	Dietrich et al. (2005)	11	3	6	3	0	23	Least
NTD prevalence	Honein et al. (2001)	8	4	5	3	1	21	Moderate
NTD prevalence	Williams et al. (2002)	8	4	5	3	0	20	Moderate
NTD prevalence	CDC (2004b)	6	3	3	4	0	16	Least
NTD prevalence	Williams et al. (2005)	9	4	5	4	0	22	Moderate
Strategy #2: Folic acid supplementation recommendations								
Awareness	Ahluwalia and Daniel (2001)	6	2	2	3	0	13	Least
Awareness, Supplement use	Chacko et al. (2003)	11	2	4	3	0	20	Least
Awareness, Supplement use	de Jong-van den Berg et al. (2005)	10	3	5	3	0	21	Least
Awareness, Supplement use	Green-Raleigh et al. (2006)	5	3	3	3	0	14	Least
Strategy #3: Consumption of folate rich foods								
NTD risk	Suarez et al. (2000)	10	3	4	3	0	20	Least

[Since this chapter was initially written, new NHANES data were released which suggest that median RBC folate levels of the US population increased between 2003–2004 and 2005–2006 after declining between 2001–2002 and 2003–2004. However, the median RBC folate level of non-Hispanic black women was significantly lower than that of non-Hispanic white and Mexican-American women; the latter two groups had similar levels (McDowell et al., 2008).]

Dietrich and colleagues (2005) reported that mandatory folic acid fortification led to significant increases in both serum and erythrocyte folate concentrations in all sex and age groups from age 20 to 60+ years but found that less than 10% of women of childbearing age reached the recommended erythrocyte folate concentration for protection against NTDs. In their analysis they excluded all survey participants who had used any nutritional supplement in the 30 days prior to their laboratory testing. The researchers concluded that fortification at the presently mandated level was probably not sufficient to prevent NTDs.

Several studies have attempted to examine the impact of fortification on NTD rates by examining the changes in NTD rates over time (see Table 5.4). Honein and colleagues (2001) utilized birth certificate date to examine the changes in spina bifida and total NTD rates from the prefortification (October 1995 to December 1996) to mandatory fortification (October 1998 to December 1999) periods. They found that the prevalence of spina bifida decreased 23% and the total NTD prevalence decreased 19% between these two time periods. Significant decreases were not seen, however, among infants born to women who received only third trimester or no prenatal care. An analysis of surveillance system data by the CDC (2004b) similarly found that the total numbers of spina bifida and anencephaly affected pregnancies decreased 27% from pre- to mandatory fortification, and the total numbers of spina bifida and anencephaly affected births decreased 26% over the time period. [Since the chapter was prepared, a new report was released by the CDC (2009) which shows that from the earliest post-fortification period, 1999–2000, to the most recent surveillance period, 2003–2005, the prevalence of births complicated by spina bifida in the United States decreased only 6.9%. Among non-Hispanic black mothers, the prevalence decreased 19.8%; no significant decrease was observed for infants with non-Hispanic white and Hispanic mothers. Contrary to previous reports, the prevalence was similar for infants born to Hispanic and non-Hispanic white mothers.]

In a study using data from 24 population-based birth defects surveillance systems in the NTD Ascertainment Project, Williams and colleagues (2002) examined the rates of spina bifida and anencephaly from 1995 to 1999. There were three time periods defined: “prefortification” from January 1995 to December 1996, “optional fortification” from January 1997 to September 1998, and “mandatory fortification” from October 1998 and on. They determined that the prevalence of spina bifida and the prevalence of anencephaly significantly decreased 31 and 16%, respectively, from the pre- to mandatory fortification periods. Williams and colleagues (2005) later attempted to determine the impact of food fortification on specific racial/ethnic groups. Between the “prefortification” and “mandatory fortification” periods, they found that the prevalence of spina bifida decreased 36% among Hispanic births (CI: 0.56–0.74) and 34% among non-Hispanic white births (CI: 0.63–0.80). No statistically significant decline, however, was observed among non-Hispanic black births. Considering these data with the above information, it appears that cereal grain folic acid fortification may have not benefited all women equally depending on the time period considered.

Strategy #2: Impact of supplementation recommendations on folic acid supplement use and awareness. The prevention strategy that has been promoted the longest is that women take exogenous folic acid as a supplement to their dietary intake. This recommendation, as previously noted, was first put forth by the U.S. Public Health Service in 1992 (CDC, 1992) and was reaffirmed by the Institute of Medicine in 1998 (Green-Raleigh, Carter, Mulinare, Prue, & Petrini, 2006). There were also several national campaigns during the early to mid-1990s that aimed to promote folic acid awareness, such as the March of Dimes “Think Ahead” campaign, as well as many state and local advertising campaigns (Ahluwalia & Daniel, 2001). Several reports, summarized in Table 5.4, have reported trends in awareness and use of folic acid supplementation.

The CDC analyzed data collected from 13 states between 1995 and 1998 through the Pregnancy Risk Assessment Monitoring System (PRAMS) to assess trends in folic acid awareness. This study found increasing awareness of folic acid across racial and ethnic groups over the 3 years with the greatest increase occurring in the Hispanic population. Data indicated no substantial differences in awareness among women of different age groups, parity, insurance coverage or WIC participation (Ahluwalia & Daniel, 2001).

In 2003, Chacko, Anding, Kozinetz, Grover, and Smith published a study aimed at assessing the awareness of folic acid prevention of NTDs and the frequency of ingestion of folate rich foods, foods fortified with folic acid, and folic acid supplements. The study population included English-speaking, low-income, women ages 13–22 years who presented to reproductive health clinics supported by public funding. Information collected through a self-administered questionnaire was analyzed and reported for 286 black non-Hispanic and 112 Hispanic women. There was no significant difference between the two groups regarding knowledge about folic acid and its use to prevent NTDs. Approximately half of the young women, irrespective of race or ethnicity, had heard of folic acid and that taking multivitamins before pregnancy can provide protection. Daily multivitamin intake was reported by only 9% of the total population, and only 7% of those taking a multivitamin reported doing so "to prevent birth defects."

From 1988 to 2002, women participating in the Slone Epidemiology Center Birth Defects Study were interviewed within 6 months of giving birth in a case-control surveillance study investigating various maternal exposures from 2 months before conception through delivery (de Jong-van den Berg, Hernandez-Diaz, Werler, Louik, & Mitchell, 2005). The database includes the mothers of 12,370 infants with congenital anomalies and 4,185 mothers of infants who were not malformed. Exploration specifically investigated the use of multivitamins containing folic acid and folic acid as a single supplement. Women were considered to be aware of the benefits of folic acid if they volunteered during the interview that it could reduce the risk of birth defects. Analyses indicate that awareness of the importance of folic acid supplementation was approximately zero in 1988 and that it began to increase sharply in 1992 reaching approximately 50% in 1996. Supplementation was relatively stable at 15–20% between 1988 and 1994 and then increased to approximately 40% thereafter. Over the entire study time period, white non-Hispanic women were significantly more likely to be aware of and to use folic acid than all other racial and ethnic groups. However, this study is limited by potential recall bias associated with retrospective assessments. In addition, because all of the women have experienced pregnancy, it is possible that their knowledge about folic acid was impacted by prenatal care thereby obscuring their ability to recall what they knew and did prior to conception.

In an effort to overcome such biases, the March of Dimes Foundation engaged the Gallup organization to conduct a series of surveys to understand trends in folic acid awareness and behaviors. Nine random digit telephone surveys were conducted with approximately 2,000 women ages 18–45 per survey year between 1995 and 2005. Analysis of the trends show that the proportion of women reporting they were aware of folic acid increased from 52% in 1995 to 84% in 2005. The percentage of women reporting that folic acid prevents birth defects increased from 4% in 1995 to 19% in 2005 and the percent of non-pregnant women who reported that they took the vitamin daily increased from 25 to 31% over the same period of time. This study also showed increasing percentages of women of childbearing age reporting awareness of folic acid in each race/ethnic category between 2000 and 2005 with the greatest increase occurring in the Hispanic population. In 2005, of women between the ages of 18–45, 87% of whites, 71% of non-whites, and 73% of Hispanics indicated knowledge of folic acid. The percent of women in the same age range who reported taking folic acid daily showed a less positive result: between 1998 and 2005, the white population increased 3% points to 36% daily use; the non-white population decreased 3% points to 23% and the Hispanic population decreased 2% points to 27%. The margin of error was ±3% (Green-Raleigh et al., 2006).

Strategy #3: Impact of dietary choices on folate levels and/or NTDs. No study was found that evaluated the impact of dietary choices on folate levels while controlling for both supplementation and fortification. Such studies under universal fortification would be difficult. The literature in this area tended to focus on describing the dietary choices of women, but did not provide links to their pregnancy outcomes. There was only one study that examined the impact of dietary choices on NTDs and no studies were found that examined the impact on maternal folate levels.

Suarez and colleagues used a case-control approach in 2000 to examine supplementation and dietary patterns in Mexican-American women experiencing a pregnancy complicated by an NTD compared to those whose pregnancies were not affected (see Table 5.4). The studied populations resided on the Texas-Mexican border, which is known to have the highest rates of NTDs in the United States. A 98-item food questionnaire was used to assess dietary intake of folic acid of 132 women whose pregnancies were complicated by an NTD and 150 control women. While the results were not statistically significant, women without an NTD-affected pregnancy were more likely to consume levels of folic acid in the upper two quartiles than women with an NTD-affected pregnancy after adjustment for age, education, obesity, and previous stillbirth or miscarriage. When total folate intake (from both diet and supplements) was considered, case women were less likely to have an average daily folate intake over 400 μg than control women. These results, however, were not statistically significant and diminished after controlling for the factors previously mentioned. The lack of significant differences in this study could be due to the small sample size in both the case and control groups. The trends of the data suggest that dietary intake and ingestion of fortified foods could be important at preventing NTDs, especially in women who do not use supplements.

Public Health Implications for Reducing Racial/Ethnic Disparities in Neural Tube Defects

Numerous studies have demonstrated that the prevalence of NTDs in this country varies by race/ethnicity with the highest rates occurring among women of Hispanic ethnicity and the lowest rates among Asian and non-Hispanic black women (Feuchtbaum et al., 1999; Hendricks, Simpson, & Larsen, 1999; Ray, Vermeulen, Meier, Cole, & Wyatt, 2004; Shaw, Velie, & Wasserman, 1997; Williams et al., 2005). Reports indicate that Hispanic rates are as much as 50% higher than non-Hispanic rates. [As stated earlier, following the completion of this chapter, a report was released that contradicts other epidemiologic and surveillance data regarding the racial and ethnic prevalence of NTDs (CDC, 2009). This study, analyzing the impact of fortification, found no difference in the prevalence of NTDs in the offspring of non-Hispanic white and Hispanic mothers; why this analysis of national Vital Statistics System data found no difference is unclear. The marked inconsistency of these data with other reports calls for restraint in interpretation.]

According to Williams and colleagues (2005), reasons for inequalities in prevalence are unknown but some of the disparity may reflect differences in genetic factors, such as the genes associated with folate metabolism. Several studies have shown that individuals who have pregnancies complicated by NTDs, as well as their affected offspring, are more likely to carry a mutation in a gene critical to converting homocysteine to methionine than parents of unaffected pregnancies and their offspring (ACOG, 2003); interruption in this conversion process is believed to play a role in the formation of NTDs. Studies have shown that the frequency of genes implicated in faulty conversion increases among racial/ethnic groups in the same pattern as the frequency of NTDs: highest in people of Hispanic ethnicity, intermediate in non-Hispanic whites, and lowest in non-Hispanic blacks (Williams et al., 2005).

In contrast to the genetic hypothesis, reviewing epidemiologic evidence, Suarez and colleagues (2000) argue that the NTD risk pattern for women of Mexican heritage most likely represents an

environmental rather than genetic cause because Mexican women migrating to the United States have an intermediate risk for NTDs. Supporting this argument is a population based case-control study of women giving birth in California (Velie et al., 2006). While Mexican immigrants were found to have a sevenfold increased risk of an NTD-affected fetus compared to the white non-Hispanic population, the researchers found the crude NTD risk for Mexican-born women compared with non-Hispanic white women to be most pronounced in those who reported moving to the US less than 2 years before conception of the index pregnancy (OR 7.2; CI 3.7–14.0) and in women who were older than 16 when they migrated to the US (OR 3.0; CI 2.0–4.3). In this study, NTD risks for second and third generation Mexican-American women were not found to be elevated when compared with non-Hispanic white women. The researchers found shorter stature among Mexico-born women associated with the increased NTD risk, and postulated that the mothers of affected children may have suffered nutritional or other insults to their own health in utero and childhood which resulted in increased risk for NTDs in their offspring. Migration at an earlier age and increased time in the U.S. may have ameliorated the impact of early biologic insults.

Evidence addressing the above environmental hypothesis may come from examining results of the first major preconceptional folic acid strategy to reduce NTDs: cereal grain fortification. It may be that all racial/ethnic groups are not equally exposed to fortified foods due to cultural preferences. Grain products in the traditional diets of the Hispanic population include masa and maize; because these foods are not commonly enriched, they fall outside the federal mandate for fortification with folic acid. For Hispanics, it may be that exposure to fortified foods and their associated protection increases with acculturation. A recent systematic review determined that little is known about the likelihood of obtaining adequate folate to prevent neural tube defects solely through food consumption, particularly in the diets of ethnic minorities (Lumley et al., 2001).

As mentioned earlier, Williams and colleagues (2005) attempted to determine the impact of food fortification on specific racial/ethnic groups. They found that the prevalence of spina bifida decreased among Hispanic births and non-Hispanic white births between the “prefortification” and “mandatory fortification” periods, but no significant decline was observed among non-Hispanic black births. These findings raise many questions about the foundations of the previously reported disparities. For instance, if the NTD prevalence differences are related to variation in genetic predisposition, why did the same amount of fortification impact Hispanic and non-Hispanic white births almost equally, and why were the rates for the non-Hispanic black population initially resistant to the population-based strategy of fortification? On the other hand, if NTD prevalence differences are related to cultural food preferences that influence exposure to fortified foods, why did Hispanics experience a significant decrease in NTDs directly after fortification, while non-Hispanic blacks did not? Does this analysis mask differences in intake between new immigrants and acculturated populations? According to Rader and Schneeman (2006), another fundamental question arises from such findings: is the amount of fortification required for prevention the same for all population subgroups? Further exploration of the disparate effects of fortification on racial/ethnic groups needs to be undertaken.

Another factor to be considered is that recent reports have demonstrated that the effect of fortification may be diminishing. An evaluation of NHANES data for 1999–2000, 2001–2002, and 2003–2004 by the CDC (2007), found a 16% decline in serum folate concentrations among women aged 15–44 from 1999–2000 to 2003–2004; RBC folate concentrations decreased 8% over the same time periods. Both of these findings were statistically significant. Unknown from these studies is whether what appeared to be a diminishing impact of fortification affected all racial/ethnic groups equally. [Since the initial literature search for this chapter was completed, new NHANES data were released which suggest that median RBC folate levels of the US population increased between 2003–2004 and 2005–2006. Of note is that the median RBC folate level of non-Hispanic black women was significantly lower than that of non-Hispanic white and Mexican-American women; the latter two groups had similar levels (McDowell et al., 2008.]

Differences in supplementation may also explain racial and ethnic disparities in NTDs. The two previously discussed studies which reported on trends in awareness and use of folic acid supplementation (de Jong-van den Berg et al., 2005; Green-Raleigh et al., 2006) analyzed data by subpopulations. The first study examined the experiences of women enrolled in the Slone Epidemiology Center Birth Defects Study from 1988 to 2002 (de John-van den Berg et al., 2005). Researchers found awareness of the benefits of folic acid was acknowledged by 60.1% of white women, 21.5% of black women, 23.2% of Hispanic women, and 26.9% of Asian women. In measuring the actual use of a folic acid supplement, 47.2% of white women reported use in the periconceptional period, compared to 15.6% of black women, 17.9% of Hispanic women, and 23.5% of Asian women. Even after adjusting for confounding factors, all minority racial/ethnic groups had decreased odds of folic acid awareness and supplement use. There were several strong independent predictors of both awareness and use, including maternal education, ethnicity, whether the pregnancy was wanted, family income, and whether a health care provider was consulted before conception.

The second study showed similar disparities in folic acid awareness and use by race/ethnicity. For all years 2000–2005, whites and non-Hispanics had higher levels of folic acid awareness and supplement use than non-whites and Hispanics, respectively. This is illustrated by the 2005 awareness levels of 87% among whites, 71% among non-whites, 73% among Hispanics, and 84% among non-Hispanics and the 2005 supplement use levels of 36% among whites, 23% among non-whites, 27% among Hispanics, and 34% among non-Hispanics (Green-Raleigh et al., 2006). Both of the studies just described have significant limitations: the Gallup survey is dependent on telephone access to women which may diminish the likelihood that poor or marginalized populations are adequately represented, and the Birth Defects Study is a retrospective study which may suffer from recall bias. Further, it is not known how factors beyond awareness of the benefits of folic acid may influence supplement use among different subpopulations. A review examining the influence of cultural practices, attitudes and beliefs on supplement use in women found scant data to describe the influences by culture or ethnicity (Jasti, Siega-Riz, & Bentley, 2003).

Finally, there is much left to be explained about the effect of food choices naturally rich in folate within racial/ethnic subpopulations and how these choices affect NTD rates. The previously detailed analysis of NHANES data conducted by Dietrich et al. (2005) briefly examined dietary choices among U.S. adults. The researchers report that before fortification the major source of dietary folate was vegetables (contributing 19.4% of total folate) but after fortification, the category "bread, rolls, and crackers" became the largest contributor (at 15.6%). The previously mentioned Chacko et al. (2003) study also considered dietary intake of folic acid among the study's adolescent and young adult participants by recording dietary intake through a food frequency recall for the previous week for 29 specific food items. This analysis found that the most frequently eaten folate-rich foods for Hispanic and non-Hispanic blacks were similar but further analysis of factors affecting food choices was not provided. In an unpublished Master's thesis, Callahan (1999) analyzed dietary intake for Hispanics and non-Hispanic black and white women through data from the 1994 to 1996 U.S. Department of Agriculture's Continuing Survey of Food Intakes by Individuals and the Diet and Health Knowledge Survey (CSFII/DHKS). Callahan found that, among women aged 19–44, Hispanic women consume more dietary folate than non-Hispanic white and black women. However, only 10% of Hispanic women and less than 8% of non-Hispanic white and non-Hispanic black women reported consumption of 400 μg of food folate daily. The proportion of that ingestion that was from foods rich in naturally occurring folate, or from fortified foods was not reported. In all these studies, no report of correlations between knowledge about folic acid, use of supplements, and food choices was offered, nor was this information linked to NTD rates.

The measurement of dietary folate intake does not address the question of whether women are aware of foods naturally high in folate and if they select food based on this awareness.

Surveys conducted by the Gallup Organization for the March of Dimes in 2005 reported that 26% of women of childbearing age who were aware of the benefits of folic acid identified green leafy vegetables as a good source of the nutrient, but only 5% identified fortified cereals as a good source (Green-Raleigh et al., 2006). This study does not link the data to actual food consumption, so there is still much to be learned about how a woman's awareness of dietary folate affect her food choices.

Summary

This chapter has examined two very different preconception interventions, both proven to reduce the risk for congenital anomalies. One of them, preconception diabetes control, is dependent on access to quality clinical health care; the other, increasing maternal folate levels, is not dependent on clinical health care but rather on awareness and personal action. No high quality research was found that reported on the differential impact of these preconceptional interventions in racial and ethnic subpopulations but this review highlights several lines of inquiry that may prove fruitful, namely: Are baseline risks between racial and ethnic groups different? Is there equitable access to information and/or interventions across subgroups? Are there differing responses to available information and interventions between groups?

An argument can be made that increasing efforts in this country to promote appropriate preconceptional care for women with preexisting medical conditions such as diabetes will, at least initially, disproportionately benefit the majority population and increase disparities. Reasoning behind this argument is that preconception care for chronic diseases is dependent on access to medical care, increasing the likelihood that those without access will be underserved by this prevention strategy. Because those with the greatest burden of disease and least likelihood to access preventive care are often women of color and ethnic minorities, they and their offspring may be disproportionately disadvantaged by services more readily available to a healthier subpopulation that is more able to access preventive services.

However, an argument can also be made that disparities could decrease – not because of improved health and access to care by minority populations, but because of worsening health and increasing barriers to preventive services by the majority population. For instance, the prevalence of obesity is increasing in all subgroups but the increase is particularly great in the white non-Hispanic population. Obesity is an important predictor of Type II diabetes; as obesity increases in population subgroups more women in those subgroups will need preconception interventions. In addition, non-Hispanic white women may increasingly find themselves uninsured and underinsured (National Center for Health Statistics, 2007; Schoen, Collins, Kriss, & Doty, 2008), decreasing the likelihood that they will be able to access the preconceptional care related to their increasing incidence of Type II diabetes. Thus the differences in burden of disease and the ability to access preventive services may narrow between subgroups, resulting in the potentially erroneous assessment that preconceptional interventions related to diabetes have no impact. The studies are clear that preconception care for *women with this disease* does make a difference, and the most important predictor of that difference may not be related to race or ethnicity but to access to appropriate preventive services.

In contrast, prevention of neural tube defects can be achieved without any medical encounter to initiate or monitor prescribed treatments. After more than a decade of government sponsored recommendations and population-based strategies to impact the occurrence of neural tube defects, almost nothing is known about whether those strategies are appropriate for all subpopulations, how they are impacting various groups, or how they will influence disparities. With mandatory fortification, a population-based prevention strategy has been implemented, assuring exposure to some level of

protection for all women who eat enriched foods. Little is known, however, about the efficacy of that strategy alone or in tandem with supplementation or a diet of naturally folate-rich foods. As well, there is minimal information about how the three avenues for increased folate levels have been differentially adopted by subpopulations. More complete data are needed to understand other factors contributing to existing disparities in NTDs. Do these disparities represent different environmental influences, different genetic risks, different rates of underlying diseases which are independently associated with NTDs, or differing access to or beliefs about pregnancy termination? Or alternatively, are these disparities a reflection of faulty ascertainment techniques or inadequate surveillance?

Preconceptional health promotion involves a future orientation – how that orientation is framed by subpopulations will be an important starting point for developing strategies that are accepted by and motivating to specific target groups. Until successful approaches are found which provide all subpopulations with the appropriate strategies to plan their reproduction and prevent the congenital anomalies associated with poor diabetes control and low folate levels, it is likely that those with the financial and personal resources, including the ability to plan and act for the future, will benefit disproportionately.

Burden of disease, access to preventive health services, and the degree to which individuals can process and act on recommendations that will impact a not-yet-conceived-child may prove to be the underlying determinants of disparities. How these three determinants interact will prove an important area for research as the preconception movement gains momentum.

With almost all new initiatives, unintended consequences surface. Some of these initiatives can prove enormously expensive, some will disproportionately burden subpopulations, and some will prove to be ineffective. The maternal and child health field is replete with examples of good intentions which were adopted as clinical best practices before adequate research had evolved to support the claims. For much of the preconceptional health agenda, sufficient research has not yet accumulated to argue that any one particular prevention approach will result in specific improvements in perinatal outcomes. Such is not the case, however, with diabetes control and increased folate levels, where the research on impact is clear. What is missing is translational research that can move the science and epidemiology into meaningful and acceptable strategies for women and the children they may someday conceive. This chapter serves to highlight that there are currently more questions than answers about moving the science of prevention to the practice of prevention, be it clinical, professional, or individual. While embracing the important opportunities of reframing the perinatal prevention paradigm to focus on primary prevention, we must pledge ourselves to be vigilant in quickly identifying and addressing the unintended consequences which will surely surface. If done well, this will prove to be public health's major contribution to narrowing disparities related to preconceptional health promotion strategies.

Conclusion

Research on the effect of preconception care on perinatal outcomes is only beginning to emerge in the literature. While there is evidence that preconceptional diabetes control decreases the risk of congenital malformations and that preconceptional folic acid interventions decrease the risk of neutral tube defects, there is little information available about the differential impact of these interventions on racial/ethnic subpopulations. It is still relatively unknown how an increasing emphasis on preconception care will impact racial/ethnic disparities in perinatal outcomes. The issues involved are complex, and conflicting theories exist about how preconceptional care could influence disparities. Only further research in this area will demonstrate the actual effect of preconceptional health promotion and preconceptional health care on perinatal outcomes.

References

Ahluwalia, I. B., & Daniel, K. L. (2001). Are women with recent live births aware of the benefits of folic acid? *Morbidity and Mortality Weekly Report*, *50* (RR-6), 3–14.

American Academy of Pediatrics & the American College of Obstetricians and Gynecologists (AAP & ACOG). (2007). *Guidelines for perinatal care* (6th ed.). Washington, DC: ACOG.

American College of Obstetricians and Gynecologists (ACOG). (2003). Neural tube defects. ACOG Practice Bulletin Number 44. *Obstetrics and Gynecology*, *102*(1), 203–213.

American College of Obstetricians and Gynecologists (ACOG). (2005). Pregestational diabetes mellitus. ACOG Practice Bulletin Number 60. *Obstetrics and Gynecology*, *105*(3), 675–685.

American Diabetes Association. (2004). Preconceptional care of women with diabetes. *Diabetes Care*,*27*(Suppl), S76–S78.

Atrash, H. K., Johnson, K., Adams, M. M., Cordero, J. F., & Howes, J. (2006). Preconception care for improving perinatal outcomes: The time to act. *Maternal and Child Health Journal*, *10*(5 Suppl), 3–11.

Botto, L. D., Moore, C. A., Khoury, M. J., & Erickson, J. D. (1999). Neural-tube defects. *New England Journal of Medicine*, *341*, 1509–1519.

Bowman, R. M., McLone, D. G., Grant, J. A., Tomita, T., & Ito, J. A. (2001). Spina bifida outcome: A 25-year prospective. *Pediatric Neurosurgery*, *34*, 114–120.

Callahan, C. G. (1999, June 21). *Prevention of neural tube defects: Assessment of folate status of Hispanic women in the United States*. Unpublished Masters thesis, UNC School of Public Health, pp. 1–29.

Casele, H. L., & Laifer, S. A. (1998). Factors influencing preconception control of glycemia in diabetic women. *Archives of Internal Medicine*, *158*, 1321–1324.

Centers for Disease Control. (1992). Recommendations for the use of folic acid to reduce the number of cases of spina bifida and other neural tube defects. *Morbidity and Mortality Weekly Report*,*41*(RR-14), 1–7.

Centers for Disease Control. (2002). Folate status in women of childbearing age, by race/ethnicity – United States, 1999–2000. *Morbidity and Mortality Weekly Report*, *51*(36), 808–810.

Centers for Disease Control. (2004a). (The Steering Committee of the National Public Health Initiative on Diabetes and Women's Health, Centers for Disease Control and Prevention). The National Public Health Initiative on Diabetes and Women's Health: Leading the way for women with and at risk for diabetes. *Journal of Women's Health*, *13*(9), 962–967.

Centers for Disease Control. (2004b). Spina bifida and anencephaly before and after folic acid mandate – United States, 1995–1996 and 1999–2000. *Morbidity and Mortality Weekly Report*, *53*(17), 362–365.

Centers for Disease Control. (2007). Folate status in women of childbearing age, by race/ethnicity – United States, 1999–2000, 2001–2002, and 2003–2004. *Morbidity and Mortality Weekly Report*,*55*(51), 1377–1380.

Centers for Disease Control. (2008). QuickStats: Spina bifida and anencephaly rates – United States, 1991, 1995, 2000, and 2005. *Morbidity and Mortality Weekly Report*, *57*(01), 15.

Centers for Disease Control. (2009). Racial/ethnic differences in the birth prevalence of spina bifida – United States, 1995–2005. *Morbidity and Mortality Weekly Report*, *57*(53), 1409–1413.

Chacko, M., Anding, R., Kozinetz, C., Grover, J., & Smith, P. (2003). Neural tube defects: Knowledge and preconceptional prevention practices in minority young women. *Pediatrics*, *112*, 536–542.

Cousins, L. (1991). The California Diabetes and Pregnancy Programme: A statewide collaborative programme for the pre-conception and prenatal care of diabetic women. *Baillière's Clinical Obstetrics and Gynaecology*, *5*443–59.

Czeizel, A. E., & Dudas, I. (1992). Prevention of the first occurrence of neural-tube defects by periconceptional vitamin supplementation. *New England Journal of Medicine*, *327*, 1832–1835.

de Jong-van den Berg, L., Hernandez-Diaz, S., Werler, M., Louik, C., & Mitchell, A. (2005). Trends and predictors of folic acid awareness and periconceptional use in pregnant women. *American Journal of Obstetrics and Gynecology*, *192*, 121–128.

Dietrich, M., Brown, C., & Block, G. (2005). The effect of folate fortification of cereal-grain products on blood folate status, dietary folate intake, and dietary folate sources among adult non-supplement users in the United States. *Journal of the American College of Nutrition*, *24*(4), 266–274.

Feuchtbaum, L. B., Currier, R. J., Riggle, S., Roberson, M., Lorey, F. W., & Cunningham, G. C. (1999). Neural tube defect prevalence in California (1990–1994): Eliciting patterns by type of defect and maternal race/ethnicity. *Genetic Testing*, *3*, 265–272.

Green-Raleigh, K., Carter, H., Mulinare, J., Prue, C., & Petrini, J. (2006). Trends in folic acid awareness and behavior in the United States: The Gallup organization for the March of Dimes Foundation surveys, 1995–2005. *Maternal and Child Health Journal*, *10*(5 Suppl), 177–182.

Harris, M. I. (2001). Racial and ethnic differences in health care access and health outcomes for adults with Type 2 diabetes. *Diabetes Care*, *24*(3), 454–459.

Heisler, M., Smith, D., Hayward, R., Krein, S., & Kerr, E. (2003). Racial disparities in diabetes care processes, outcomes, and treatment intensity. *Medical Care*, *41*, 1221–1232.

Hendricks, K. A., Simpson, J. S., & Larsen, R. D. (1999). Neural tube defects along the Texas-Mexico border, 1993–1995. *American Journal of Epidemiology*, *149*, 1119–1127.

Herman, W. H., Janz, N. K., Becker, M. P., & Charron-Prochownik, D. (1999). Diabetes and pregnancy: Pre-conception care, pregnancy outcomes, resource utilization and costs. *Journal of Reproductive Medicine*, *44*, 33–38.

Holing, E. V., Brown, Z. A., Beyer, C. S., & Connell, F. A. (1998). Why don't women with diabetes plan their pregnancies? *Diabetes Care*, *21*(6), 889–895.

Honein, M. A., Paulozzi, L. J., Mathews, T. J., Erickson, J. D., & Wong, L. Y. C. (2001) Impact of folic acid fortification of the US food supply on the occurrence of neural tube defects. *Journal of American Medical Association*, *285*, 2981–2986.

Institute of Medicine. (IOM). Committee to Study the Prevention of Low Birth Weight. (1985). *Low birth weight*. Washington, DC: National Academy Press.

Institute of Medicine. (IOM). Standing Committee on the Scientific Evaluation of Dietary Reference Intakes and Its Panel on Folate, Other B Vitamins, and Choline. (1998). *Dietary reference intake: Folate, other B vitamins and choline*. Washington, DC: National Academy Press.

Janz, N. K., Herman, W. H., Becker, M. P., Charron-Prochownic, D., Shayna, V. L., Jacober, S. J., et al. (1995). Diabetes and pregnancy: Factors associated with seeking preconception care. *Diabetes Care*, *18*(2), 157–165.

Jasti, S., Siega-Riz, A. M., & Bentley, M. E. (2003). Dietary supplement use in the context of health disparities: Cultural, ethnic and demographic determinants of use. *The Journal of Nutrition*, *133*(6), 2010S–2013S.

Johnson, K., Posner, S. F., Biermann, J., Cordero, J. F., Atrash, H. K., Parker, C. S., et al. (2006) Recommendations to improve preconception health and health care – United States: A report of the CDC/ATSDR Preconception Care Work Group and the Select Panel on Preconception Care. *Morbidity and Mortality Weekly Report*, *55*(RR06), 1–23.

Kim, C., Ferrara, A., McEwen, L. N., Marrero, D. G., Gerzoff, R. B., Herman, W. H., et al. (2005). Preconception care in managed care: The translating research into action for diabetes study. *American Journal Obstetrics and Gynecology*, *192*, 227–232.

Kitzmiller, J. L., Gavin, L. A., Gin, G. D., Jovanovic-Peterson, L., Main, E. K., & Zigrang, W. D. (1991). Pre-conception care of diabetes: Glycemic control prevents congenital anomalies. *Journal of the American Medical Association*, *265*, 731–736.

Korenbrot, C., Steinberg, A., Bender, C., & Newberry, S. (2002). Preconception care: A systematic review. *Maternal and Child Health Journal*, *6*(2), 71–73.

LeMaster, J. W., Chanetsa, F., Kapp, J. M., & Waterman, B. M. (2006). Racial disparities in diabetes-related preventive care: Results from the Missouri Behavioral Risk Factor Surveillance System. *Preventing Chronic Disease Public Health Research, Practice, and Policy*, *3*, 1–8.

Lumley, J., Watson, L., Watson, M., & Bower, C. (2001). Periconceptional supplementation with folate and/or multivitamins for preventing neural tube defects. *Cochrane Database of Systematic Reviews,* 2001, Issue 3. Art. No.: CD001056. DOI: 10.1002/14651858.CD001056.

Mathews, T. J., Honein, M. A., & Erickson, J. D. (2002). Spina bifida and anencephaly prevalence – United States, 1991–2001. *Morbidity and Mortality Weekly Report 51*(RR13), 9–11.

McDonnell, G. V., & McCann, J. P. (2000). Why do adults with spina bifida and hydrocephalus die? A clinic-based study. *European Journal of Pediatric Surgery*, *10*(Suppl), 31–32.

McDowell, M., Lacher, D., Pfeiffer, C., Mulinare, J., Picciano, M. F., Rader, J., et al. (2008) Blood folate levels: The latest NHANES results. NCHS Data Brief. No. 6. May 2008.

McElvy, S. S., Miodovnik, M., Rosenn, B., Khoury, J. C., Siddiqi, T., Dignan, P. S., et al. (2000). A focused preconceptional and early pregnancy program in women with Type I Diabetes reduces perinatal mortality and malformation rates to general population levels. *Journal of Maternal-Fetal Medicine*, *9*, 14–20.

Medical Research Council Vitamin Study Research Group. (1991). Prevention of neural tube defects: Results of the Medical Research Council Vitamin Study. *Lancet*, *338*, 131–137.

Mills, J. L., & Signore, C. (2004). Neural tube defect rates before and after food fortification with folic acid. *Birth Defects Research. Part A, Clinical and Molecular Teratology*, *70*(11), 844–845.

Moos, M. K. (2006). Preconception health: Where to from here? *Women's Health Issues*, *16*(4), 156–158.

Moos, M. K., & Cefalo, R. C. (1987). Preconceptional health promotion: A focus for obstetric care. *American Journal of Perinataology*, *4*(1), 63.

Mulinare, J., Cordero, J. F., Erickson, J. D., & Berry, R. J. (1988). Periconceptional use of multivitamins and the occurrence of neural tube defects. *The Journal of the American Medical Association*, *260*, 3141–3145.

National Center for Health Statistics. (2003). *Health, United States, 2003*. Hyattsville, MD: National Center for Health Statistics.

National Center for Health Statistics. (2007). *Health, United States, 2007*. Hyattsville, MD: Public Health Service.

National Diabetes Information Clearinghouse. (2005). National diabetes statistics. Retrieved December 20, 2006, from http://diabetes.niddk.nih.gov/dm/pubs/statistics/index.html.

Neuhouser, M. L., & Beresford, S. A. (2001). Folic acid: Are current fortification levels adequate? *Nutrition*,*17*, 868–872.

Ogden, C. L., Carroll, M. D., Curtin, L. R., McDowell, M. A., Tabak, C. J., & Flegal, K. M. (2006). Prevalence of overweight and obesity in the United States, 1999–2004. *The Journal of the American Medical Association*, *295*(13), 1549–1554.

Owens, M. D., Kieffer, E. C., & Chowdhury, F. M. (2006). Preconception care and women with or at risk for diabetes: Implications for community interventions. *Maternal Child Health Journal*, *10*, S137–S141.

Rader, J. I., & Schneeman, B. O. (2006). Prevalence of neural tube defects, folate status, and folate fortification of enriched cereal-grain products in the United States. *Pediatrics*, *117*, 1394–1399.

Ray, J. G., O'Brien T. E., & Chan, W. S. (2001). Preconception care and the risk of congenital anomalies in the offspring of women with diabetes mellitus. *QJM: Monthly Journal of the Association of Physicians*, *94*, 4435–4444.

Ray, J. G., Vermeulen, R. J., Meier, C., Cole, D. E., & Wyatt, P. R. (2004). Maternal ethnicity and risk of neural tube defects: A population-based study. *Canadian Medical Association Journal*, *171*, 343–345.

Rosenn, B., Miodovnik, M., Combs, C. A., Khoury, J., & Siddiqi, T. A. (1991). Pre-conception management of insulin-dependent diabetes: Improvement of pregnancy outcome. *Obstetrics and Gynecology*, *77*, 846–849.

Saurez, L., Hendricks, K. A., Cooper, S. P., Sweeny, A. M., Hardy, R. J., & Larsen, R. D. (2000). Neural tube defects among Mexican Americans living on the US-Mexico border: Effects of folic acid and dietary folate. *American Journal of Epidemiology*, *152*(11), 1017–1023.

Schoen, C., Collins, S. R., Kriss, J. L., & Doty, M. M. (2008). How many are underinsured? Trends among U.S. adults, 2003 and 2007. *Health Affairs*, *27*(4), w298–w309.

Sequist, T., Adams, A., Zhang, F., Ross-Degnan, D., & Ayanian, J. (2006). Effect of quality improvement on racial disparities in diabetes care. *Archives of Internal Medicine*, *166*, 675–681.

Shaw, G. M., Velie, E. M., & Wasserman, C. R. (1997). Risk for neural tube defect-affected pregnancies among women of Mexican descent and white women in California. *American Journal of Public Health*, *87*, 1467–1471.

The Diabetes Control and Complications Trial Research Group. (1996). Pregnancy outcomes in the diabetes control and complications trial. *American Journal of Obstetrics and Gynecology*, *174*, 1343–1353.

Thompson, J. E., Walsh, L. V., & Merkatz, I. R. (1990). The history of prenatal care: Cultural, social and medical contexts. In I. R. Merkatz & J. E. Thompson (Eds.), *New perspectives on prenatal care*. New York: Elsevier.

Velie, E. M., Shaw, G. M., Malcoe, L. H., Schaffer, D. M., Samuels, S. J., Todoroff, K., et al. (2006). Understanding the increased risk of neural tube defect-affected pregnancies among Mexico-born women in California: Immigration and anthropometric factors. *Paediatric and Perinatal Epidemiology*, *20*, 219–230.

Willhoite, M. B., Herman, W. H., Bennert, H. W., Williams, J. R., Palomaki, G. E., Spear, N. H., et al. (1993). The impact of preconception counseling on pregnancy outcomes. *Diabetes Care*, *16*, 450–455.

Williams, L. J., Mai, C. T., Edmonds, L. D., Shaw, G. M., Kirby, R. S., Hobbs, C. A., et al. (2002). Prevalence of spina bifida and anencephaly during the transition to mandatory folic acid fortification in the United States. *Teratology*, *66*, 33–39.

Williams, L. J., Rasmussen, S. A., Flores, A., Kirby, R. S., & Edmonds, L. D. (2005). Decline in the prevalence of spina bifida and anencephaly by race/ethnicity: 1995–2002. *Pediatrics*, *116*, 580–586.

Chapter 6
Infertility Status and Infertility Treatment: Racial and Ethnic Disparities

Sylvia Guendelman and Laura Stachel

Difficulties in conceiving or carrying a child to term affect 12% of the 62 million American women between ages 15–44 (Chandra, Martinez, Mosher, Abma, & Jones, 2005). Involuntary childlessness often leads to disappointment and despair, contributing to depression, marital strife and social stigma. Although infertility status, access to infertility care, and response to medical treatment are not distributed uniformly among women of reproductive age, few studies have examined the experiences of women of color. Evidence suggests important racial/ethnic disparities in the diagnosis and prevalence of infertility. Furthermore, the privatization of infertility clinical services and the high costs of these services have contributed to wide racial and socioeconomic inequalities in utilization of treatment.

Only 36% of infertile women seek infertility care despite increasing public awareness of infertility and a proliferation of U.S. infertility clinics (Chandra et al., 2005). Couples seeking care begin with infertility counseling and are offered a succession of diagnostic tests. Motivated couples with adequate emotional and financial resources proceed to infertility treatments such as gynecologic surgery, ovulation induction and artificial insemination. While many couples are able to conceive as a result of these treatments, some require assisted reproductive technology (ART) to achieve a pregnancy. In vitro fertilization (IVF) accounts for 99% of ART. This invasive procedure involves surgical extraction of eggs, manipulation of sperm and egg in a laboratory to effect fertilization, incubation of resulting embryos, and transfer of one or more embryos to the uterus. Unlike other reproductive interventions reviewed in this book that are directed at reducing high risk births, current evidence suggests that ART may actually increase the rate of preterm deliveries and multiple births (Reddy, Wapner, Rebar, & Tasca, 2007).

Infertility care is often expensive and is rarely covered by insurance. Women of color obtain fewer infertility interventions than white women, though they report higher rates of infertility (Stephen & Chandra, 2006). White, affluent, well-educated women are more likely to resort to IVF (Chandra et al., 2005), and thereby incur higher risks of preterm delivery, multiple births and adverse maternal outcomes. Ironically, the racial/ethnic differentials in IVF use may contribute to narrowing the racial gap in prematurity and low birth weight.

This chapter has several objectives. First, we review prevalence rates of infertility, impaired fecundity, and utilization of infertility services among racial and ethnic populations in the United States. Next, we review the effectiveness of IVF in producing live births and provide an evidence-based summary of reproductive responses to IVF by race and ethnicity. We focus on this highly technologic intervention because of its rapid growth in market share, its contribution to rising health care costs, and its association with adverse pregnancy outcomes and multiple births. Finally, we discuss public health implications of infertility treatment.

S. Guendelman (✉)
School of Public Health, University of California, Berkeley, CA 94720-7360, USA
e-mail: sylviag@berkeley.edu

A. Handler et al. (eds.), *Reducing Racial/Ethnic Disparities in Reproductive and Perinatal Outcomes: The Evidence from Population-Based Interventions*, DOI 10.1007/978-1-4419-1499-6_6,

Infertility Prevalence Rates

The National Survey of Family Growth (NSFG) is the main source of population data on infertility prevalence and infertility service utilization in the U.S. (Chandra et al., 2005). This survey uses specific demographic criteria for defining infertility and distinguishes "infertility" from "impaired fecundity." Information on these conditions is limited to 15–44 year old women. Infertility applies to *married or cohabiting* women, excluding those with surgical sterility,[1] who are unable to conceive after 12 or more consecutive months of unprotected intercourse.[2] Impaired fecundity is more broadly defined, and applies to women without surgical sterility of *any marital or cohabitation status* who have a history of problems conceiving or carrying a pregnancy to term (Chandra et al., 2005). This includes women who report that it is physically impossible, difficult or dangerous to carry a pregnancy to term, and women whose partners have fertility impairment. For either category, medical verification is not required.

In 2002, the most recent year for which NSFG data are available, 7.4% of married women, or 2.1 million women, were estimated to be infertile (Chandra et al., 2005). Advancing age increases the risk of infertility. Although the overall rate of infertility declined from 1982 to 1995 and held constant from 1995 to 2002, the absolute numbers of women reporting infertility has risen as more women, in particular the large cohort of aging "baby boomers," have delayed marriage and childbearing (Chandra & Stephen, 1998).

Infertility rates by race/ethnicity among married women in the 2002 NSFG survey are shown in Fig. 6.1. Black non-Hispanics reported higher infertility rates (11.5%) than white non-Hispanics (7.0%) or Hispanics of any race (7.7%).[3] A similar pattern was noted in an examination of pooled data from 1982, 1988, 1995, and 2002 NSFG cycles (Bitler & Schmidt, 2006). Infertility rates were inversely related to income and education. Considering that blacks and Hispanics tend to have lower income and education than whites in the U.S., the higher prevalence rates of infertility for women of color is not unexpected.

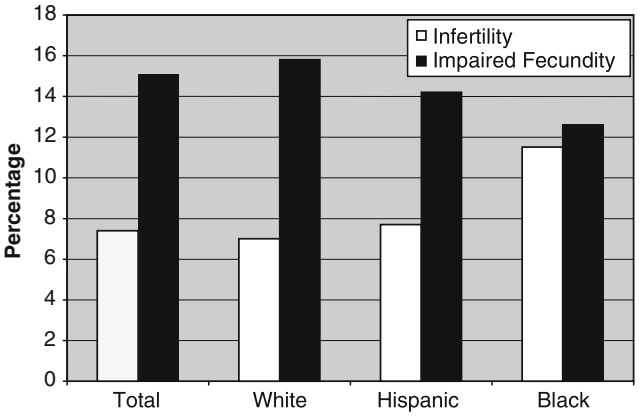


Fig. 6.1 Prevalence rates of infertility and impaired fecundity among married women aged 15–44: United States, 2002 [Source: 2002 National Survey of Family Growth (Chandra et al., 2005)]

[1] A woman is defined as surgically sterile if she or her current husband or cohabiting partner had an unreversed sterilizing operation, such as a tubal ligation or vasectomy.

[2] It should be noted that the NSFG constructs the outcomes "infertility" and "impaired fecundity" from responses to items such as marital status, sexual activity, contraceptive use, and sterility.

[3] Hereafter, black non-Hispanics, white non-Hispanics and Hispanics of any race will be referred to as blacks, whites and Hispanics. Asians are included in the NSFG in the "Other" category. The "Other" category combines women of several racial and ethnic origins and is not presented here.

Among married women, rates of impaired fecundity (i.e., difficulty conceiving or carrying a pregnancy to term) have been rising steadily since 1988 (Chandra et al., 2005). This increase may reflect changes in reporting and/or clinical care. For example, greater public awareness of fertility problems and increased social acceptability of infertility may contribute to increases in reporting. Likewise, advances in gynecologic practice, including the promotion of medical procedures that spare the removal of reproductive organs, have enabled some women with reproductive difficulties to retain reproductive capacity. For example, women nowadays may face impaired fecundity as a result of a myomectomy (a surgery to remove fibroids from the uterus), while in the past women with similar reproductive histories might have undergone a hysterectomy and been rendered sterile.

Among married women surveyed in the 2002 NSFG, rates of impaired fecundity were comparable for whites (15.8%) and Hispanics (14.2%) and slightly, but not significantly, lower for blacks (12.6%) (Fig. 6.1) (Chandra et al., 2005).[4] Although rates of impaired fecundity were higher for married/cohabiting women (a group more likely to be actively trying to conceive), the patterns across ethnic/racial groups remained similar when all women, regardless of marital status, were compared. However, when the sample was restricted to married women without children, women of color appeared to have higher rates of impaired fecundity (30.8% of Hispanics, 32.4% of blacks and 25.1% of whites).

Risk Factors for Infertility

There are numerous social, lifestyle, and health factors that contribute to infertility, including age-related patterns of childbearing, past exposure to sexually transmitted infections, environmental toxins, tobacco, alcohol and drug abuse, stress, strenuous exercise, poor nutrition, chronic disease and medical treatment for cancer (Office of Technology Assessment, 1988). In addition, cultural expectations for a large family size and prior fertility history can influence the recognition of secondary infertility [i.e., difficulty conceiving after one or both members of a couple have previously conceived a child (King & Davis, 2006; Kalmuss, 1987)].

Accurate assessment of factors contributing to infertility necessitates a complete medical evaluation for both partners. Causes of infertility can be traced to female factors approximately 30% of the time, male factors 30% of the time, and a combination of male and female factors 30% of the time (American Society of Reproductive Medicine, 2006). Approximately 10–20% of infertility is "unexplained," which implies that no specific abnormality can be detected in either partner. Infertility treatments are selected to address the specific cause of infertility. For example, ovulatory disorders may be treated with drugs to induce ovulation. Fallopian tube occlusion, on the other hand, may be treated with surgical repair or IVF. Descriptions of common causes of male and female infertility and their treatments are depicted in Table 6.1.

Population-based estimates of medical conditions causing infertility are difficult to ascertain. In the NSFG, prevalence rates of medical conditions are based on self report and some conditions such as tubal blockage or male-factor infertility can be estimated only for couples who have undergone infertility evaluations. Couples without knowledge of where or how to access infertility services may be unable to report causal factors, and couples who lack the emotional or economic resources to complete a comprehensive infertility evaluation may have incomplete or inaccurate information. According to 2002 NSFG findings, approximately 36% of women with current fertility problems (impaired fecundity or infertility) reported ever having used medical services for infertility (Chandra et al., 2005). Among users, 33% of women with current fertility problems had recognized

[4] Although the 2002 NSFG data reports Hispanic origin as well as race (black, white), the other clinical studies reviewed in this chapter do not distinguish Hispanic origin apart from race/ethnicity, and in these studies racial/ethnic groups are commonly classified as "white," "black," "Asian" or "Hispanic."

Table 6.1 Summary of major infertility diagnoses, associated factors, diagnostic tests and potential treatment strategies

Infertility factor	Description	Associated factors[a]	Diagnostic test	Examples of treatments
Diminished ovarian reserve	Reduced number of oocytes capable of fertilization	Aging with delayed childbearing, premature ovarian failure	Assessment of ovulation by hormone testing and ultrasound, day 3 FSH level	Assisted reproductive technology (using egg donor if needed)
Ovulatory dysfunction	Impairment of monthly ovulation	Polycystic ovarian syndrome, obesity, elevated androgen production, elevated insulin levels, thyroid imbalance, pituitary tumor	Assessment of ovulation, serum progesterone, serum gonadotropins (FSH, LH), serum insulin and androgens. Exclude other treatable conditions with thyroid and prolactin testing	Ovulation induction with clomiphene citrate or gonadotropin treatment, Metformin to enhance insulin sensitivity, laparoscopic ovarian drilling to reduce androgens and improve ovulatory function
Endometriosis	Growth of endometrial tissue outside of the uterus with attachment to other pelvic structures	Endometriomas (ovarian "chocolate" cysts) or adhesions (scar tissue) resulting from endometriosis	Laparoscopy to visualize pelvic organs	Surgical ablation or resection of endometriosis plus laparoscopic adhesiolysis (removal of scar tissue), ART
Uterine abnormalities	Congenital or acquired abnormalities that impair implantation or increase the risk of miscarriage	Fibroids, uterine polyps, uterine change due to DES prenatal exposure	Pelvic exam, ultrasound and/or hysteroscopy to assess for uterine abnormalities	Surgical removal of fibroids or polyps, Gestational surrogate for severe abnormalities
Fallopian tube obstruction	Blockage of tubes or damage to tubal lining that impair passage of gametes through fallopian tube	Prior sexually transmitted infections, previous ectopic pregnancy	Assessment for tubal damage with HSG (hysterosalpingogram) or laparoscopy, screening for sexually transmitted infections	Antibiotic treatment of active infection, Microsurgical repair of fallopian tubes (high risk of subsequent ectopic pregnancy), in vitro fertilization
Male factor	Decrease in quantity or quality of sperm; includes absence of sperm (azoospermia), decreased quantity of sperm (oligospermia), or impaired function (poor mobility or inability to fertilize)	Aging, lifestyle, environmental exposures	Semen analysis, male hormone testing	Surgical correction of epididymal blockage, surgical sperm recovery, artificial insemination, in vitro fertilization, intracytoplasmic sperm injection (ICSI), donor insemination
Unexplained infertility	No obvious male or female factor		Normal test results for ovulation, sperm production, fallopian tube patency	Ovulation induction combined with intrauterine insemination, In vitro fertilization

[a] Environmental exposures and lifestyle may be risk factors for many of these conditions

ovulatory difficulties. Other reported conditions included endometriosis (12%), occluded fallopian tubes (5%), and male factor problems (7.6%) (Chandra et al., 2005). Racial and ethnic breakdowns at the population level are unavailable.

Another source of data on medical conditions is provided by infertility clinics. In a chart review of 756 women presenting to two infertility centers in Ohio, black women were more likely than whites to have tubal factor infertility (41 vs. 13.8%), and less likely to have ovarian factor and male factor infertility (Green, Robins, Scheiber, Awadally, & Thomas, 2001). A Massachusetts study also reported a higher rate of tubal factor infertility in blacks (Jain, 2006). Studies of women seeking in vitro fertilization in the U.S. have also reported a higher prevalence of tubal factor infertility in blacks – ranging from 61 to 73% – compared to 23 to 40% among whites (Feinberg, Larsen, Catherino, Zhang, & Armstrong, 2006; Nichols, Higdon, Crane, & Boone, 2001; Sharara & McClamrock, 2000). These data have the advantage of being medically verified. However they are subject to selection bias; infertility care recipients are more likely to have medical insurance, advanced education and higher income than the underlying population of infertile adults (Jain, 2006). Nevertheless, the higher rates of tubal factor infertility in blacks are consistent with population data on risk factors for tubal disease, such as racial disparities in the prevalence of sexually transmitted infections and pelvic inflammatory disease (PID) (Centers for Disease Control, 2004; Chandra et al., 2005).

Utilization of Services

A range of treatment options, listed in Table 6.1, is available to individuals or couples who seek infertility care. Since services are usually expensive and often not covered by insurance, access is restricted to those with adequate economic resources. In 2002, approximately 12% of reproductive-age women (13.8% of whites, 8.2% of Hispanics and 8.4% of blacks) reported ever having received any infertility service (Chandra et al., 2005). Whites and women of higher income reported higher rates of utilization. As shown in Fig. 6.2, Hispanics and blacks reported lower utilization than whites at both higher and lower income levels.

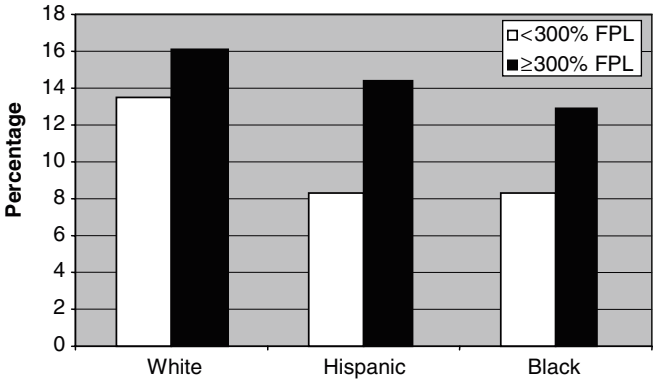


Fig. 6.2 Percentage of women aged 20–44 years who ever received any fertility service (infertility service includes service to achieve pregnancy and to prevent miscarriage), by race and ethnicity and percent of poverty level: United States, 2002 [Source: 2002 National Survey of Family Growth (Chandra et al., 2005) (Data generously adapted for this table by Anjani Chandra)]

Table 6.2 Percentage of women with current fertility problems[a] who ever received any infertility service, and percentage who received specific infertility services, by race/ethnicity: United States, 2002

	Any infertility service	Advice	Testing	Ovulation drugs	Surgery or treatment for blocked tubes	Artificial insemin.	ART	Medical help to prevent miscarriage (%)
TOTAL	35.7	22.9	19.5	14.8	3.3	5.5	1.0	12.2
White, non-hispanic	40.2	28.0	25.1	18.2	4.3	6.8	1.3	12.6
Black	26.1	10.9	5.0	5.6	1.4	1.9	0.8	14.5
Hispanic	28.5	13.2	11.7	11.3	1.9	2.9	0.6	7.9
Non-hispanic other	25.7	16.6	10.1	6.7	0.0	5.2	0.0	12.5

Source: National Survey of Family Growth, 2002 (Adapted from Stephen and Chandra, 2006 poster, Table 3)
[a]Current fertility problems defined as 12-month infertility plus impaired fecundity

Among women with "current fertility problems" (a combination of 12-month infertility and impaired fecundity), rates of service utilization were much higher; nonetheless they reflect the same pattern of racial-ethnic differences (Chandra et al., 2005). In 2002, white women were more likely to have reported receiving medical services (40.2%) than Hispanic (28.5%) or black (26.1%) women (Table 6.2). Specifically, whites were more likely than Hispanics or blacks to have received fertility advice, infertility testing, ovulation drugs, surgery or treatment for blocked tubes, artificial insemination and/or assisted reproductive technologies (ART). The only exception to this pattern was the higher utilization of services to prevent miscarriage by black women, which perhaps is an indication of better access to prenatal care than infertility services. ART was only utilized by 1% of women with current fertility problems, and was more likely to be used by whites than women of color (1.3% of whites, 0.8% of blacks and 0.6% of Hispanics).

The percentage of women who utilized each infertility service increased with advancing age, being married, higher levels of education, and higher income (Chandra et al., 2005). Women with private health insurance (42.9%) were twice as likely to obtain services as women with Medicaid (21.2%) or women without insurance (21.9%) (Chandra et al., 2005). Racial/ethnic differentials in health insurance coverage and in type of coverage (private or public) likely contribute to racial-ethnic disparities in service utilization. Financial barriers may also help explain why the proportion of women who seek infertility services has remained stable despite increases in availability of services (Stephen & Chandra, 2006). With regards to ART, most health plans exclude coverage, mainly on grounds that this service is not medically necessary (Jain, 2006). But the greater range of ART techniques and the increased number of available infertility specialists have resulted in a rise in the number of visits and in the amount of money spent on infertility by a select group of patients (Mosher & Bachrach, 1996).

Several states have passed legislation requiring health insurers "to offer" plans with infertility coverage, or "to cover" infertility services within specified health plans. Bitler and Schmidt (2006) examined whether infertility service utilization was higher in states with legal mandates to provide some form of infertility coverage. While it could be assumed that legal mandates to insure infertility services would allow women of lower socioeconomic status to obtain infertility services, their study found no evidence that mandates increased the probability that non-white women obtained treatment. In fact, the impact of state mandates in expanding access to infertility services was confined to older, more educated, white women. We speculate that this group of women may have been more likely to have medical insurance, better able to access existing services, and/or more motivated to

obtain infertility care. Although infertility services can be covered by Medicaid under the rubric of family planning (Gold, Richards, Ranji, & Salganicoff, 2007), poor women on Medicaid have little access to infertility treatments as few states' Medicaid programs actually cover infertility diagnosis and even fewer cover treatment (Kaiser Commission on Medicaid and the Uninsured and the Women's Health Policy Program, 2001).

A study of families within the military health care system suggests that utilization of infertility services by certain minorities is higher when traditional barriers to specialty care such as low socioeconomic status and lack of health insurance are reduced (Feinberg et al., 2006). Feinberg et al. (2006) compared utilization of ART within the Department of Defense to utilization of ART within the U.S. civilian population. They showed that blacks in the military health care system had a fourfold increase in relative utilization of ART services relative to black civilians. By contrast, Hispanic families in the military appeared to "underutilize" ART services based on demographics (Feinberg, Larsen, Wah, Alvero, & Armstrong, 2007). Although 9% of the Department of Defense self reported as Hispanic, only 4% of ART users were Hispanic (Feinberg et al., 2007). Feinberg et al. (2007) concluded that factors other than access to care influenced the utilization of infertility services by Hispanic families.

The effects of social and cultural factors on infertility treatment-seeking behaviors by women of color have been explored only in studies using small convenience samples. Several barriers to care have been identified including: distrust of the medical community (Jenkins, 2005; White, McQuillan, & Greil, 2006; Becker, Castrillo, Jackson, & Nachtigall, 2006); religious beliefs that discourage use of high-tech conception interventions that run "counter to God's wishes" (Jenkins, 2005; Parham & Hicks, 2005); an extended family ethos that encourages other parenting options including adoption (Jenkins, 2005; White et al., 2006; Parham & Hicks, 2005); fear of being labeled "infertile" and losing hope and motivation to try to conceive (Becker et al., 2006); and, language difficulties (Inhorn & Fakih, 2006; Jenkins, 2005).

The extent to which cognitive barriers such as having awareness of an infertility problem, knowledge about reproduction and access to information about infertility services influence help-seeking behaviors has not been adequately studied (White et al., 2006). For instance, women who previously conceived a child may not recognize secondary infertility in a timely manner, and women with little familiarity with the medical system or knowledge of their insurance benefits may be unaware of the services that exist to treat fertility problems.

Motivational factors tied to the desire for children also influence utilization of infertility services. In one study, motivation to seek services was lower among couples who had higher parity and in couples in which either partner had a child from a prior marriage (Kalmuss, 1987). In contrast, stronger motivation to conceive has been found among women who delay childbearing until an advanced age (Mosher & Bachrach, 1996).

Provider attitude, knowledge and behavior can also affect patient utilization of care. Some primary care providers may not provide early recognition of and treatment of diseases that threaten fertility (Office of Technology Assessment, 1988). Others, aware of a couple's financial constraints, may choose not to refer them to infertility specialists (Heitman, 1995). Infertility service providers may have difficulty adapting testing and treatment protocols to meet the diverse social and cultural needs of their clients (Blenner, 1991). As infertility specialists are permitted to use their judgment in determining whom they will treat, some have refused treatment to unmarried women (Jenkins, 2005). Black and Hispanic women are less likely to be married than white women, and this could result in fewer women of color receiving services.

Combining these factors may help to explain why women of color, especially black and Hispanic women access care less frequently, and may experience longer duration of infertility before seeking care (Sharara & McClamrock, 2000; Bendikson, Cramer, Vitonis, & Hornstein, 2005). It seems plausible that delaying access to infertility services diminishes the likelihood of achieving a successful pregnancy.

In summary, the evidence suggests that black and Hispanic women have unequal access to infertility services, despite their higher rates of infertility. Timely access to services can potentially impact the chance of becoming pregnant and of delivering a healthy child.

In Vitro Fertilization

With this background in mind, we now turn to examination of one infertility treatment – in vitro fertilization (IVF). While IVF represents a first-line treatment for some causes of infertility (e.g., tubal factor infertility and severe male infertility), for most infertile couples it provides a last option for treatment following a succession of increasingly invasive and expensive procedures. The mean cost of one cycle of IVF is $12,400 (Grayson, 2003). Given an average success rate of less than 37% (Wright, Chang, Jeng, & Macaluso, 2006), many couples attempt several cycles of IVF before becoming pregnant or abandoning treatment.

IVF Clinical Protocol

A typical IVF treatment cycle includes ovarian stimulation, egg retrieval, fertilization, and embryo transfer. Initially, medication is administered to a woman to stimulate the production of multiple eggs (ovarian stimulation). The ovarian response is monitored with ultrasound and hormone testing. If ovarian follicles and hormone levels are appropriate, ultrasound guided vaginal retrieval is performed (egg retrieval). Thereafter, sperm is collected and combined with extracted eggs (oocytes) in the laboratory (in vitro) to effect fertilization.[5] If fertilization is successful, the resulting embryos are incubated for 2 to 5 days. Next, one or more embryos are selected for transfer into the uterus (embryo transfer).[6] In general, implantation is more successful when a higher number of embryos are used. However, transfers of more than one embryo are associated with an increased likelihood of a multiple gestation (Wright et al., 2006). When implantation is successful, an early pregnancy can be detected by a blood test (chemical pregnancy). Once a gestational sac develops and can be visualized by ultrasound, the pregnancy is defined as a clinical pregnancy. The most successful clinical pregnancies result in the delivery of one or more live-born infant.

The numbers of live births resulting from IVF represent a fraction of cycles that are initiated. Cycles may fail or be cancelled at different stages. For example, some IVF cycles may not lead to egg retrieval and others may not result in egg fertilization. A portion of embryo transfers result in miscarriage or ectopic pregnancies. In 2005, 97,442 ART cycles were initiated with fresh non-donor eggs, resulting in 85,713 egg retrievals, 78,797, embryo transfers, 33,101, clinical pregnancies, and 27,047 live-birth deliveries [Centers for Disease Control and Prevention (CDC), American Society for Reproductive Medicine (ASRM), & Society for Assisted Reproductive Technology (SART), 2007].[7]

[5] Approximately half of U.S. ART procedures utilize intracytoplasmic sperm injection (ICSI), a technique in which a single sperm is injected into an egg (Wright, Schieve, Reynolds, & Jeng, 2005).

[6] Two other modifications of this procedure are (1) gamete intrafallopian transfer (GIFT) in which gametes are transferred to the fallopian tube immediately following egg retrieval and (2) zygote intrafallopian transfer (ZIFT) in which a fertilized egg (Zygote) is transferred to the fallopian tube by laparoscopy the day following egg retrieval (Steinberg, Holtz, Sullivan and Villar, 1998).

[7] For up to date statistics on ART cycles, refer to the following website: http://www.cdc.gov/art/.

Success Rates from IVF

The success of a cycle can be reported in different ways, which can lead to some confusion between providers and consumers. A clinic may report success based on the likelihood of achieving clinical pregnancy following successful embryo transfer, while an infertile couple may be more interested in the probability of a live birth per initiated cycle. Success rates based on initiated cycles will be lower than rates based on embryo transfer, because a number of cycles never reach the stage of embryo transfer. To understand success rates from a clinical standpoint, it is important to be clear on the definition of the numerator (the clinical outcome) and the denominator. The Society for Assisted Reproductive Technology (SART) has been working closely with the Centers for Disease Control and Prevention to collect and report national data in order to comply with the Fertility Clinic Success Rate and Certification Act, a 1992 federal mandate to standardize the reporting of ART success rates.[8] It is important to note that these data largely reflect IVF cycles which account for 99% of ART procedures.[9] National 2005 IVF success rates (using fresh non-donor eggs) were reported in the following ways: a 34.0% clinical pregnancy rate per cycle, a 27.8% live birth rate per cycle, a 31.6% live birth rate per egg retrieval, and a 34.3% live birth rate per embryo transfer (CDC et al., 2007). From a public health perspective, the primary outcome of interest as an indicator of success of IVF is live birth deliveries.

Effectiveness of IVF

IVF results vary according to patient factors as well as treatment factors. Age is the most influential patient variable; chances for a successful response to IVF diminish as a woman ages. In 2005, 37% of IVF cycles initiated on women under the age of 35 resulted in live births. In contrast, 11% of women aged 41–42, and only 4% of women older than 42 achieved live birth through IVF (CDC et al., 2007). Other factors influencing IVF response include parity and infertility diagnosis. A prior successful pregnancy is associated with a higher likelihood of IVF success, while certain infertility diagnoses including diminished ovarian reserve, uterine abnormalities, and combined male and female factor infertility are associated with lower success rates (CDC et al., 2007). Additionally, first cycles of IVF are more likely to be successful than subsequent cycles.

Treatment factors that influence IVF outcomes include the number of days the embryo is cultured, the number of embryos transferred, whether an intracytoplasmic sperm injection (ICSI) is utilized, whether the embryo is fresh or cryo-preserved, and whether a donor egg is used.

The number of IVF cycles in the United States doubled from 1996 to 2004. Live birth rates have improved as IVF techniques have been refined. From 1996 to 2005, the live birth rate per embryo transfer resulting from fresh non-donor cycles increased from 28 to 34% (CDC et al., 2007).

Response to IVF by Race/Ethnicity

In this section, we present a detailed summary of U.S. and British studies that examine responses to IVF by race and/or ethnicity. Comparing IVF success rates between studies would be misleading because the infertility centers that were studied varied in their patient composition, clinical treatment

[8] The Fertility Clinic Success Rate and Certification Act includes a voluntary certification program for ART laboratories to ensure quality of care, and provides consumers with standardized information on ART success rates from individual ART clinics.

[9] For ease of reading, this paper will report all national ART results as IVF results.

protocols, and years of study. Rather, these studies can indicate whether pregnancy and live birth rates resulting from fresh, non-donor IVF vary by race and ethnicity within a particular study population. A total of 11 studies (9 peer-reviewed and 2 poster presentations) were identified using Medline, Cochrane Library, CINAHL and Popline search engines in 2006 and again in 2008. The identified articles span the years 1995 through 2007. All of the identified studies are included in this review and are presented in chronological order (Table 6.3).

The first report on IVF response by race/ethnicity compared the outcomes of IVF cycles conducted in a British university IVF center between 1987 and 1993 in 44 Indian women and 88 white women, retrospectively matched for age, body mass index (BMI) and year of treatment (Mahmud, Lopez Bernal, Yudkin, Ledger, & Barlow, 1995). Compared to whites, Indian women had a longer duration of infertility prior to treatment and IVF cycles were marked by a higher discontinuation of treatment (cancellation rate of 22.7 vs. 9.1%). Indians had a lower clinical pregnancy rate per cycle (18.2 vs. 27.3%) and a higher miscarriage rate, resulting in a lower live birth rate (9.1 vs. 22.7%) per cycle. There were, however, no differences in the clinical protocol used, the number of retrieved eggs, the rate of fertilization, or the number of embryos transferred.

In a subsequent British study, Lashen, Afnan, and Sharif (1999) compared the ovarian response to controlled ovulation induction in 108 first-generation Indian or Pakistani women and 216 White women undergoing IVF from 1994 to 1997. The groups were matched for age, early follicular phase follicle stimulating hormone (FSH) levels, cause of infertility, dose of gonadotropin and year of treatment. The authors found higher implantation and clinical pregnancy rates in the whites, but this did not achieve statistical significance. Data on miscarriage and live birth rates were not provided. The groups did not differ in the duration of stimulation, number of eggs or embryos produced, fertilization rate, or treatment cancellation rate. The authors concluded that the response to controlled ovarian stimulation was comparable between matched first-generation Indian or Pakistani and British whites when IVF clinical protocols were identical and women were matched for FSH levels.

The first report of IVF response by race/ethnicity in the United States was from a Maryland inner city, university-based IVF program (Sharara & McClamrock, 2000). In this retrospective comparison, 95 whites (undergoing 121 IVF cycles) and 37 blacks (undergoing 47 IVF cycles) treated from 1997 and 1999 were compared. Women with hydrosalpinges (blocked fallopian tubes filled with fluid), elevated early FSH, intrauterine polyps or large fibroids, or age greater than 40 were excluded. Blacks were characterized by higher rates of tubal factor infertility, higher BMI, and longer duration of infertility than whites. Whites had a higher incidence of endometriosis and male factor infertility. Compared to whites, blacks had lower implantation rates (9.8 vs. 23.4%; $p < 0.001$) and clinical pregnancy rates per initiated cycle (19.2 vs. 42.2%; $p = 0.009$). There appeared to be a higher rate of early miscarriage among blacks than among whites; however, the small sample size in this report limited the power of this analysis. When the sample was restricted to first IVF cycles, the racial/ethnic differences in clinical pregnancy rate per initiated cycle and per embryo transfer remained highly significant. While many features of the IVF cycles were similar, blacks required a more aggressive clinical protocol (microdose flare GnRH cycles, an approach used in women who do not respond well to the standard IVF drug protocols) than whites (70.2 vs. 43%; $p = 0.01$), suggesting more effort was needed to achieve a clinical response.

Nichols et al. (2001) conducted a retrospective study of 316 women undergoing IVF procedures in a South Carolina hospital based practice between 1996 and 2000. In this study, 24 blacks (25 cycles) and 273 white women (333 cycles) were included – all under 41 years old and without current uterine intracavitary defects or large fibroids. Black women had a higher average BMI than white women, and were more likely to be parous and have tubal-factor infertility. Whites were more likely than blacks to have endometriosis and ovarian dysfunction as the primary diagnosis. In contrast to the Sharara and McClamrock study (2000), the implantation rate was higher in blacks than whites, contributing to a higher pregnancy rate in blacks. After adjusting for the differences in

Table 6.3 Survey of the evidence regarding response to in-vitro fertilization (IVF) by race/ethnicity

Author (date)/ study design	Publication	Description of IVF clinic location and years studied	Population studied and sample size	Addressed disparities in which populations	Key findings related to intervention effectiveness (OR w/CI or *p* values)	Adequacy of data – caveats/strengths
Mahmud et al. (1995)/ retrospective comparison	Published article	IVF clinic in England 1987–1993	44 Indian, 88 white, matched for age, BMI and year of treatment, 1st cycle only	British Indian and British white	*Cancellation rate* higher in Indian women *Pregnancy rate* lower in Indian Women • Longer duration of infertility in Indian women (41% Indian with >8 years infertility vs. 22% of whites) *No difference in*: • # ampules gonadotropins • # oocytes retrieved • Fertilization rate • # embryos transferred	*Selection*: First cycles *Exclusions*: Not specified *Matching*: Age, BMI, and year of treatment *IVF Protocol*: Consistent *Analyses*: No adjustment for group differences in duration of fertility, primary vs. secondary infertility, infertility diagnosis *Outcome*: Lower clinical pregnancy and live birth rate in Indian women due to higher rate of cancelled cycles and higher miscarriage rate *Caveats*: Small sample size
Lashen et al. (1999)/ retrospective comparison	Published article	IVF clinic in England 1994–1997	108 Indian or Pakistani immigrants, 216 white, matched for age, early follicular phase FSH, cause of infertility, gonadotropin dose and year of treatment	British Indian or Pakistani and British white	*Cancellation rate* similar in both groups *Implantation rate* did not show statistically significant difference *Pregnancy rate* did not show statistically significant difference *Indian/Pakistani women* • Longer duration of infertility • Lower implantation rate per cycle (13% vs. 17%) • Lower pregnancy rate per cycle (16% vs. 23%) *No difference in*: • Duration of stimulation • # eggs or embryos produced • Fertilization rate • Treatment cancellation rate	*Selection*: First cycles *Exclusions*: Not specified *Matching*: Age, day 3 FSH level, infertility diagnosis (male factor or non-male factor), dose of gonadotropin, and year of treatment *IVF Protocol*: Consistent *Analysis*: No adjustment for group differences in duration of infertility, and differences in female infertility diagnoses *Outcome*: Similar ovarian response rate; lower, but not statistically significant, implantation and clinical pregnancy rates in Indians/Pakistani women *Caveats*: Small sample size and low pregnancy rates in both groups – could not rule out Type II error

(continued)

Table 6.3 (continued)

Author (date)/ study design	Publication	Description of IVF clinic location and years studied	Population studied and sample size	Addressed disparities in which populations	Key findings related to intervention effectiveness (OR w/CI or *p* values)	Adequacy of data – caveats/strengths
Sharara & McClamrock (2000)/ retrospective comparison	Published article	IVF inner city university based program in Maryland, April 1997–July 1999	95 white (121 IVF cycles), 37 black (47 IVF cycles)	Black and white	*Clinical pregnancy rate* lower in blacks than whites (19.2 vs. 41.2%) *Implantation rate* lower in blacks (9.8 vs. 23.4%) *Blacks* • Higher % tubal factor • Higher BMI • Longer duration of infertility • Required more aggressive stimulation protocol (70 vs. 43%) *Whites* • Higher % endometriosis and male factor *No statistical differences*: • # ampules gonadotropin • Duration of stimulation • Endometrial thickness • Cancellation rate • # retrieved oocytes • # transferred pre-embryos	*Selection*: All cycles, subgroup analysis of first cycles *Exclusions*: age>40, day 3 FSH>10, hydrosalpinges, intracavitary uterine abnormalities, mixed race *Matching*: Not done *IVF Protocol*: Variable *Analysis*: No adjustment for group differences in BMI, infertility diagnosis, duration of infertility, and protocol variations; subgroup analysis of first IVF cycle performed *Outcome*: Blacks maintained lower pregnancy rate overall and in first cycle analysis *Caveats*: Small sample size
Nichols et al. (2001)/ retrospective comparison	Published article	South Carolina Hospital based 1996–2000	273 white (333 cycles), 24 black (25 cycles)	Black and White	*Pregnancy rate* higher in blacks than whites (71 vs. 48%); OR 3.3 (1.3–8.6) remained higher after adjusting for parity, BMI, and tubal factors *Implantation rate* higher in blacks than whites (35 vs. 23%) *Blacks* • Higher % tubal factor • More likely to have prior child • Higher BMI *Whites* • Higher % prior cycles	*Selection*: All cycles *Exclusion Criteria*: Age >40, mixed ethnicity, FSH>10, myomas >4 cm *Matching*: Not done *IVF Protocol*: Variable *Analyses*: Multiple logistic regression controlling for parity, BMI, presence of tubal factor infertility *Outcome*: Higher pregnancy rate for blacks in crude and adjusted *Caveats*: Higher proportion of white patients with poor prognosis and whites were more likely than blacks to be undergoing multiple IVF attempts (28 vs. 4%), small sample size

Bendikson et al. (2005)/ retrospective comparison	Published article	Boston Collaborative IVF Study, 3 IVF Centers 1994–1998	1,039 white, 43 black, 18 Hispanic, 35 Asian	White, black, Hispanic and Asian	*Chemical pregnancy rate* and *live birth rate* – no statistical differences by race/ethnicity *Blacks* • Higher % tubal factor infertility than whites • Higher BMI than whites or Asians • Higher peak estradiol *Hispanics* • Longer duration of infertility than whites and blacks *Asians* • Higher peak estradiol *No differences*: • Day 3 FSH • # oocytes retrieved • # embryos transferred • Miscarriage rate • Live births	*Selection*: First cycle only *Exclusions*: Mixed race clients *Matching*: Not done *IVF Protocol*: Variable, but # gonadotropins per cycle and microflare cycles per ethnic group were not significantly different *Analysis*: No adjustments for group differences in BMI, gravity, duration of infertility, infertility diagnoses *Outcome*: No differences in pregnancy and live birth rate *Caveats*: Potential selection bias – 65% of couples who were approached for study agreed to participate; limited # of non-white participants
James et al. (2002)/ retrospective comparision	Poster session, *Fertility and Sterility* [abstract]	University of Alabama IVF program 1995–2001, 1st cycle only	87 white, 41 black, matched for age, parity, infertility diagnosis and year of service	Black and white	*Clinical pregnancy* and *live birth rate* – no statistical differences by race *Blacks* • Higher BMI • Higher % fibroids *No differences*: • Number of eggs retrieved • Number of embryos transferred	*Selection*: First cycle only *Exclusions*: Not specified *Matching*: Age, date of service, parity and primary infertility diagnosis *IVF Protocol*: Not specified *Analysis*: No adjustment for group differences in BMI and fibroids *Outcome*: No difference in clinical pregnancy and live birth rate after controlling for age, parity, and infertility diagnosis *Caveats*: Small sample in one clinic

(continued)

Table 6.3 (continued)

Author (date)/ study design	Publication	Description of IVF clinic location and years studied	Population studied and sample size	Addressed disparities in which populations	Key findings related to intervention effectiveness (OR w/CI or *p* values)	Adequacy of data – caveats/strengths
Feinberg et al. (2006)/ retrospective comparision	Published article	Department of Defense population, Walter Reed Army Medical Center 1999–2003	2,650 cycles reviewed for 1,387 patients: 974 white, 252 black, 56 Hispanic, 94 Asian/PI, 10 native American, only first cycle, fresh, non-donor	Black, Hispanic, Asian/PI and white	*Clinical pregnancy rate* – no statistical differences by race or ethnicity *Live birth rate* – lower in blacks than whites (29 vs. 35.8%, OR 0.83, 95% CI 0.67–1.02) *Implantation Rate* – no difference *Blacks* • Higher % tubal factor (60.5 vs. 31.2%); RR 1.91(1.66–2.18) • Higher % leiomyoma (30.8 vs. 10.7%); RR 2.85 (2.06–3.95) *Whites* • More likely to have male factor, anovulation, endometriosis and unexplained infertility	*Selection*: First cycle only *Exclusions*: Age>41, day 3 FSH >11, "other" races *Matching*: Not done *IVF Protocol*: Variable *Analysis*: Stratified by presence of fibroids *Outcomes*: No significant difference in clinical pregnancy rate between blacks and whites; however, blacks had lower live birth rate; when adjusting for fibroids, live birth rate was similar in both groups *Caveats*: Higher miscarriage rate in blacks (OR 1.57, 95% CI 1.05–2.36); lower implantation rate and higher miscarriage rate associated with presence of fibroids. Blacks utilized IVF services at rate proportional to the population; Hispanics under-utilized IVF services by more than 50%
Feinberg et al. (2007)/ retrospective. comparison, using same data set as previous study	Published article	Department of Defense population, Walter Reed Army Medical Center 1999–2003	2,650 cycles reviewed, 1,457 patients met inclusion criteria	Hispanic, white	*Clinical pregnancy rate, live birth rate, implantation rate* and *spontaneous abortion rate*– no significant differences between Hispanic and white patients *Clinical Diagnoses* – similar for Hispanic and white patients	*Selection*: First cycle *Exclusions*: Women > 41, FSH > 11, "other" races *Matching*: Not Done *Analysis*: No adjustments *Outcome*: Similar rates of clinical pregnancy and live birth rate in Hispanics and whites *Caveats*: Lower than expected use of IVF services by Hispanics based on military demographics; Hispanics comprise 9% of military population, but comprise only 4% of IVF utilizers

Grainger et al. (2004)/ retrospective comparison	Poster session *Fertility and Sterility* [abstract]	SART national registry 1999–2000, included data from clinics reporting on race/ ethnicity in more than 90% of clients and with more than 50 ART cycles per year	80,196 ART cycles: 68,607 white, 4,338 Hispanic, 3,585 Asian, 3,666 black, first and subsequent cycles included	White, Hispanic, Asian, black	*Live birth rates* lower inblacks (18.7%) and Asians (20.7%) than whites (26.3%) and Hispanics (26.7%) *Asians* • Mean age older than black, white, Hispanic • Higher # mean embryos transferred than white *Hispanics* • Lower mean ratio of day 3 serum FSH to lab upper normal	*Selection*: First and subsequent cycles *Exclusions*: Cycles which involved donor eggs and gestational carriers *Matching*: Not done *IVF Protocol*: Variable *Analysis*: Stratified by age *Groups differed by*: Age distribution, day 3 FSH and clinic *Outcome*: When stratified by age group, blacks and Asians had lower live birth rate than whites for all age groups up to 40; for 41–42 year olds, blacks had lower live birth rate than all other groups *Caveats/Strengths*: Large sample size; only clinics reporting race in more than 90% of clients and conducting ≥50 cycles/year were included; racial differences in clinical pregnancy rate seen only in fresh cycles not in cycles using cryopreserved embryos
Seifer et al. (2007)/ retrospective comparison	Published article	SART national U.S. registry 1999–2000, included clinics reporting race in >95% of cycles and conducting >50 cycles per year	80,309 cycles, 68,607 white, 3,666 black, 8,036 other race/ethnicity	White and black	*Clinical pregnancy rate* lower in blacks than whites (27.7% vs. 33.6%, $p < 0.001$) *Live birth rates* lower in blacks than whites (unadj. rates for first IVF cycle 20.7% vs. 28.4%; $p < 0.001$). This disparity remained sig. after adjusting for multiple patient and clinic level factors Blacks had longer duration of infertility, higher likelihood of tubal and uterine factor infertility. Whites were more likely to have male factor infertility, endometriosis or ovulatory disorders. Blacks were more likely to receive treatment at clinics with l ower median number of ART cycles per year and lower success rates overall	*Selection*: First and subsequent cycles *Exclusions*: Cycles with donor eggs or gestational carriers *Matching*: Not done *IVF Protocol*: variable *Analysis*: Multivariate analysis controlled for age, parity, day 3 FSH ratio, etiology of infertility, clinic level factors (i.e., clinic volume and overall clinic pregnancy rate) *Strengths*: Large sample size, multivariate analysis with individual and clinic level factors, inclusion of significance levels.

(continued)

Table 6.3 (continued)

Author (date)/ study design	Publication	Description of IVF clinic location and years studied	Population studied and sample size	Addressed disparities in which populations	Key findings related to intervention effectiveness (OR w/CI or *p* values)	Adequacy of data – caveats/strengths
Purcell et al. (2007)/ retrospective comparison for two data sets, including data set used in Grainger study	Published article	(1) SART national registry, 1999–2000 Data from clinics reporting on race in >90% of clients and conducting >50 ART cycles per year (2) University of California, San Francisco (UCSF) Jan 2001– Dec 2003	SART natural Database-187 clinics with >50 cycles National Registry: 25,843 white cycles 1,429 Asian cycles First cycles only UCSF Database: 567 cycles 197 Asian 370 white	Asian and white	*Clinical pregnancy rate* lower in Asians SART – OR 0.71 (0.64–0.80). UCSF – OR=0.69 (0.49–0.99) *Live birth rate* lower in Asians SART – OR 0.69 (0.61–0.77) UCSF – OR 0.67 (0.46–0.98) SART – differences in age, nulliparity, nulligravidity, diagnosis of diminished ovarian reserve; similar d3 FSH levels UCSF – similar mean age and d3 FSH level; Asians less likely to have diagnosis of diminished ovarian reserve, more freq ovarian dysfunction and unexplained infertility, higher estradiol levels	*Selection*: First cycle *Exclusions*: Cycles cancelled due to lack of follicular response, lack of oocyte retrieval or failed fertilization *Matching*: Not done *IVF Protocol*: Multiple *Analysis of SART data*: Univariate and multivariable logistic regression, controlling for age, infertility diagnosis, parity, day 3 FSH, use of ICSI and number of embryos transferred *Analysis of UCSF data*: Multivariate analysis controlled for age, primary or secondary infertility, prior spontaneous or therapeutic abortions, number of prior IVF attempts, IVF stimulation protocol, specific type of female infertility, days of gonadotropin stimulation, embryo fragmentation score, physician performing embryo transfer and difficulty of transfer *Outcome in SART and UCSF data*: Asians had lower pregnancy rates than whites, difference remained in multivariate analysis *Strengths*: Includes large, national dataset

parity, presence of tubal disease, and BMI, the odds of pregnancy remained elevated for blacks (OR 3.3; 95% CI 1.3–8.6). However, it should be noted that other factors were not controlled that could have affected the outcome. Almost all (96%) of blacks were attempting their first IVF cycle compared to 72% of whites, and first cycles tend to be more successful than subsequent ones. South Carolina does not have mandated infertility coverage, and the socioeconomic profile of these patients was likely more affluent than that of patients in the Maryland study. In contrast, the Maryland study population (Sharara and McClamrock, 2000) was drawn from an IVF program in a state with mandated infertility coverage, and likely encompassed more diverse socioeconomic backgrounds. Blacks enrolled in the Maryland study also required an aggressive IVF protocol more frequently than those in the South Carolina study, suggesting that the Maryland study included a higher proportion of "poor responders" among their black patients.

Bendikson et al. (2005) compared ethnic differences in IVF responses in Massachusetts, another state with mandated infertility coverage. This retrospective cohort study of three IVF centers in Boston reported on first IVF cycles in 1,039 white (91.5%), 43 black (3.8%), 18 Hispanic (1.6%), and 35 Asian women (3.1%) who received treatment between 1994 and 1998. Mean age was similar in all groups, but mean BMI was significantly higher in blacks than whites and Asians. Blacks also had significantly higher parity than whites and Asians, and Hispanics had significantly longer duration of infertility than whites and blacks. Blacks were more likely than whites to have tubal factor infertility. The rates of chemical and ectopic pregnancies, miscarriage, and live birth rates did not differ by race/ethnicity. Furthermore, the number of cycles requiring aggressive protocol (microdose flare cycles), the total gonadotropin dose, the days of stimulation, the number of oocytes retrieved and the number of embryos transferred were comparable between groups. This study restricted the analysis to first IVF treatment cycles; however, it was limited by its retrospective nature, the small samples of minority patients, and the absence of control for patient characteristics such as parity, socioeconomic status (SES), and date of procedure.

James, Hammond, and Steinkampf (2002) at the University of Alabama, Birmingham, retrospectively matched 41 black women with 87 white women who underwent IVF between 1995 and 2001. Matching criteria were age, date of IVF procedure, parity, and primary infertility diagnosis. Blacks had a higher mean BMI than whites, and had higher prevalence of uterine fibroids. Despite these contrasts, there were no significant differences in the clinical pregnancy rate or the live birth rate. The number of eggs retrieved, number of embryos transferred, and the chemical pregnancy rate were comparable.

Feinberg et al. (2006, 2007) examined racial disparities in IVF outcomes within the Department of Defense (DoD), reviewing 2,650 cycles for 1,387 patients. The study included 974 whites, 252 blacks, 56 Hispanics, 94 Asian or Pacific Islanders and 10 Native Americans. Exclusion criteria for IVF included age over 41, a day 3 FSH greater than 11 and "other races". Intracavitary fibroids greater than 3 cm were removed prior to implementing IVF procedures.

In the 2006 study, blacks were more likely than whites to have tubal factor infertility as well as fibroids. Whites were more likely to be diagnosed with male factor infertility, ovulatory dysfunction, endometriosis and unexplained infertility. A clinically significant difference in the live birth rate in black women compared to white women (29.6 vs. 35.8%) was observed which did not reach statistical significance at the 0.05 level. The authors observed a lower implantation rate and a higher miscarriage rate in women with fibroids. When the population was stratified by the presence or absence of fibroids, the clinical pregnancy rate, the implantation rate and the live birth rate among black and white women with fibroids was comparable. Blacks had a higher relative risk for spontaneous abortion than whites [1.57 (95% CI 0. 1.05–2.36)]. This difference disappeared after controlling for fibroids.

Hispanics receiving IVF services in the military had the same distribution of clinical diagnoses as whites. There were no significant differences in the rates of clinical pregnancy, live birth, implantation and spontaneous abortion between these two groups (Feinberg et al., 2007).

Several analyses of racial disparities in IVF have utilized the SART 1999–2000 national register data set. This database collects information on outcomes of individual IVF cycles (rather than outcomes per patient) from SART member clinics throughout the United States.

Grainger et al. (2004) examined the outcomes of 80,196 fresh non-donor IVF cycles from SART member clinics. Only clinics reporting racial/ethnic data on at least 90% of their clients and conducting over 50 cycles of IVF per year were included. Cycles which involved donor eggs and gestational carriers were excluded. The racial/ethnic composition of patients was 85.5% white, 4.6% black, 4.5% Asian, and 5.4% Hispanic, reflecting a much lower proportion of women of color compared to the US population. The live birth rate per cycle was lower for blacks (18.7%) and Asians (20.7%) than for whites (26.3%) and Hispanics (26.7%). This pattern was consistent in each age group examined with the exception of women over 40. For 41–42 year olds, blacks had lower birth rates than all other groups while Asians and Hispanics had similar rates. This study also examined cycles with frozen eggs, and found no racial/ethnic differences in the success rates of these procedures**.**

In this study, Asians had an older mean age than other groups. Hispanic women had lower day 3 serum FSH than Asians, whites, and blacks. Asians also had higher mean number of embryos transferred than whites, but there were no significant differences in embryo transfers between the other racial/ethnic groups. This analysis did not control for infertility diagnosis, duration of infertility, BMI, SES and other patient and clinic characteristics that could have affected the live birth rates. Additionally, the clinic selection criteria suggest that only larger, more resourceful IVF centers were included.

In a subsequent in depth analysis of black-white differentials in the SART data set, Seifer, Frazier and Grainger (2008) found that black women undergoing their first IVF cycle using fresh non-donor embryos were more likely to have a longer duration of infertility and an infertility diagnosis related to tubal disease or a uterine anatomic problem. White women were more likely than blacks to be nulliparous and carry a diagnosis of male factor infertility, endometriosis, or diminished ovarian reserve. Cycle outcomes differed by race: clinical pregnancies were less likely for blacks than whites (27.7 vs. 33.6%, $p < 0.001$), were more likely to result in spontaneous abortion (20.4 vs. 13.8%, $p < 0.001$), and were less likely to achieve a live birth per initiated cycle (20.7 vs. 28.4%, $p < 0.001$). Multivariate analyses adjusting for potential individual confounders (maternal age, cycle day 3 FSH ratio, etiology of infertility, parity and number of embryos transferred) and clinic level factors (i.e., clinic volume, and clinic overall pregnancy rate) revealed that black women with no prior ART had a 24% (95% CI 1.12–1.36) increased risk of not achieving a live birth compared to whites. Blacks with a prior ART cycle had a 38% (95% CI 1.20–1.57) higher risk than whites. This study also analyzed clinic-specific factors and found that "more black women sought treatment at smaller centers that had somewhat lower overall success rates during 1999–2000." Education, insurance coverage, and socioeconomic status of IVF recipients were not included in this database.

Purcell et al. (2007) compared Asian and white IVF outcomes utilizing two data sets: (1) the large 1999–2000 SART dataset described above; and, (2) data from a California university IVF center. In the SART data set, cycles cancelled due to lack of follicular response, inadequate oocyte retrieval or failed fertilization were excluded. This SART sample included 25,843 white and 1,429 Asian first IVF cycles. The groups differed in age, parity, gravidity and diagnosis of diminished ovarian reserve, but had similar cycle day 3 FSH levels. Among Asians, the odds of achieving a clinical pregnancy rate was 29% lower than whites (OR 0.71, 95% CI 0.64–0.80), and the live birth rate was 31% lower in unadjusted analyses (OR 0.69, 95% CI 0.61–0.77). The odds of a live birth remained reduced for Asians after controlling for age, infertility diagnosis, parity and day 3 FSH levels, use of ICSI and number of embryos transferred (OR 0.76, 95% CI 0.66–0.88). When the SART data were stratified by age, the live birth rates were notably lower for Asians in all age groups, with the exception of 41–42 year olds.

The California university sample in the Purcell et al. (2007) study included 370 white cycles and 197 Asian first cycles from 2001 to 2003. The two groups had similar mean age and day 3 FSH levels, but Asians were less frequently given a diagnosis of ovarian dysfunction and unexplained infertility. In this sample, the odds of live birth for Asians was decreased compared to whites (OR

0.61; 95% CI 0.96–0.98). This decreased live birth rate persisted after adjusting for multiple covariates. The investigators noted that estradiol levels were higher in Asians, despite a similar number of follicles, but otherwise did not find significant differences in cycle characteristics. The authors hypothesized about possible reasons for the lower pregnancy rates in Asians, including underlying biologic or genetic differences as well as lifestyle and dietary differences that might have impaired successful implantation.

Several methodological limitations apply to the studies presented in Table 6.3. All use retrospective study designs based on data drawn from chart reviews. Given ethical considerations, no prospective or randomized trials have been conducted to assess the effect of IVF by race/ethnicity. Since subjects are drawn from IVF clinics, and civilian IVF users tend to be women of higher socioeconomic status, the civilian studies in particular are prone to selection bias. Blacks and Hispanics are underrepresented in all civilian clinics, and in some studies their low numbers preclude meaningful analysis. In addition, success rates from IVF use are affected by patient and treatment factors. The ideal study would control, either by design or statistical analysis, for a number of patient demographic characteristics (e.g., age, parity, SES, BMI), medical factors (e.g., infertility diagnosis, duration of infertility, day 3 FSH, prior IVF failure, primary or secondary infertility, prior miscarriage), and treatment factors (e.g., IVF protocol, number of embryos transferred, use of ICSI, IVF center, year of service). Most studies do not control for these features. In the last column of Table 6.3 we have identified important covariates and the extent to which these were controlled for through design or analysis in each study. Caution should be observed when comparing results between clinics that serve different populations, maintain different patient exclusion criteria, and utilize different clinical protocols. Some clinics exclude patients with specific conditions that are known to reduce IVF success rates, and thereby boost reported success rates (Schulman, 2007). Additionally, IVF technologies have improved over time; recent clinical success rates are likely to be higher than those of previous years.

The SART national registry has the largest sample size, and provides a rich source of data for future studies. Currently, IVF clinics provide success rates *per cycle* rather than *per patient*; success rates per cycle may underestimate the true per-patient success rate since the former over-counts patients who undergo more than one procedure in a given year. Also, this registry is subject to selection bias; in 2003, only 32% of the IVF clinics that complied with voluntary reporting provided racial and ethnic data.

Despite these limitations, these studies present a preliminary picture of racial/ethnic differences in response to IVF, and suggest that more work is needed in this area. The six American studies comparing IVF response in blacks and whites are inconsistent in their findings. Compared to whites, pregnancy rates for blacks were noted to be lower in three studies (Sharara & McClamrock, 2000; Grainger et al., 2004; Seifer et al., 2008), higher in one study (Nichols et al., 2001), and comparable in two studies (Bendikson et al., 2005: James et al., 2002). While the military study showed no differences in clinical pregnancy rates (Feinberg et al., 2006), the increase in spontaneous abortions in blacks reduced their live birth rate compared to whites. Two studies included substantial numbers of Asians, and found decreased rates of clinical pregnancy and live birth rates compared to whites (Grainger et al., 2004; Purcell et al., 2007). The two studies that included adequate numbers of Hispanics showed no difference in IVF response compared to whites (Feinberg et al., 2007; Grainger et al., 2004).

Consequences of the Use of IVF for Maternal and Infant Health

There is an increased rate of pregnancy complications associated with IVF pregnancies compared to naturally conceived gestations, largely but not exclusively due to the high rate of multiple gestations. In 2003, 51% of infants conceived through IVF were born in multiple birth deliveries. Consequently, the 1% of US infants conceived through IVF accounted for 18% of multiple births nationwide (Wright et al., 2006). Multifetal pregnancies are associated with increased risks of adverse maternal

health outcomes including pre-eclampsia, preterm labor, hemorrhage, uterine atony and increased surgical deliveries (Gambone, 2006). Infants from multifetal pregnancies are more likely to be delivered preterm and have low birth weight, and these conditions increase the risk of numerous short and long-term sequelae including increased neonatal mortality (four times greater for twins than for singletons) and morbidity (Davis, 2004). In 2003, of the IVF pregnancies that resulted in live births, 14.7% singleton births were born preterm; higher rates were reported in twin gestations (64.9%) and triplets (97.0%) (Wright et al., 2006). Low birth weight rates were also elevated; 9% of singleton births, 56% of twin births and 93.4% of higher order births were of newborns weighing less than 2,500 g (Wright et al., 2006).

The medical risks associated with multiple births have contributed to escalating health care costs. In 2000, the cost per family of multiple deliveries conceived with IVF was estimated to range from $58,865 for twins to $170,282 for triplets (ESHRE Capri Workshop Group, 2000). Another concern about IVF is the increased risk of genetic and structural abnormalities in offspring conceived using this intervention. There is a twofold increased rate of structural abnormalities associated with in vitro conceptions, and a higher rate of diagnosed chromosomal abnormalities (Allen, V.M., & Wilson, R.d., et al., 2006; Schieve, Rasmussen, et al., 2004). It is hard to discern whether these complications are entirely due to IVF technology, or whether they reflect increased risks associated with underlying infertility (Allen, Wilson, et al., 2006; Schieve, Rasmussen, et al., 2004).

Racial/Ethnic Disparities in IVF Birth Outcomes

Recent work at the CDC by Wright et al. (2006) indicates that among IVF singleton births in 2003, blacks were more than twice as likely to deliver infants of low birth weight as whites (16.0 vs. 7.0%) and over four times more likely to deliver very low birth weight infants (5.2 vs. 1.4%). Preterm births were also more likely among blacks (17.7%) than whites (11.7%). The disparities in preterm birth and low birth weight between blacks and whites using IVF reproduce racial disparities found among women who deliver in the general US population.

Hispanic women using IVF in 2003 were more likely to deliver preterm singletons compared to whites (16.2 vs. 11.7%) and to deliver low birth weight infants (12.1 vs. 7%) (Wright et al., 2006). This disparity is larger than found in national estimates for these outcomes. In the general population, Hispanics tend to have preterm birth and low birth weight rates approximately 25% higher than whites (CDC, 1999). The larger disparity in the subpopulation of infertile women may be related to the fact that the mean age of Hispanic IVF users is higher than that of their fertile counterparts. Finally, compared to whites, Asian IVF users had lower rates of preterm births (10%) but higher rates of low birth weight (13%). These data reflect the outcomes of IVF singleton births without controlling for potential confounders. To date, there are no studies of IVF multiple births with rates and outcomes stratified by race/ethnicity**.**

Discussion

According to national estimates, black and Hispanic women have a higher prevalence of infertility than whites, but lower access to infertility services. These disparities pose several public health challenges that must be addressed to narrow inequalities. First, expensive infertility treatments combined with inadequate or unavailable health insurance present major economic deterrents to infertility service utilization. This has resulted in a two-tiered system of infertility care in the U.S. Infertile women with financial privilege are able to seek and obtain treatment, while women facing

financial constraints, including a large proportion of ethnic/racial minorities, many of whom rely on Medicaid, are often unable to access care. The public health challenge is to improve access to timely infertility services to avoid costly and irreversible conditions.

In 2002, only 36% of women with fertility problems reported having ever sought medical attention (Chandra et al, 2005). While 40% of whites with fertility problems reported having received an infertility service, this proportion was far lower for women of color. Among black women, less than one out of three had sought care, while among Hispanics, only one in four had done so. Among service seekers, the majority obtained advice and testing; fewer advanced to medical or surgical treatment of infertility conditions, and only 1% used IVF. Limited evidence suggests that women of color may experience a delayed onset of infertility treatment. In fact, some evidence suggests that women of color utilize fewer infertility services, regardless of whether they have high or low income. However, scant data exist on factors associated with racial/ethnic disparities in utilization and much of the evidence is limited to economic factors.

IVF is predominantly utilized by white affluent women according to the NSFG survey and the CDC dataset (CDC et al., 2007; Chandra et al., 2005). The evidence suggests that when access to IVF is improved, as it is for Department of Defense beneficiaries, racial inequities in service utilization decrease for certain minorities (Feinberg et al., 2006, 2007). Improved access to infertility services by underserved women will require increased access to health insurance as well as more comprehensive insurance coverage. Current state mandates requiring insurers "to offer" rather than "to cover" infertility services have not gone far enough to remove structural barriers to care. A mandate "to offer" requires that health insurance companies make available for purchase a policy which offers coverage of infertility diagnosis and treatment. However, the law does not require employers to purchase such policies, and therefore does not resolve financial barriers. Currently, only 15 states have issued mandates; these vary widely in their scope. Coverage for IVF is specified in nine mandates (Insurance information accurate as of 6/08). These have varying eligibility requirements; one state restricts IVF coverage to couples with 5 years of infertility. Some mandates apply *only* to HMOs, others exclude HMOs. At this time of this writing, employer self-insured health plans were exempt; in 2007 federal legislation (H.R. 2892) was proposed that would require all group health plans that cover obstetric services to include infertility insurance coverage but no further action has been taken on this legislation. None of the current mandates address the plight of infertile women who are unable to obtain health insurance. Poor women on Medicaid have virtually no access to infertility treatment. Prior to 1992, at least 17 states covered low-tech infertility treatment through Medicaid; none covered IVF. Following public outcry about infertility expenditures for the poor, all states eliminated drug treatments for infertility (King & Harrington-Meyer, 1997). Regardless of economic barriers to access, utilization of infertility services also depends upon the extent to which women of color perceive infertility to be a treatable medical problem. Not enough is known about the extent to which poor, underserved women experience infertility as a response to general life adversities and succumb to or accept this condition without seeking care. In contrast, more affluent women may be empowered to overcome infertility and may be more strongly motivated to seek aggressive medical care.

A broader question is whether we should encourage increased access to highly technologic treatment or advocate for greater prevention of risk factors associated with infertility. The public health challenge is to increase preventive efforts towards the reduction of sexually transmitted diseases, optimizing body weight, improving nutritional deficiencies, reducing environmental and work related exposures to reproductive toxins, and counseling all women about the risks of delaying childbearing (Office of Technology Assessment, 1988). Persons requesting family planning should be fully advised about the fertility ramifications of contraceptive choices. Sexually active individuals should be counseled about the risk of sexually transmitted infection, the importance of early treatment of infection, and the implications for future fertility. Persons requesting sterilization procedures should be adequately counseled to minimize subsequent requests for surgical

reversal or IVF. Individuals who desire biologic children should be informed about age-related declines in fertility, ways to optimize fertility, and appropriate guidelines for when to seek medical help. Special attention should be given to specific racial/ethnic groups who are at high risk for acquiring certain conditions or of delaying childbearing. On the part of providers, primary care physicians need to know when to send women for infertility screening and avoid lengthy delays that increase the risk of diminished ovarian function. Early detection and intervention of conditions linked to infertility may prevent the use of more costly, high tech treatments such as in-vitro fertilization.

Another public health challenge is to consider ways to address infertility risk factors on a systemic level outside the healthcare system. It behooves us to ask, for instance, whether policies that limit maternity leave and make it difficult to combine working and parenting influence the decision of many women to postpone childbearing, a choice that results in diminished fertility. Furthermore, have efforts to promote healthy work environments and neighborhoods gone far enough to minimize exposure to reproductive toxins that could impair fertility?

In addition to examining the likelihood of achieving a live birth, it is important from a public health perspective to assess the quality of the live birth associated with IVF. When comparing the incidence of preterm delivery by race/ethnicity among women who used IVF and delivered singleton births, it appears that blacks and Hispanics are at a disadvantage compared to whites (Schieve, Ferre, et al., 2004; Wright et al., 2006). With respect to low birth weight, blacks, Hispanics and Asians all seem to be at a disadvantage. It is not known whether other adverse outcomes such as birth defects and chromosomal abnormalities in offspring conceived through IVF also vary by race/ethnicity. Studies that examine the effects of IVF on birth outcomes are just emerging. Most report crude outcomes and do not adjust for multiple factors.

Multiple births are one of the most critical public health concerns associated with IVF use, due to associated levels of morbidity, mortality, and health care costs. Some advocates argue for the use of singleton live birth rates as an important indicator to monitor success of IVF procedures. Data are not available to determine whether rates of multiple birth among IVF users vary by race and ethnicity.

The higher use of IVF by white women and their greater exposure to procedures that rely on multiple embryo transfers have contributed to recent dramatic increases in multiple births and related poor pregnancy outcomes (Reddy et al., 2007; Russell Petrini, Damus, Mattison, & Schwarz, 2003; Zhang, Meikle, Grainger, & Trumble, 2002). Ironically, this may mean that white-black racial disparities in birth outcomes are narrowing, but for the wrong reasons. Instead of striving for improving birth outcomes of vulnerable black women, we are contributing to increasing adverse birth outcomes among white, socioeconomically advantaged women who can afford expensive infertility treatments. Future studies need to closely monitor trends in birth outcome differentials associated with IVF among ethnic populations and identify ways of closing the gap by improving outcomes for all women.

As IVF has achieved greater pregnancy success, the practice of multiple embryo transfer has been questioned. There is a movement towards single-embryo transfer as a means of reducing rates of high risk pregnancies associated with IVF (Nygren, 2007). Single-embryo transfer may require additional treatment cycles to achieve live birth rates comparable to those of multiple embryo transfer; however, multiple pregnancy rates are almost eliminated when single embryo transfer is utilized (Pandian, Templeton, Serour, & Bhattacharya, 2005). Because of the potential for additional cycles and higher expense associated with single-embryo transfer, this protocol may receive greater endorsement when couples have insurance coverage. It has been demonstrated that in states where IVF services are covered by insurance mandates, a lower average number of embryos are transferred (Reynolds, Schieve, Jeng, & Peterson, 2003). Failure to provide insurance coverage of infertility care may not be cost-effective in the long run. While the cost of an individual IVF cycle may be covered out-of-pocket, the expense of medical care for a high-risk mother and her offspring most often does not escape the insurer. Future studies need to monitor whether disparities in access to single-embryo transfer technologies begin to narrow among racial/ethnic groups and if so, whether

expanded access to this type of IVF technology results in favorable outcomes for both minority and non-minority women.

Future Research

Future studies are needed to collect racial and ethnic specific data on infertility population-level risk factors and on a range of infertility services and responses to treatment. If Asians were over sampled in future rounds of the NSFG, prevalence and utilization rates for that population could be estimated. Furthermore, within Asian and Hispanic populations, data are required that will allow us to disaggregate sub-populations by country of origin, immigration status and time in the United States. We also need better population-level data on cultural, cognitive and motivational factors associated with infertility status and associated medical conditions. As for IVF data, we need to ensure that all IVF clinics report data by race/ethnicity to allow for an analysis of clinical responses of all clinic users. The national CDC/SART database could be enhanced by including IVF outcomes by the patient (and not only by the cycle), as well as information on race, ethnicity, insurance, socioeconomic status and neighborhood factors, and by providing information (e.g., confidence intervals, standard deviations) that would permit an assessment of statistical significance when differences in outcome rates are observed (Steinberg et al., 1998). Given racial/ethnic inequalities in infertility status, risk factors, access to infertility services and responses to treatment, it is imperative that we continue to monitor these dimensions with high quality data.

Conclusion

Research focusing on racial/ethnic disparities in infertility and its treatment is just emerging and much work remains to be done. Current evidence points to marked differentials in outcomes. Relative to white women, black and Hispanic women have higher rates of infertility, yet face more barriers to infertility care. Far less is known about infertility status and care of other minority women, including Asians, Native Americans and Pacific Islanders. Among women who obtain the most intensive and expensive treatment, i.e., IVF, not all succeed in achieving a live birth. Aggregate national data suggest higher IVF live birth rates in whites and Hispanics and lower rates in blacks and Asians. However, reports from individual IVF clinics yield conflicting results, particularly with regards to black patients undergoing IVF. Caution must be exercised when evaluating disparities in responses to IVF because women of color are underrepresented in infertility clinics and success rates are affected by patient and treatment characteristics that have not been controlled in most studies. In sum, it is possible that differential access and use of infertility services by white women and women of color may ultimately affect pregnancy outcomes in ways that research is just beginning to evaluate.

References

Allen, V. M., & Wilson, R. D., et al. (2006). Joint Society of Obstetricians and Gynaecologists of Canada and Canadian Fertility and Andrology Society Guideline: Pregnancy outcomes after assisted reproductive technology. *Journal of Obstetrics and Gynaecology Canada*, *173*, 220–233.

Bendikson, K., Cramer, D. W., Vitonis, A., & Hornstein, M. D. (2005). Ethnic background and in vitro fertilization outcomes. *International Journal of Gynecology and Obstetrics*, *88*(3), 342–3466.

Bitler, M., & Schmidt, L. (2006). Health disparities and infertility: impacts of state-level insurance mandates. *Fertility and Sterility*, *85*(4), 858–865.

Centers for Disease Control (CDC). (1999). Preterm singleton births – United States, 1989–1996. *Mortality and Morbidity Weekly Report*, *48*, 185–189.

Centers for Disease Control (CDC). (2004). *STD Surveillance 2004*. Retrieved August 15, 2006 from http://www.cdc.gov/std/stats/toc2004.htm.

Centers for Disease Control (CDC), American Society for Reproductive Medicine (ASRM), and Society for Assisted Reproductive Technology (SART). (2007). *2005 assisted reproductive technology success rates. Atlanta, GA: US Department of Health and Human Services*. CDC, National Center for Chronic Disease Prevention and Health Promotion.

Chandra, A. (2005). The National Survey of Family Growth as an extension of the Vital Statistics System. *National Center for Health Statistics, Powerpoint presented at NAPHSIS*. Cincinnati, Ohio, June 8, 2005. Available at http://www.naphsis.org/NAPHSIS/files/ccLibraryFiles/Filename/000000000176/NAPHSIS-NSFG%2005.ppt#1.

Chandra, A., Martinez, G. M., Mosher, W. D., Abma, J. C., & Jones, J. (2005). Fertility, family planning, and reproductive health of U.S. women: Data from the 2002 National Survey of Family Growth. *Vital Health Statistics*, *23*(25), 1–160.

Chandra, A., & Stephen, E. (1998). Impaired fecundity in the United States: 1982–1995. *Family Planning Perspectives*, *30*(1), 34–42.

Davis, O. K. (2004). Elective single-embryo transfer – has its time arrived? *New England Journal of Medicine*, *351*(23), 2440–2442.

ESHRE Capri Workshop Group. (2000). Multiple gestation pregnancy. *Human Reproduction*, *15*(7), 1856–1864.

Feinberg, E. C., Larsen, F. W., Catherino, W. H., Zhang, J., & Armstrong, A. Y. (2006). Comparison of assisted reproductive technology utilization and outcomes between Caucasian and African American patients in an equal-access-to-care setting. *Fertility and Sterility*, *85*(4), 888–894.

Feinberg, E. C., Larsen, F. W., Wah, R. M., Alvero, R. J., & Armstrong, A. Y. (2007). Economics may not explain Hispanic underutilization of assisted reproductive technology services. *Fertility and Sterility*, *88*(5), 1439–1441.

Gambone, J. C. (2006). Are adverse pregnancy and fetal outcomes more common with assisted reproductive technologies? What should patients be told? *Clinical Obstetrics and Gynecology*, *49*(1), 123–133.

Gold, R. B., Richards, C. L., Ranji, U., & Salganicoff, A. (2007). *Medicaid's role in family planning*. Kaiser Family Foundation, Issue Brief. (Report #7064–03.)

Grainger, D. A., Seifer, D. B., Frazier, L. M., Rall, M. J., Tjaden, B. L., & Merrill, J. C. (2004). Racial disparity in clinical outcomes from women using advanced reproductive technologies (ART): Analysis of 80,196 ART cycles from the SART database 1999 and 2000. *Fertility and Sterility*, *82*, s37–s38.

Grayson, C. (2003). *In vitro fertilization*. Retrieved July 1, 2006 from http://www.webmd.com/content/article/73/87998.htm

Green, J. A., Robins, J. C., Scheiber, M., Awadally, S., & Thomas, M. A. (2001). Racial and economic demographics of couples seeking infertility treatment. *American Journal of Obstetrics and Gynecology*, *184*, 1080–1082.

H.R. 2892. (2007). *The Family Building Act of 2007*. Library of Congress THOMAS. Retrieved March 31, 2008 from http://thomas.loc.gov/cgi-bin/query/z?c110:H.R.2892.

Heitman, E. (1995). Infertility as a public health problem: why assisted reproductive technologies are not the answer. *Stanford Law and Policy Review*, *6*(2), 89–102.

Inhorn, M. C., & Fakih, M. H. (2006). Arab Americans, African Americans, and infertility: Barriers to reproduction and medical care. *Fertility and Sterility*, *85*(4), 844–852.

Jain, T. (2005). Disparities in access to infertility in a state with mandated insurance coverage. *Fertility and Sterility*, *85*(1), 221–223.

Jain, T. (2006). Socioeconomic and racial disparities among infertility patients seeking care. *Fertility and Sterility*, *85*(4), 876–881.

James, C. E., Hammond, K. R., & Steinkampf, M. P. (2002). Race and assisted reproduction: A case-controlled study of outcomes in African-American and Caucasian women [abstract]. *Fertility and Sterility*, *78*(suppl 1), S123.

Jenkins, R. L. (2005). Ensuring access to education and services on infertility for the underserved. *Journal of the National Cancer Institute. Monographs*, *34*, 101–103.

Kaiser Commission on Medicaid and the Uninsured and the Women's Health Policy Program. (2001). *Medicaid coverage of family planning services: Results of a national survey.* [Electronic version]. Publication No. 2216. Retrieved March 3, 2008 from http://www.kff.org/womenshealth/2216-index.cfm.

Kalmuss, D. S. (1987). The use of infertility services among fertility-impaired couples. *Demography, 24*(4), 575–585.

King, R. B., & Davis, J. (2006). Introduction: health disparities in infertility. *Fertility and Sterility*, *85*(4), 842–843.

King, L., & Harrington-Meyer, M. (1997). The politics of reproductive benefits. *Gender and Society*, *11*(1), 8–30

Lashen, H., Afnan, M., & Sharif, K. (1999). A controlled comparison of ovarian response to controlled stimulation in first generation Asian women compared with white Caucasians undergoing in vitro fertilization. *British Journal of Obstetrics and Gynaecology*, *106*(5), 407–409.

Mahmud, G., Lopez Bernal, A., Yudkin, P., Ledger, W., & Barlow, D. H. (1995). A controlled assessment of the in vitro fertilization performance of British women of Indian origin compared with white women. *Fertility and Sterility*, *64*(1), 103–106, 199.

Mosher, W. D., & Bachrach, C. A. (1996). Understanding U.S. fertility: Continuity and change in the National Survey of Family Growth, 1988–1995. *Family Planning Perspectives*, *28*(1), 4–12.

Nichols, J. E. Jr., Higdon, H. L. 3rd, Crane, M. M. 4th, & Boone, W. R. (2001). Comparison of implantation and pregnancy rates in African American and white women in an assisted reproductive technology practice. *Fertility and Sterility*, *76*(1), 80–84.

Nygren, K. G. (2007). Single embryo transfer: The role of natural cycle/minimal stimulation IVF in the future. *Reproductive Biomedical Online*, *14*(5), 626–627.

Office of Technology Assessment. (1988). *Infertility: Medical and social choices*. Washington, DC: U.S. Government Printing Office.

Pandian, Z., Templeton, A., Serour, G., & Bhattacharya, S. (2005). Number of embryos for Transfer after IVF and ICSI: A Cochrane review. *Human Reproduction*, *20*(10), 2681–2687.

Parham, G. P., & Hicks, M. L. (2005). Racial disparities affecting the reproductive health of African-American women. *The Medical Clinics of North America*, *89*(5), 935–943.

Purcell, K. J., Schembri, M., Frazier, L. M., Rall, M. J., Shen, S., Croughan, M., et al. (2007). Asian ethnicity is associated with reduced pregnancy outcomes after assisted reproductive technology. *Fertility and Sterility*, *87*(2), 297–302.

Reddy, U. M., Wapner, R. J., Rebar, R. W., & Tasca, R. J. (2007). Infertility, assisted reproductive technology and adverse pregnancy outcomes: executive summary of a National Institute of Child Health and Human Development workshop. *Obstetrics and Gynecology*, *109*(4), 967–977.

Reynolds, M. A., Schieve, L. A., Jeng, G., & Peterson, H. B. (2003). Does insurance coverage decrease the risk for multiple births associated with assisted reproductive technology? *Fertility and Sterility*, *80*(1), 16–23.

Russell, R. B., Petrini, J. R., Damus, K., Mattison, D. R., & Schwarz, H. (2003). The changing epidemiology of multiple births in the United States. *Obstetrics and Gynecology*, *101*(1), 129–135.

Schieve, L. A., Ferre, C., Peterson, H. B., Macaluso, M., Reynolds, M. A., & Wright, V. C. (2004). Perinatal outcome among singleton infants conceived through assisted reproductive technology in the United States. *Obstetrics and Gynecology*, *103*(6), 1144–1153.

Schieve, L. A., Rasmussen, S. A., Buck, G. M., Schendel, D. E., Reynolds, M. A., & Wright, V. C. (2004). Are children born after assisted reproductive technology at increased risk for adverse health outcomes? *Obstetrics and Gynecology*, *103*(6), 1154–1163.

Schulman, J. D. (2007). "What's your success rate?": Understanding IVF pregnancy statistics: Part I. Retrieved July 7, 2007 from http://www.givf.com/library/whatisyoursuccessrate1.cfm.

Seifer, D. B., Frazier, L. M., & Grainger, D. C. (2008). Disparity in assisted reproductive technologies outcomes in black women compared to white women. *Fertility and Sterility*, *90*(5), 1701–1710.

Sharara, F., & McClamrock, H. (2000). Differences in in vitro fertilization (IVF) outcome between white and black women in an inner-city, university based IVF program. *Fertility and Sterility*, *73*(6), 1170–1173.

Steinberg, E. P., Holtz, P. M., Sullivan, E. M., & Villar, C. P. (1998). Profiling assisted reproductive technology: outcomes and quality of infertility management. *Fertility and Sterility*, *69*(4), 617–623.

Stephen, E. H., & Chandra, A. Infertility service utilization among women aged 15–44 in the United States, 2002. Poster presented at the annual meeting of the Population Association of America, March 30–April 11, Los Angeles, CA, 2006.

Stephen, E. H., & Chandra, A. (2006). Use of infertility services in the United States: 1995. *Family Planning Perspectives*, *32*(3), 132–137.

Wang, J. X., Norman, R. J., & Kristiansson, P. (2002). The effect of various infertility treatments on the risk of preterm birth. *Human Reproduction*, *17*(4), 945–949.

White, L., McQuillan, J., & Greil, A. L. (2006). Explaining disparities in treatment seeking: the case of infertility. *Fertility and Sterility*, *85*(4), 853–857.

Wright, V. C., Chang, J., Jeng, G., & Macaluso, M. (2006). Assisted reproductive technology surveillance – United States, 2003. *Morbidity and Mortality Weekly Report: Surveillance Summaries*, *55*(SS04), 1–22.

Wright, V. C., Schieve, L. A., Reynolds, M. A., & Jeng, G. (2005). Assisted reproductive technology surveillance – United States, 2002. *Morbidity and Mortality Weekly Report: Surveillance Summaries*, *3*52(SS02), 1–24.

Zhang, J., Meikle, S., Grainger, D. A., & Trumble, A. (2002). Multifetal pregnancy in older women and perinatal outcomes. *Fertility and Sterility*, *78*(3), 562–568.

Chapter 7
Public Health Interventions for Perinatal HIV and STI Screening in Pregnancy

Whitney B. You, Ann E.B. Borders, Beth F. Pellettieri, and Patricia M. Garcia

Over the last two decades, much of the public health focus regarding sexually transmitted diseases in the perinatal period has been on the prevention of HIV transmission from mothers to infants. The development of the evidence base in support of screening policies for HIV infection during pregnancy has a rich and complex history (especially given the relatively short time period), and there continue to be policy challenges related to the implementation of screening programs. This chapter focuses on both the history and body of evidence supporting various approaches to perinatal HIV screening, with a particular focus on the United States.

We also present background information on screening policies and practices for other STIs, in order to place the debates that have surrounded perinatal HIV screening within the historical context of general STI policies. This perspective highlights a key commonality: decisions on whether and how to screen for HIV and other STIs have taken place against a backdrop of societal views on morality and stigma associated with sexuality, especially in relation to women.

Background

Though screening for certain sexually transmitted infections (STIs) is now considered an expected part of prenatal care for many women, the best way to incorporate HIV screening has been the subject of debate for nearly two decades. It is helpful to remember that similar conflicts arose over early efforts to screen for other STIs (especially syphilis). The history of successful public health efforts to control STIs is as much a history of the battle against stigma as it is the history of developments in diagnostic testing and treatment. Revisiting the social history of STI screening and contemporary guidelines provides a useful background for reviewing the evidence and screening guidelines for HIV during pregnancy.

The devastating impact of STIs on the health of women and children was first recognized in the Victorian Era (Brandt, 1987). Syphilis was considered a major concern because of its diagnostic complexity, severe symptoms, and profound effects on infected infants (Poirier, 1995). Wives and children were seen as "innocent victims" of promiscuous men who consorted with prostitutes before and during marriage. It was this issue of morality that brought sexually transmitted infections into the public eye, and prompted government responses towards prevention and control (Brandt; Poirier).

A.E.B. Borders (✉)
Division of Maternal-Fetal Medicine, Department of Obstetrics and Gynecology,
Institute for Healthcare Studies, Northwestern University Feinberg School of Medicine, Chicago, IL, USA
e-mail: abryant@md.northwestern.edu

A. Handler et al. (eds.), *Reducing Racial/Ethnic Disparities in Reproductive and Perinatal Outcomes: The Evidence from Population-Based Interventions*, DOI 10.1007/978-1-4419-1499-6_7,

Many of the “anti-venereal disease” (i.e., anti-STI) campaigns at the turn of the century targeted prostitutes and their customers, and were characterized by messages of stigma and fear (Poirier, 1995). The first effective screening test for syphilis (the Wasserman test) became available in 1906; however, at that time there was no effective or widely available treatment (Morabia & Zhang, 2004). Both stigma and the lack of acceptable treatment options discouraged the screening of pregnant women as part of routine obstetric care (Faden, Geller, & Powers, 1991). Meanwhile, with the start of World War I, the Armed Services began to educate soldiers about venereal diseases, labeling failure to avoid them, “a neglect of duty” (Poirier). However, despite the high prevalence of STIs, soldiers during World War I were not universally screened for syphilis (Faden et al.; Morabia & Zhang).

In 1936, the U.S. Surgeon General announced a five-point plan aimed at controlling and preventing syphilis. This included mass education, as well as recommended screenings before marriage and early in pregnancy (Faden et al., 1991; Poirier, 1995). The increased awareness also focused attention on congenital syphilis, which was thought to be the leading cause of spontaneous abortion and stillbirth (Faden et al.). Despite the recommendations, many pregnant women still were not screened for syphilis, presumably in part because physicians were afraid of offending patients with such a recommendation (Poirier). Public health professionals argued tirelessly that successful control of venereal diseases required treating them like any other infectious disease, rather than as a failure of morality (Brandt, 1987).

Eventually, due in large part to an aggressive public education campaign spearheaded by the Surgeon General, New York in 1938 became the first state to legislate mandatory syphilis screenings for pregnant women, and other states soon followed (Faden et al., 1991; Green, Talbot, & Morton, 2001). With the discovery of penicillin in 1944, all soldiers in WWII were screened and treated for syphilis, resulting in an almost complete eradication of this infection in the U.S. armed forces (Morabia & Zhang, 2004). By 1945, as penicillin became more widely available to the public, 36 states passed laws requiring prenatal screening for syphilis. A decade later, congenital syphilis was no longer deemed a critical public health issue (Faden et al.).

The history of syphilis screening and treatment in the U.S. has striking parallels to recent developments related to screening/treatment for HIV/AIDS. In both cases, association of the disease with stigmatized persons and activities resulted in reticence of individuals to be screened, as well as fierce social and political battles over public policy related to screening and treatment during pregnancy. Both situations also highlight the important role of advocacy on behalf of women and children in the implementation of screening recommendations and policies. With changes in disease incidence and prevalence, and evolving recognition of association of genital infections with adverse pregnancy outcomes, health professionals and policy makers are continually challenged to reassess the need for and appropriate timing of STI screening in pregnancy.

Though syphilis and HIV are particularly dramatic examples of serious health threats, other STIs also pose substantial dangers to women and infants. Many sexually transmitted infections are asymptomatic in women, and when undetected and untreated can cause cancer, pelvic inflammatory disease, ectopic pregnancy, infertility, and other serious health consequences (Eng & Butler, 1997). Failure to treat sexually transmitted infections during pregnancy is associated with premature rupture of membranes and preterm delivery, with accompanying increase in risk for stillbirth and low birth weight (Locksmith & Duff, 2004).

Additional risks of untreated STIs in pregnancy include postpartum infection and puerperal sepsis in the mother, and neonatal conjunctivitis, pneumonia, neonatal sepsis, acute hepatitis, neurologic damage, and other congenital abnormalities in the infant (Eng & Butler, 1997; Newell & McIntyre, 2000). For many women who lack access to routine preventive health care, prenatal screening is often the first opportunity for diagnosis and treatment of sexually transmitted infections. Prenatal STI screening allows for a variety of interventions, including treatment of the disease condition in the mother, modification of delivery practices to minimize the risk of vertical transmission, and postnatal treatment or prophylaxis for the infant.

Below we provide a more detailed overview of the history and current status of routine prenatal screening for four STIs: syphilis, chlamydia, gonorrhea, and hepatitis B. These were selected because they are relatively common and/or reportable STIs characterized by racial/ethnic disparities, because their sequelae affect both mothers and infants, and because multiple professional organizations have published screening guidelines to address them. Other STIs that pose risk for adverse perinatal outcomes such as hepatitis and herpes simplex virus (HSV) are not dealt with here. Bacterial vaginosis, which is not a "classic" STI but can be related to sexual activity, is discussed in Chap. 9. Following this brief overview of screening for STIs during pregnancy, we provide a more systematic review of the evidence informing screening and treatment for HIV/AIDS during pregnancy, a topic which has been the focus of several landmark clinical and public health studies over the last 20 years. In both the overview of STI screening/treatment and the review of HIV screening/treatment literature, we focus on both the evidence base for the effectiveness of screening and treatment in pregnancy for preventing adverse outcomes, and the importance of considering the role of stigma when evaluating alternative screening options. This chapter also briefly addresses the potential of increased perinatal HIV screening to reduce racial/ethnic disparities in health status for women and infants.

Recent History of STI Screening In Pregnancy

In 1989, the U.S. Department of Health and Human Services published "Caring for Our Future: The Content of Prenatal Care," which concluded that sufficient evidence existed to recommend routine screening of all pregnant women for gonorrhea, hepatitis B, and syphilis, along with targeted chlamydia screening for gravid women deemed to be at increased risk. Since that time, a number of agencies and associations, including the American College of Obstetricians and Gynecologists (ACOG), the American Academy of Family Physicians (AAFP), the U.S. Preventive Services Task Force (USPSTF), and the Centers for Disease Control and Prevention (CDC) have provided STI screening guidelines and recommendations to clinicians providing prenatal care (CDC, 2006a; Meyers et al., 2008). Although there is substantial overlap, there is some variation in recommendations among these groups. These differences reflect varying degrees of emphasis on the evidence base for efficacy of screening, including criteria such as disease severity, prevalence in the population, sensitivity, specificity, cost of screening, and potential harms related to inaccurate diagnoses and unnecessary treatments. The discussion below focuses on the recommendations laid out in the 2006 CDC Sexually Transmitted Diseases Treatment Guidelines, since these tend to be the most inclusive, and those of the U.S. Preventive Services Task Force (Meyers et al.), since these place strong emphasis on the evidence base. Following are more specific details on the guidelines and evidence for selected STI screening practices, followed by a more detailed discussion of screening for HIV.

Syphilis

Syphilis is a bacterial infection caused by the spirochete *Treponema pallidum*, and is almost always transmitted sexually or vertically from mother to infant. Syphilis can be transmitted through the placenta at any time during pregnancy, and recommendations have been consistent over time and across agencies that all pregnant women should be screened for syphilis at the first prenatal visit. The USPSTF gave this recommendation an 'A' rating, citing good evidence that screening of pregnant women results in a reduction in proportion of infants with positive serologies and clinical

manifestations of syphilis (USPSTF, 2004a). The CDC further recommends that in communities and populations at elevated risk, testing of women should be repeated at 28 weeks gestation and at birth. Penicillin is used both to treat the mother and prevent congenital syphilis in the infant.

Universal screening for syphilis in pregnancy has been a public health success story, with the congenital syphilis incidence rate declining steadily over a 14 year period beginning in 1990. However, there has been a disturbing resurgence in recent years, with the incidence rate increasing by 28% between 2005 and 2007, reflecting an overall increase in rates of primary and secondary syphilis in the general population. There is substantial racial and ethnic disparity in this resurgence, with African-Americans having an incidence rate of primary and secondary syphilis seven times that of Whites, and a rate of congenital syphilis 14 times that of Whites, and twice as high as Hispanics (CDC, 2009a). Some have suggested that the increase in syphilis infections is related to the large global and domestic focus on and prioritization of funding for HIV, which has marginalized prevention and control efforts for other infectious diseases (López-Zambrano, Briceño, & Rodriguez-Morales, 2009).

The resurgence in congenital syphilis despite policies of universal screening can be explained in part by the failure of the U.S. health system to engage all pregnant women in adequate prenatal care. In an analysis of national data from 2002, the CDC found that 29% of infants with congenital syphilis were born to women who did not receive prenatal care. Of those women in the sample who did receive prenatal care, only 30% did so in the first trimester (CDC, 2004). These figures are particularly relevant for addressing racial and ethnic disparities in congenital syphilis and other adverse perinatal outcomes associated with syphilis infection, given that African-American women are less likely than White women to receive early or any prenatal care and are therefore less likely to have a syphilis infection detected and treated in a timely fashion (Martin et al., 2009).

Inadequate adherence to screening recommendations by health care providers has also been identified as contributing to elevated rates of congenital syphilis (CDC, 2004; Trepka & Bloom, 2006). A national survey of a random sample of 7,300 private-sector physicians found that only 32% screened their pregnant patients for syphilis. The rate for obstetricians/gynecologists was significantly higher at about 85%, but still substantially below the recommendation for universal screening (St. Lawrence et al., 2002). These findings suggest that outside the STI clinic environment, there is still reluctance on the part of physicians to offer syphilis testing to their patients.

There is evidence of change over time in patterns of screening by race and ethnicity. A study of Medicaid-covered deliveries in Florida found that in 1995, African-Americans had the highest rates of any, early, and repeat syphilis screenings compared to other racial and ethnic groups (Fowler, Gavin, Adams, Tao, & Chireau, 2008). By 2000, screening rates among all groups had increased; however, Whites emerged as having the highest rates of any and early screening, and African-Americans had the lowest rates. Rates for repeat screening were similar across all groups. The authors found that two factors influenced prenatal syphilis screening rates: the timing of women's enrollment in Medicaid and the source of the prenatal care. These findings are particularly noteworthy, as Florida law requires health care providers to screen all women for syphilis twice in the prenatal period (during the first and third trimester), and again at delivery, if a previous test was positive (Hollier, Hill, Sheffield, & Wendel, 2003). Further studies are needed in order to understand the extent to which disparities in screening practices underlie disparities in rates of primary, secondary and congenital syphilis.

Chlamydia

Chlamydia, caused by the bacterium *Chlamydia trachomatis*, is the most common STI in the United States, with approximately three million new cases occurring annually (Coonrod et al., 2008).

Between 2003 and 2007, rates increased steadily for both men and women, reflecting increased reporting and possibly a true increase in morbidity (CDC, 2009a). In 2007, surveillance data showed a median state-specific test positivity of 6.9% for women age 15–24 screened at family planning clinics, and 7.4% for women in the same age group screened at prenatal clinics (CDC). There are substantial racial, ethnic and socioeconomic disparities in chlamydia infection rates. Using data from Wave I (1994–1995) and Wave III (2001–2002) of the National Longitudinal Study of Adolescent Health, Miller and colleagues found the prevalence of chlamydia to be six times higher in young African-American men and women than in young White adults (Miller et al., 2004). High rates in pregnancy have been found in underserved communities. For example, a repeat-testing study among pregnant women attending an inner-city clinic in New Orleans found an overall test positivity of 17.8%, with a substantial proportion of women testing positive after an initial negative test, and others experiencing treatment failure and/or re-infection (Miller, Maupin, & Nsuami, 2005).

Screening is especially important for chlamydia, since 70–90% of infected women are asymptomatic (Coonrod et al., 2008) and untreated chlamydia infection frequently leads to pelvic inflammatory disease (PID), infertility, and an elevated risk of contracting or transmitting HIV. During pregnancy, transmission of chlamydia from mother to infant occurs in 60–70% of births among infected women (Merkatz & Thompson, 1990; Newell & McIntyre, 2000). An untreated infection in pregnancy increases the neonate's risk of preterm delivery, stillbirth, pneumonia and eye infection, and the mother's risk for endometritis (Beem & Saxon, 1977; Berenson, Hammill, Martens, & Faro, 1990; Hallberg, Mårdh, Persson, & Ripa, 1979). There is evidence of racial/ethnic disparities both in prevalence of chlamydia infections and the associated risks. In a large prospective cohort study, pregnant African-American women were more likely than other groups to have lower genital tract infections including chlamydia at enrollment, and given an existing infection they were more likely to experience a preterm birth (Hitti et al., 2007). A retrospective cohort study found chlamydia infection to be associated with an increased risk of preterm birth and premature rupture of membranes (PROM), and to be more common in non-White women (Blas, Canchihuaman, Alva, & Hawes, 2007).

Chlamydia testing during pregnancy allows for appropriate antibiotic therapy and thus avoidance of perinatal and postpartum complications. However, there is not complete consensus on chlamydia screening recommendations among those agencies that provide practice guidelines. The CDC (2006a) and ACOG (2007) now recommend chlamydia screening of all pregnant women in the first trimester, with repeat screening near term for at-risk women and those under the age of 25. This is a departure from earlier recommendations that only at-risk pregnant women should be screened (CDC, 2002a). The 2007 USPSTF statement and the 2005 American Academy of Family Physicians (AAFP) statement recommend routine screening in pregnancy only for high-risk women (ACOG/AAP, 2007; CDC, 2006a; Kirkham, Harris, & Grzybowski, 2005; Meyers et al., 2008). The difference in recommendations seems to reflect different emphases on the impact of new screening technologies, as well as on risks associated with screening. The nucleic acid amplification test (NAAT) for chlamydia has the benefit of a relatively high sensitivity, and because it can be used with a urine sample it does not require a pelvic exam (CDC, 2002b; Kohl, Markowitz, & Koumans, 2003). The CDC and ACOG seem to have reached the conclusion that these advantages, combined with the rising prevalence of chlamydia and the health threats to mothers and infants tilt the balance in favor of routine testing as part of prenatal care. The USPSTF, in contrast, focuses on the paucity of studies providing direct evidence on the effectiveness of routine screening in low-risk populations for reducing adverse health outcomes. Based on a systematic review conducted in 2005 (Meyers, Halvorson, & Luckhaupt, 2007), the USPSTF concluded that evidence is good for the effectiveness of screening for at-risk non-pregnant women, defined as women age 24 years and younger, those engaging in high-risk sexual activity, and those living in a high prevalence community (Meyers et al., 2008). The USPSTF also utilizes these findings for at-risk non-pregnant women as a basis for a recommendation to routinely screen at-risk pregnant women, even though "there are no studies evaluating the effectiveness of

screening for chlamydial infection in pregnant women who are at increased risk." For pregnant women at low to moderate risk, however, USPSTF concluded that the rate of false positives would be unacceptably high, and that the accompanying risks of anxiety, relationship problems, and unnecessary treatment mitigate against routine testing of those women (Meyers et al.).

Gonorrhea

Gonorrhea, caused by the bacterium *Neisseria gonorrhoeae*, is the second most common STI, with an estimated 700,000 new cases occurring in the U.S. each year. Between 1975 and 1997, the incidence rate declined 74%, and has remained relatively stable since then. Among 15–25 year-old pregnant women, the rate of gonorrhea was 0.8% in 2007 (CDC, 2009a). As with other STIs, there are racial and ethnic differences in the prevalence of gonorrhea. Rates in non-pregnant African-American women are 15 times higher than in White women (CDC). There is a high co-infection rate of gonorrhea and chlamydia (CDC, 2006a).

Transmission of gonorrhea from mother to infant occurs in 30% of births among infected women (Merkatz & Thompson, 1990). For infants, gonorrhea is the most common cause of neonatal conjunctivitis, which can lead to blindness (Merkatz & Thompson). Treatment of the mother's infection can prevent endometritis and infertility, poor pregnancy outcomes (such as stillbirth and prematurity), and transmission of infection from mother to infant (Newell & McIntyre, 2000). Use of ocular prophylaxis prevents conjunctivitis in infants in cases where infection in the mother was not detected or adequately treated during pregnancy.

Prior to 1998, gonorrhea screening was recommended for all pregnant women. However, more recent guidelines from the CDC recommend that only women who have personal risk or live in an area of high prevalence should be tested for *Neisseria gonorrhea* in the absence of symptoms. Where risk is considered high, repeat testing in the third trimester is recommended (CDC, 1998, 2006a). The USPSTF has declined to make a recommendation on universal gonorrhea screening in pregnancy, citing insufficient evidence to make a recommendation either for or against the practice (an 'I' rating). Even with high test specificity, the USPSTF reasons that two thirds of positive screening tests would be expected to be false positives, thus making it difficult to argue that the benefits of screening would outweigh the possible harms (USPSTF, 2005).

In addition to prenatal screening of women considered to be at elevated risk for gonorrhea, the CDC recommends routine use of ocular prophylaxis to prevent conjunctivitis in all infants, and use of ocular prophylaxis is required in most states by law (CDC, 2006a). Despite these legally binding recommendations, there is limited evidence on the extent of adherence to recommendations for routine newborn prophylaxis (Moyer & Butler, 2004). Furthermore, it could be argued that by requiring universal treatment of infants but not universal screening of mothers, undue emphasis is placed on the health and welfare of infants to the neglect of the health and welfare of women.

Hepatitis B

Hepatitis B is a blood-borne disease caused by infection with the hepatitis B virus (HBV), which is most commonly transmitted through sexual contact, injection drug use, or from mother to child during pregnancy or birth. It can be a mild acute illness that lasts a few weeks followed by a full recovery, or it can become a chronic infection, with the carrier often remaining asymptomatic but sometimes becoming severely ill with liver disease or liver cancer. North America and Northern and Western Europe have the lowest prevalence globally; the highest rates are found in sub-Saharan

Africa and parts of Asia and South America (Newell & McIntyre, 2000). In the United States, there was an 81% decrease in the incidence of hepatitis B between 1990 and 2006, with the most marked decline (98%) among persons under the age of 15, reflecting increasing immunization rates (CDC, 2006b). Currently, the incidence of hepatitis B in the U.S. is 1.6 cases per 100,000 (CDC). Foreign-born Asian/Pacific Islanders have a Hepatitis B prevalence rate of 10%, dramatically higher than other racial/ethnic groups in the U.S. (Chang & So, 2007).

The risk of developing chronic hepatitis B infection is higher when infection occurs at younger ages. While a majority of persons infected as adults recover from an HBV infection, an estimated 80–90% of infected infants remain chronically infected if untreated (Newell & McIntyre, 2000). Most infants with hepatitis B are born to mothers with chronic, asymptomatic hepatitis B. Women with a positive surface antigen (HBsAg) and e antigen (HBeAg) have a 70–90% chance of transmitting the hepatitis B virus to their infant (ACOG/AAP, 2007; Newell & McIntyre).

Routine screening of pregnant women is a proven and cost-effective strategy for prevention of perinatal HBV transmission (Arevalo & Washington, 1988). All major professional bodies including CDC, USPSTF, ACOG and AAFP recommend HBV screening at the first prenatal visit; the recommendation has an 'A' rating from the USPSTF (Meyers et al., 2008; USPSTF, 2004b). Repeat testing is recommended for women at increased risk, defined as those with multiple partners, those with a known HBsAg-positive partner, and those who have engaged in intravenous-drug use (Greenspoon, Martin, Greenspoon, & McNamara, 1989). Routine vaccination of infants is recommended, along with administration of hepatitis B immune globulin (HBIG) within 12 h after birth for infants of infected mothers (ACOG/AAP, 2007; Newell & McIntyre, 2000). This combination of passive and active immunization has 85–95% efficacy for preventing vertical transmission (ACOG, 2007).

Adherence to recommendations for HBV screening in pregnancy has improved dramatically over the last decade. A study published in 2000 reported that nearly half of hospitals surveyed had no policies in place regarding HBV screening in pregnancy (Bath et al., 2000). In contrast, more recent studies by Schrag et al. (2003) and Sheikh, Sarnquist, Grieb, Sullivan, and Maldonado (2009) report screening rates of around 95%, among the highest for all infectious diseases. Lack of prenatal care is the factor most consistently associated with failure of women to be screened during pregnancy (Schrag et al.; Sheikh et al.). In the study by Schrag and colleagues, African-American women were less likely to receive prenatal care and be tested for hepatitis B, and when tested were more likely to test positive for the infection.

STI Screening Summary

As noted above, while screening during pregnancy is recommended for the most common sexually transmitted infections, there is variation in whether the recommendations are for universal screening as a part of routine prenatal care, or selective screening based on the risk profile of the patient. While for the most part there is consensus among recommendations of professional organizations, there is some variation such as in the case of chlamydia, where CDC and ACOG recommend universal screening of pregnant women, while the USPSTF and AAFP recommend risk-based screening. There are differences for gonorrhea as well, with CDC recommending risk-based screening and the USPSTF taking no position. Where risk-based screening is recommended, risk criteria vary, ranging from age, which is relatively simple for the clinician to determine, to factors like number and risk status of sexual partners, or community-level disease prevalence, which can be more problematic to ascertain. Risk factors tend to be stigmatized behaviors about which the clinician may be uncomfortable inquiring, and the patient may be uncomfortable disclosing them. In a qualitative study conducted in Australia to inform development of a chlamydia screening program, young women indicated support for age-based screening but not screening based on risk behavior. They objected

to being asked for their sexual history as part of routine medical care, and in some cases, indicated that they probably would not respond truthfully to such requests. The women expressed a strong desire for STI testing to be "normalized," which they felt would reduce stigma, increase adherence, and ultimately improve the health status of the population (Pavlin, Parker, Fairley, Gunn, & Hocking, 2008). While this study did not specifically address the issue of STI testing in pregnancy, it highlights stigma as an impediment to effective control of sexually transmitted infectious diseases. It can be argued that studies such as this constitute "evidence" that should be considered in policy decisions about STI screening along with more traditional evidence from randomized controlled trials and data on disease prevalence and test sensitivity and specificity.

Perinatal HIV Screening During Pregnancy

The Centers for Disease Control and Prevention (CDC) estimate that approximately one million U.S. residents are living with HIV, and that about one quarter are not aware of their infection (CDC, 2008). Over 56,000 persons were estimated to have been newly infected with HIV in 2006. People of color continue to bear a disproportionate burden of this disease. Nearly half of those with HIV diagnoses are Black, though Blacks make up only 13% of the population. Black women have a rate of HIV diagnosis 19 times that of White women, and for Hispanic women the rate is five times that of Whites. The burden of perinatal HIV infection is also disproportionately born by women and infants of color (CDC, 2007). An estimated 142 cases of perinatal HIV transmission occurred in the United States in the year 2005. Data on race and ethnicity are available for 111 of the cases; 96 were infants of color.

Maximal reduction of perinatal HIV infection is one of the four primary goals of CDC's Advancing HIV Prevention initiative (CDC, 2003). There has been substantial progress to date, a result of rapid scientific advancement coupled with efficient translation of these findings to the clinical setting. The use of maternal and neonatal antiretroviral therapy, cesarean section, and avoidance of breastfeeding have diminished the perinatal transmission rate to 1%. This impressive achievement is the result of government leadership along with the development of practice guidelines by professional organizations. However, obstacles still remain giving way to missed opportunities for prevention. In an era in which effective clinical interventions exist to dramatically decrease the risk of perinatal HIV transmission, infants continue to be born with HIV infections due to failure to appropriately diagnose and treat HIV infected pregnant women. This failure contributes to the persistent racial and ethnic disparities in the burden of HIV disease in the U.S.

The remainder of this chapter provides: (a) an overview of the evolution of policies in the U.S. that have addressed perinatal HIV transmission, (b) examples of various approaches that individual states have taken toward HIV screening in pregnancy, (c) the evidence supporting these national and state policies, and (d) implications of various policies and approaches for reducing HIV disparities.

Evolution of Policy

Table 7.1 outlines the evolution of events and policies that have influenced how the United States has addressed perinatal HIV transmission.

Based on the results of the "PACTG 076" perinatal transmission trials, the CDC issued recommendations in 1994 supporting maternal and neonatal administration of zidovudine (ZDV) for prevention of vertical transmission. Effective prevention of perinatal transmission not only requires appropriate therapy but identification of those who warrant that therapy. Timely and accurate

Table 7.1 Evolution of policies to address HIV in pregnancy

Year	Event
1981	AIDS is first detected in the United States
1982	First pediatric case of AIDS reported
1990	The Ryan White CARE Act – an act that would provide grants to improve the quality and availability of care for individuals and families with HIV disease – is passed in congress
1992	Approximately 945 cases of perinatally acquired AIDS cases in the United States
1994	A large European study on mother to child transmission shows that delivering via C-section halves the rate of HIV transmission
1994	ACTG 076
1995	The USPHS develops guidelines that call for the counseling of all pregnant women about the risk of AIDS and the benefits of HIV testing (opt-in approach)
1996	The Ryan White CARE Act Amendments include a provision for having state funding contingent upon mandatory testing of newborns
1999	IOM releases its report on evaluating states' efforts in reducing perinatal transmission of HIV and recommends universal testing as a routine part of prenatal care (opt-out approach)
1999	ACOG and AAP recommend an opt-out approach to HIV testing in pregnancy
2002	FDA approves rapid HIV test that can be used to test mothers during childbirth
2004	ACOG and CDC issue recommendations for rapid HIV testing during labor and delivery

diagnosis of an HIV infection during pregnancy allows for initiation of antiretroviral therapy in the antepartum period and zidovudine administration intrapartum. Identification of seropositive mothers additionally permits appropriate administration of ZDV to infants and counseling regarding the avoidance of breastfeeding. As such, in 1995 the CDC recognized that specific services and treatment must be offered to HIV-infected pregnant women to prevent perinatal transmission, and issued guidelines that recommended perinatal counseling and voluntary testing (CDC, 1995).

In 1999, the Institute of Medicine (IOM) issued a report stating that while progress had been made, the number of children born with HIV "continued to be far above what is achievable." The IOM recommended "the adoption of a national policy of universal HIV testing, with patient notification, as a routine component of prenatal care." The IOM stated that this testing strategy offers several advantages in that it (a) avoids the need for women to disclose high-risk behaviors that opens them to potential stigmatization, (b) reduces the need for extensive pretest counseling and discussions about personal risks (c) decreases the physicians' risk of incorrectly concluding that a women is not at risk for HIV infection, (d) eliminates the stigma associated with being "singled out" for testing, and (e) overcomes the problem that women are missed when a risk-based or prevalence-based testing strategy is employed (IOM, 1999).

The release of this IOM report marked a shift from an opt-in to an opt-out strategy for HIV testing in pregnancy. The opt-in approach is characterized by offering an HIV test once informed consent has been obtained formally and pretest counseling provided. With opt-out testing, women are informed that their standard battery of prenatal labs includes an HIV test. Patients are then given the opportunity to opt-out of the HIV test (Branson et al., 2006). The opt-out strategy makes HIV testing routine and does not require formalized pre-test counseling or written informed consent. This strategy also has the potential to reach thousands of women who would not have been identified as high-risk by practitioners using the opt-in method.

In 1999 ACOG along with AAP recommended the opt-out approach to prenatal HIV testing The CDC followed with the same recommendations in 2006 (Branson et al., 2006). Despite these recommendations, the majority of states have failed to put this strategy into practice (CDC, 2006a). As of 2009, 22 states have opt-out testing policies, and 28 plus Puerto Rico and the District of Columbia have opt-in policies (Kaiser Family Foundation, 2009).

Although these practice guidelines and policy statements resulted in significant strides in the effort to reduce perinatal HIV transmission, the CDC reported that between 2001 and 2004, 7% of HIV-infected women remained undiagnosed at the time of delivery (CDC, 2006a). With the knowledge that abbreviated courses of antiretroviral therapy could reduce perinatal HIV transmission, a concerted effort was undertaken to develop approaches to assess maternal HIV status for women who remained untested at the time of delivery.

In 2002, the FDA approved a rapid HIV test kit that allows for an accurate and efficient determination of a patient's serostatus (CDC, 2002c). Thus, women who had received no prenatal care or did not receive HIV testing could be tested during labor and delivery with the goal of administering intrapartum maternal and then neonatal zidovudine to HIV positive patients and their infants. In 2001, even before FDA approval, the CDC recommended rapid testing where available for all women with undocumented HIV status during labor and delivery; ACOG echoed these recommendations in 2004. ACOG additionally recommended repeat testing in the third trimester for facilities located in areas where HIV prevalence among women of child-bearing age was 5 per 1,000 (0.5%) or greater (ACOG, 2004).

Evidence Behind Perinatal HIV Testing Policies

The evolution of HIV testing policies over the last 10 years has changed the way prenatal care providers approach HIV in a pregnant population. Initially, there was no therapy that could offer an effective treatment to either promote health in a pregnant woman or prevent infection of her fetus; therefore, there was little utility in prenatal and/or neonatal screening other than for surveillance purposes. However, with the development of zidovudine treatment and demonstration of its effectiveness in controlling HIV disease progression and preventing vertical transmission, the importance of timely diagnosis of HIV (prior to or during pregnancy) became more relevant. Perinatal HIV policies have been aimed at providing timely diagnosis and appropriate treatment of HIV for infected pregnant women. To document the evidence justifying these policies and evaluate their ability to accomplish this goal, we conducted a systematic literature review in February 2008. The literature search terms utilized for this review are detailed under each subheading. All searches utilized MEDLINE, CINAHL (Cumulative Index to Nursing and Allied Health Literature), Popline, WHO Reproductive Health Library, Web of Science and Cochrane Database. The searches were limited to human studies, the English language, and the years 1990 through 2007. Aside from the research addressing zidovudine (ZDV), only studies conducted in North America were included and analyzed. The authors additionally used bibliographies from reviewed studies to further identify pertinent literature and selected additional content under the guidance of expert opinion.

Use of ZDV

Tables 7.2 and 7.3 summarize the studies that support the use of ZDV in pregnancy. A literature search was conducted utilizing the search words; *AZT* and *pregnancy*, *AZT* and *perinatal*, *zidovudine* and *pregnancy*, *zidovudine* and *perinatal*, and *vertical disease transmission*. Only studies that compared zidovudine to placebo or no therapy were included in this analysis. We chose to limit our analysis to ZDV as it is the most studied antiretroviral in pregnancy and has been the focus of national and state policy. Exploring the use of other HIV specific medications would be a primarily clinical comparison and outside the scope of this review.

Table 7.2 Major outcomes associated with studies of zidovudine (ZDV) intervention

Author (year)	Study design	Study type	Description of intervention what, how and where	Populations studied (ages included, race and ethnicity) and sample size	Address disparities (yes/no)	Key findings related to intervention effectiveness (OR with CI or p values reflecting the intervention-outcome relationship)	Caveats/biases	Findings support the intervention? (yes/no); For which populations?
Connor et al. (1994)	Double-blinded, placebo-controlled, randomized study	Journal article	Antepartum, intrapartum, and neonatal ZDV for reducing vertical HIV transmission in the United States and France	477 HIV positive gravid women between 14 and 34 weeks gestation: – 51% Black – 29% Hispanic – 19% White	No	Treatment resulted in 67.5% relative risk reduction of vertical transmission p = 0.0006	The study was limited to other wise healthy women with CD4+ counts ≥200	Yes in gravid HIV positive women with CD4+ count ≥200
Boyer (1994)	Prospective cohort	Journal article	Antepartum or intrapartum ZDV therapy to reduce vertical transmission of HIV in Los Angeles, CA (university center and two hospital affiliates)	63 HIV-seropositive gravid women: – 51% Black – 17% Hispanic – 22% White	No	ZDV treatment in the antepartum and/or intrapartum period reduced the risk of vertical transmission p = 0.01	1. Compliance was not monitored 2. Medical treatment was not standardized 3. Small sample size 4. Study was not designed to specifically evaluate ZDV therapy 5. Study was not randomized	Yes in gravid HIV positive women
Matheson (1995)	Retrospective cohort	Journal article	Antepartum ZDV to reduce vertical transmission among women with various CD4+ counts in New York City	321 HIV positive gravid women: – 54% Black – 39% Hispanic – 7% White or other	No	Benefits of ZDV are independent of CD4 count	1. Compliance was not monitored 2. Medical treatment was not standardized	Yes in gravid HIV positive women with or without CD4+ lymphocyte depression
Gorsky (1996)	Decision Analysis	Journal article	Maternal counseling and voluntary testing for HIV followed by treatment of seropositive patients with ZDV	n/a	No	Voluntary testing for HIV during pregnancy along with ZDV treatment for seropositive patients is cost effective	1. Analytical study based on assumptions 2. Costs vary over time	Yes in communities where maternal seroprevalence is equal to or greater than 1.7/1,000

(continued)

Table 7.2 (continued)

Author (year)	Study design	Study type	Description of intervention what, how and where	Populations studied (ages included, race and ethnicity) and sample size	Address disparities (yes/no)	Key findings related to intervention effectiveness (OR with CI or p values reflecting the intervention-outcome relationship)	Caveats/biases	Findings support the intervention? (yes/no); For which populations?
Aleixo (1997)	Prospective cohort	Journal article	Antenatal, intrapartum, and neonatal ZDV to reduce vertical transmission in Florida	42 HIV positive gravid women: – 77.5% Black – 5.58% Hispanic – 16.7% White – 2.4% unidentified	No	ZDV reduced the risk of transmission by 78% p = 0.017	1. Not randomized 2. Small sample size	Yes in a predominately Black nonbreastfeeding HIV positive pregnant population
Wade et al. (1998)	Retrospective cohort	Journal article	ZDV prophylaxis for prevention of vertical transmission in New York	939 HIV exposed infants: – 58.3% Black – 27.1% Hispanic – 10.5% White – 4.2% other	No	Rates of perinatal transmission were dependant on when ZDV prophylaxis was begun: *Antepartum*: 6.1% (4.1–8.9%) *Intrapartum*: 10% (3.3–21.8%) *48 h of life*: 9.3% (4.14–17.5%) *Day 3 or greater* 18.4% (7.7–34.4%) *No ZDV*: 26.6% (21.1–32.7%)	1. Observational study 2. Relatively small numbers 3. Exact date of initiation of zidovudine therapy could not be determined 4. Patient compliance could not be determined	Yes
Dabis et al. (1999)	Double-blinded, placebo-controlled, randomized study	Journal article	ZDV administered weeks 36–38 for prevention of vertical transmission in Africa	431 gravid African women, ages 18 and older	No	ZDV was superior to placebo. Transmission was reduced by 42% (5–65%) p = 0.026	Compliance could not be confirmed	Yes
Shaffer et al. (1999)	Double-blinded, placebo-controlled, randomized study	Journal article	Antepartum ZDV starting at 36 weeks (short course) in combination with intrapartum ZDV to prevent vertical HIV transmission in Bangkok, Thailand	397 gravid HIV positive women	No	ZDV reduced the risk of vertical transmission by 50% (15.4–70.6%) p = 0.006	Compliance could not be confirmed	Yes in gravid HIV positive women who do not breastfeed

Wiktor (1999)	Double-blinded, placebo-controlled, randomized study	Journal article	Antepartum and intrapartum ZDV therapy	280 HIV positive, breastfeeding, African women who were 36 weeks pregnant	No	At 4 weeks of age 21% of infants in the placebo group were found to be infected with HIV-1 as compared to 12% in the treatment arm. At 3 months there was not statistical difference in the 2 arms p = 0.05	1. At 3 months of age the difference in HIV-1 infected infants was not statistically significant between the treatment and the control groups 2. Only 17% of eligible patients enrolled in the study 3. Adherence to the intrapartum component of the study was poor	No
Rovira (2001)	Retrospective cohort	Journal article	Antepartum and/or intrapartum ZDV therapy to prevent vertical transmission of HIV in Spain	52 low income, low education, and poorly compliant gravid women	No	Transmission rates were 6.6% with the use of zidovudine. Prior to the initiation of ZDV transmission rate were 14.3%	1. Not randomized 2. Not standardized	Yes in a low income, low education, and poorly compliant HIV positive pregnant population
Songok (2003)	Prospective cohort	Journal article	Short course ZDV therapy to prevent vertical transmission in Kenya	194 gravid women	No	ZDV reduced the risk of transmission (59.2% vs. 29.6% disease free survival) p = 0.0005	1. Not randomized 2. Differences in gestation period and maternal level of education 3. Differences in the number of days AZT was administered	Yes

Table 7.3 Quality rating of studies associated with zidovudine

	Author (year)	Reporting	External validity	Internal validity – bias	Internal validity – confounding	Power	Total quality score ≤14 = poor 15–19 = fair ≥20 = good	Suitability of study to assess effectiveness: greatest, moderate, least
Health status out-come #1	Connor et al. (1994)	12	4	7	6	2	31	Greatest
	Boyer (1994)	10	2	6	3	0	21	Moderate
	Matheson (1995)	10	4	6	4	0	24	Moderate
	Aleixo (1997)	10	1	5	2	0	18	Least
	Wade et al. (1998)	10	4	6	4	0	24	Moderate
	Dabis et al. (1999)	13	3	7	6	1	30	Moderate
	Shaffer et al. (1999)	12	4	2	6	1	25	Moderate
	Wiktor et al. (1999)	11	1	5	4	1	22	Moderate
	Rovira (2001)	5	4	5	2	0	16	Least
	Songok (2003)	13	4	6	4	0	27	Greatest

The main study that led the CDC to recommend the use of zidovudine in pregnancy was the Pediatric AIDS Clinical Trial Group (PACTG) 076 trial. This was a randomized, double-blind, placebo-controlled trial in which zidovudine therapy was initiated between 14 and 34 weeks of pregnancy, administered in labor, and dispensed to neonates for the first 6 weeks of life. The trial was stopped after the first interim analysis established the ability of ZDV to effectively reduce the risk of maternal-infant HIV transmission by 67% (Connor et al., 1994).

In a retrospective review, Wade and colleagues found that transmission rates were approximately 10% when prophylaxis was initiated in the intrapartum period, and within the first 48 h of life for the newborn. When newborn treatment did not start until the third day of life, transmission increased to 18.4% and was 26.6% in the absence of therapy (Wade et al., 1998).

PACTG 076 was not the only randomized control trial that found that zidovudine could effectively reduce perinatal HIV transmission. In 1998, the CDC conducted a randomized placebo-controlled trial in Thailand. This study evaluated the use of zidovudine starting at 36 weeks of pregnancy and continuing until delivery. This abbreviated regimen reduced the risk of transmission by 51% (Shaffer et al., 1999). Wiktor et al. found the same regimen to be effective in Abidjan, Côte d'Ivoire with a 37% reduction in transmission at 3 months of age (Wiktor et al., 1999). In 1999, Dabis and colleagues published results of a randomized placebo-controlled trial carried out in two African countries where zidovudine was initiated between 36 and 38 weeks of pregnancy (Dabis et al., 1999). ZDV reduced the rate of vertical transmission by at least 30% when the infants were followed for 15 months.

Additional studies support the conclusion that zidovudine significantly reduces the risk of vertical transmission (Aleixo, Goodenow, & Sleasman, 1997; Boyer et al., 1994; Rovira et al., 2001; Songok et al., 2003). Matheson et al. found that the benefits of ZDV therapy are independent of CD4 count (Matheson et al., 1995). In addition, Gorsky et al. found that a program of voluntary testing and zidovudine treatment in pregnancy is cost-effective (Gorsky et al., 1996).

Evolution to Opt-Out Testing for HIV

Tables 7.4 and 7.5 summarize the studies that support the opt-out testing strategy and demonstrate its public health impact. A search was conducted utilizing the terms *opt* and *HIV*, *opt* and *AIDS*, and *opt-out*.

In 1991, Barbacci, Repke, and Chaisson reported that utilizing a risk-based strategy for HIV testing in pregnancy identified only 57% of pregnant women with HIV. This report highlighted the failure of risk-based testing to identify a critical number of HIV positive patients and predicted the shortcomings that were inherent in the 1995 CDC recommendations.

Following a congressional mandate, the IOM (1999) examined the extent to which states had effectively reduced perinatal HIV transmission and explored the barriers to such reduction. The results of their investigation led the IOM to conclude that the most effective strategy for reducing perinatal HIV transmission would be to increase the number of women in prenatal care who both are offered HIV testing by their providers and accept it. At the time of the report (1999) it was estimated that, at most, 75% of providers offered HIV testing to every pregnant woman. Increasing the rate at which providers offered HIV testing to 100%, and increasing the proportion of women who accepted it from 80 to 100%, would reduce the number of HIV-infected infants by 386 per year (33%). Based on these projected data, the IOM recommended universal testing with patient notification as a routine component of prenatal care (i.e. an opt-out testing strategy; IOM).

Several researchers have been able to demonstrate the public health impact of opt-out testing. Jayaraman, Preiksaitis, and Larke (2003) demonstrated that the adoption of an opt-out strategy in

Table 7.4 Major outcomes associated with studies of opt-out testing

Opt-out testing

Author (year)	Study design	Study type	Description of intervention what, how and where	Populations studied (ages included, race and ethnicity) and sample size
Stringer et al. (2001)	Prospective cohort	Journal article	Opt-out HIV testing strategy in a predominantly Black Alabama population	Intervention group = 3,415 gravid women seeking prenatal care Historical control = 3,778 women seeking prenatal care the year prior to the intervention
CDC (2002d)	Cross-sectional	Journal article	Opt-out HIV testing strategy in the in the United States and Canada	Population based data were obtained by three different methods: (1) analyzing labor and delivery charts, (2) laboratory data, and (3) patient interviews
Jayaraman et al. (2003)	Retrospective cohort	Journal article	Opt-out HIV testing strategy in Canada	Approximately 20,000 women eligible for prenatal HIV screening in the province of Alberta
Breese et al. (2004)	Retrospective cohort	Journal article	Verbal opt-out HIV testing strategy in an urban Denver public hospital	12,221 pregnancies resulting in delivery in a largely Hispanic population (82%)
Yudin et al. (2007)	Two part study with a retrospective cohort and a prospective cohort	Journal article	Opt-out testing in a single Canadian clinic	1,140 gravid women: – 11% black – 18–19% Hispanic – 39–42% Caucasian – 23% Asian – 6–8% other – Majority of women were insured

Address disparities (yes/no)	Key findings related to intervention effectiveness (OR with CI or p values reflecting the intervention-outcome relationship)	Caveats/biases	Findings support the intervention? (yes/no); For which populations?
Non-White/non-Black (predominately Hispanic) women were less likely to be tested	With an opt-out policy HIV testing increased from 75 to 88% $p < 0.001$	1. Historical controls were utilized 2. Unable to determine who declined HIV testing and why they did so	Yes in a predominantly Black population
No	States and provinces utilizing opt-out approach had higher testing rates as compared to those utilizing an opt-in approach	1. Maternal self-reported data may be subject to recall bias 2. Uncertain if prenatal-care providers were aware of a patient's HIV	Yes in a pregnant population
No	HIV testing rates increased by 28% immediately following implementation of opt-out testing	1. Not specifically designed to look at pregnant patients 2. Controls are historical	Yes in a pregnant population
Yes – Those not screened were more often African-American or Caucasian, were more likely to speak English, and were U.S. citizens	98.2% of pregnancies resulting in delivery were screened for HIV by the time they presented to labor and delivery	Unable to ascertain why screening was not accomplished (refusal vs. failure to offer screening) No comparison with alternative strategy	Yes in a predominately Hispanic population
Yes – Asians were significantly LESS likely to accept testing Hispanics were significantly MORE likely to accept testing Those fluent in English were significantly MORE likely to accept testing	Clinic rates were significantly higher than provincial rates $p < 0.001$	Results are representative of only a single clinic	Yes in a pregnant largely insured population

Table 7.5 Quality rating of studies associated with opt-out testing

	Author (year)	Reporting	External validity	Internal validity – bias	Internal validity – confounding	Power	Total quality score ≤14 = poor 15–19 = fair ≥20 = good	Suitability of study to assess effectiveness: greatest, moderate, least
Health status out-come #1	Stringer et al. (2001)	11	3	4	3	0	21	Moderate
	CDC (2002d)	8	4	4	4	0	20	Moderate
	Jayarman et al. (2003)	7	4	7	2	0	20	Moderate
	Breese et al. (2004)	10	4	3	3	0	19	Moderate
	Yudin et al. (2007)	9	4	5	2	0	20	Moderate

Alberta, Canada increased the number of prenatal HIV tests by 28% province-wide. A Canadian study by Yudin, Moravac, and Shah (2007) discovered that the testing rate of a cohort of patients offered opt-out testing was significantly higher than the provincial average in an area where opt-in screening remained the standard of care. Stringer and colleagues evaluated eight clinics in Alabama and discovered that the rate of HIV testing increased from 75 to 88% with the implementation of the opt-out strategy (Stringer, Cliver, Goldenberg, & Goepfert, 2001).

In 2002, the CDC also reviewed the efficacy of opt-in vs. opt-out strategy by evaluating the HIV testing rates in U.S. states and Canadian provinces. States using the opt-in method had testing rates that varied between 25 and 83%. In a study of two U.S. states that were using the opt-out method, Tennessee was shown to have an 85% testing rate; Arkansas increased testing rates from 57 to 71% when an opt-out strategy became mandatory. In Canada, the opt-out strategy yielded a 94–98% testing rate whereas those provinces utilizing the opt-in approach only had a 54–83% testing rate (CDC, 2002d).

Initiation of Rapid Testing

Tables 7.6 and 7.7 delineate the studies that supported the initiation of rapid HIV testing during labor. A search was conducted utilizing the terms "rapid (HIV or AIDS), and pregnancy."

Initially, four studies which used decision-analysis methods were published that supported rapid HIV testing in a pregnant population, concluding that such a strategy is accurate and cost-effective (Doyle, Levison, & Gardner, 2005; Grobman & Garcia, 1999; Mrus & Tsevat, 2004; Stringer & Rouse, 1999).

Following these studies, the Mother-Infant Rapid Intervention At Delivery (MIRIAD) Study Group conducted a prospective study that included 17 facilities in 6 cities (Bulterys et al., 2004). The goal was to determine the feasibility and acceptance of rapid HIV testing among women in labor. The group determined that rapid testing was feasible, acceptable, and accurate.

Table 7.6 Major outcomes associated with studies of rapid testing

Rapid testing								
Author (year)	Study design	Study type	Description of intervention what, how and where	Populations studied (ages included, race and ethnicity) and sample size	Address disparities (yes/no)	Key findings related to intervention effectiveness (OR with CI or p values reflecting the intervention-outcome relationship)	Caveats/biases	Findings support the intervention? (yes/no); For which populations?
Stringer and Rouse (1999)	Decision analysis	Journal article	1. No rapid HIV testing and no ZVD in labor 2. Rapid testing then ZVD if seropositive 3. Universal ZVD for all pregnant patients with unknown serostatus	Women presenting to labor and delivery without prenatal care	No	Rapid testing with ZDV therapy for HIV positive patients is the most cost-effective approach	1. Analytical study based on assumptions 2. Costs vary over time	Yes for previously unregistered women residing in communities where the prevalence of HIV is at least 1%
Grobman and Garcia (1999)	Decision Analysis	Journal article	Rapid HIV testing during labor for women with inadequate prenatal care	Women with undetermined HIV serostatus presenting to labor and delivery	No	Rapid testing in labor offers significant economic and health benefits	1. Analytical study based on assumptions 2. Costs vary over time	Yes for women presenting with undocumented HIV serostatus
Cohen et al. (2003)	Prospective cohort	Journal article	Point-of-care HIV testing in four Chicago hospitals	514 women presenting to labor and delivery without documentation of HIV status	No	Point-of-care HIV testing is faster than laboratory testing ($p = 0.0001$) and feasible on labor and delivery units	Adequate training and quality assurance procedures are required	Yes in women with undocumented HIV serostatus

(continued)

Table 7.6 (continued)

Rapid testing

Author (year)	Study design	Study type	Description of intervention what, how and where	Populations studied (ages included, race and ethnicity) and sample size	Address disparities (yes/no)	Key findings related to intervention effectiveness (OR with CI or p values reflecting the intervention-outcome relationship)	Caveats/biases	Findings support the intervention? (yes/no); For which populations?
Mrus and Tsevat (2004)	Decision Analysis	Journal article	1. No HIV testing 2. Rapid HIV testing and treatment for women presenting without prenatal care	Women presenting to labor and delivery without prenatal care	No	Rapid testing and treatment of patients who are HIV positive reduces transmission and is cost effective	1. Analytical study based on assumptions 2. Costs vary over time	Yes for previously unregistered women residing in communities where the HIV prevalence rate for unregistered patients is >2/1,000
Bulterys et al. (2004)	Prospective cohort	Journal article	Rapid HIV testing on labor and delivery in 16 US hospitals in 6 states	5,744 women presenting to labor and delivery with undocumented HIV serostatus	Hispanics were more likely to accept testing	Rapid testing is faster than EIA ($p = 0.001$) and feasible on labor and delivery units	Favorable results may be difficult to reproduce outside of a research setting	Yes in women with undocumented HIV serostatus
Doyle et al. (2005)	Decision Analysis	Journal article	Rapid testing vs. ELISA in labor and delivery settings	Low risk Mexican American population in labor with undocumented HIV serostatus	No	Rapid testing is cost-effective	1. Analytical study based on assumptions 2. Costs vary over time	Yes in low risk Mexican-American women with undocumented serostatus residing in a population where the incidence of HIV is 0.05%

Table 7.7 Quality rating of studies associated with rapid testing

	Author (year)	Reporting	External validity	Internal validity – bias	Internal validity – confounding	Power	Total quality score ≤14 = poor 15–19 = fair ≥20 = good	Suitability of study to assess effectiveness: greatest, moderate, least
Health status out-come #1	Cohen et al. (2003)	8	2	5	1	0	16	Moderate
	Bulterys et al. (2004)	11	3	7	3	0	24	Greatest

In 2003 Cohen and colleagues reviewed the data from Chicago area hospitals participating in the MIRIAD study and found that point-of-care HIV testing was effective and efficient (Cohen et al., 2003). Cohen et al.'s study found that such HIV testing was three times faster than conventional testing and provided an opportunity for prophylactic antiretroviral administration.

Third Trimester Repeat Screening

Tables 7.8 and 7.9 provide information on the studies that support prenatal third trimester HIV screening and its public health implications. The search terms utilized were: *third trimester* (*HIV or AIDS*), *test*; *third trimester* (*HIV or AIDS*), *screen*; *repeat* (*HIV or AIDS*), *screening, pregnancy*; and *repeat* (*HIV or AIDS*), *screening*, *perinatal*.

In 2003, Sansom and colleagues published a decision analysis demonstrating that a second HIV test performed in the third trimester would result in a $5.2 million savings in communities where the estimated HIV incidence was 6.2 per 1,000 person years (Sansom, Jamieson, Farnham, Bulterys, & Fowler, 2003). Nesheim and colleagues reviewed the MIRIAD data and discovered that 11% of those found to be positive at the time of rapid testing had evidence of seroconversion during pregnancy (Nesheim et al., 2007). These data suggest that for some populations an initial HIV screening test early in pregnancy may not be adequate.

Summary of Evidence Related to HIV Testing During Pregnancy and Delivery

This review of the literature on HIV testing during pregnancy and delivery provides strong evidence that the ideal method for minimizing vertical transmission of HIV is to provide zidovudine therapy to all HIV-infected pregnant women and their newborns. Furthermore, the evidence suggests that the best way to achieve this goal is to offer universal opt-out testing to women early in pregnancy as a part of routine prenatal care, and when appropriate, to repeat testing in the third trimester. Evidence also suggests that women who reach labor and delivery with an undocumented HIV status should receive rapid testing at that time so that intrapartum and neonatal zidovudine treatments can be offered.

Table 7.8 Major outcomes associated with studies of third trimester screening

Third trimester screening								
Author (year)	Study design	Study type	Description of intervention	Populations studied and sample size	Address disparities (yes/no)	Key findings related to intervention effectiveness (OR with CI or p values reflecting the intervention-outcome relationship)	Caveats/biases	Findings Support the intervention? (yes/no); For which populations?
Sansom et al. (2003)	Decision-Analysis	Journal article	Third trimester screening	Gravid women	No	Repeat HIV testing in the third trimester was cost-effective in populations were the incidence of HIV was 6.2/1,000 person years	1. Analytical study based on assumptions 2. Costs change over time	Yes
Nesheim et al. (2007)	Observational	Journal article	Third trimester screening	8,165 gravid women in their third trimester	No	54 women were diagnosed with HIV utilizing rapid testing in labor. Eleven percent of these infections occurred during pregnancy.	1. Secondary analysis 2. Small sample size	Yes

Table 7.9 Quality rating of studies associated with third trimester screening

	Author (year)	Reporting	External validity	Internal validity – bias	Internal validity – confounding	Power	Total quality score ≤14 = poor 15–19 = fair ≥20 = good	Suitability of study to assess effectiveness: greatest, moderate, least
Health status out-come #1	Nesheim et al. (2007)	10	3	7	3	0	23	Moderate

State Implementation of Policies to Address Perinatal HIV Transmission

Although the CDC now recommends universal opt-out screening with repeat third trimester screening in high prevalence communities as well as rapid testing in labor and delivery for all women who lack evidence of HIV testing in pregnancy, there are still a significant number of states that have not yet formulated any state-wide policies regarding perinatal HIV testing. Among those that have developed policies, there is substantial variability in how they are formulated and implemented. To demonstrate this diversity, we provide examples from five states of policies and strategies aimed at minimizing and potentially eliminating perinatal HIV transmission. (See text boxes.)

Use of Legislation

California

In 2006 Kropp and colleagues published a study aimed at recognizing missed opportunities to prevent perinatal HIV transmission in the Hispanic population in California. They found that Hispanic women were significantly less likely to initiate prenatal care in a timely manner or receive offer for an HIV test. According to the California Department of Health, in 2001 the greatest number of pediatric HIV infections occurred in Hispanic children (37.5%). In a state where the Hispanic population is growing, the disparity that exists between Hispanic and non-Hispanic populations is of particular concern (Kropp, Sarnquist, Montgomery, Ruiz, & Maldonado, 2006).

A California State law was passed in 2003 that mandated an "opt-out" testing strategy and rapid testing for women presenting to labor and delivery with unknown serostatus. Following the passage of this legislation the percentage of patients offered prenatal HIV testing significantly increased in California. In this opt-out setting, failure to receive prenatal HIV testing and appropriate treatment was found to be most associated with not receiving prenatal care, demonstrating the connection between adequate access to prenatal care and access to or utilization of perinatal HIV testing (Sarnquist, Cunningham, Sullivan, & Maldonado, 2007).

Standard of Care

New Jersey

The New Jersey Department of Health and Senior Services (NJDHSS) issued an objective to decrease the rate of HIV infection in pregnant women to 0.10% by 2010 (NJDHSS, 2001). As part of this initiative, New Jersey state law mandates prenatal counseling and voluntary testing of all pregnant women (New Jersey State Nurses Association, 2004). Additionally, the NJDHSS made counseling and the offer of a rapid or expedited HIV test the standard of care for all women presenting to labor and delivery with no known prenatal care or unknown/undocumented HIV status (NJDHSS, 2002).

The rate of pediatric HIV infection dropped dramatically (75%) between 1991 and 2000. This decline was a result of a decrease in both the prevalence of HIV in women of childbearing age and the introduction of zidovudine to reduce perinatal transmission. Improved zidovudine administration was confirmed through a blinded population-based newborn serum screening program that demonstrated that the proportion of HIV-exposed neonates (i.e. screened positive for maternal HIV antibodies) appropriately treated with zidovudine (screened positive for zidovudine) had increased to 91.7% in 2004, the highest proportion recorded ever in the state (NJDHSS, 2005).

Despite the advances made in reducing perinatal HIV transmission, African-American women of childbearing age remained disproportionately affected, with HIV prevalence rates four times that of Hispanic women and 30 times that of White women. Because perinatal exposure is the most common source of HIV exposure for children, this disparity translates into a significant disparity in the number of African-American children infected (NJDHSS, 2005).

Legislation and Targeted Interventions

Florida

Florida law mandates HIV counseling of pregnant women and the offer of an HIV test during the initial prenatal visit. If a patient refuses HIV testing, the refusal must be documented in the prenatal record. Because mandated prenatal testing is only effective to the extent that women receive prenatal care, Florida has developed programs to improve access to prenatal care for high-risk women.

In 1999, Florida adopted a program known as the Targeted Outreach for Pregnant Women Act (TOPWA) to identify women deemed to be at risk of giving birth to an HIV-infected or substance-exposed infant, and to link those women with prenatal care (Clark et al., 2006). The program works through community-based outreach organizations at venues where HIV-positive pregnant women are likely to be found, and it has been expanded into jails because of the high prevalence of substance abuse among this population. While 6% of the female prison population in Florida is reported to be pregnant, limited prenatal care has historically been provided in prison. From 2002 to 2004, TOPWA jail programs were able to identify 515 gravid women; 38% had not received prenatal care, and 3.7% were found to be HIV positive through testing and disclosure. Sixteen women delivered by April of 2005 and only one infant (6%) was reported to be infected (Clark et al.).

Legislation and Statewide Programs

New York

New York State has had the largest number of HIV infected women of childbearing age in the country, putting both women and children in the state at risk (NYSDOH, 2004). The state has responded to this fact with the implementation of a comprehensive program aimed at reducing both HIV infection in women of childbearing age and vertical transmission.

In 1995, the New York State Department of Health (NYSDOH) launched several initiatives to reduce the incidence of perinatal infection. Funding was provided for prenatal counseling and testing as well as community and medical education. All providers received training and educational materials related to consensus guidelines and best practices for physicians. Access to antiretroviral therapy was secured by Medicaid and the AIDS Drug Assistance Program.

In 1996, NYSDOH-regulated facilities were required to provide HIV counseling and consented testing as mandated by New York State (NYS) statute and subsequent regulations. Counseling and testing were deemed standard of care for all other prenatal sites. The Newborn Screening program was initiated in 1997 and required the testing of all newborns (Wade et al., 2004). Expedited testing was initiated for all women who presented to labor and delivery with unknown serostatus. Newborn testing was implemented for all infants whose mothers were not tested prior to delivery (Wade et al.).

Wade et al. (2004) examined the rate of perinatal transmission in New York between 1997 and 2000 and found a decline from 11 to 2.7% over the 3 years. Interestingly, this study also found that the rate of transmission was highest among women over the age 40 years, those who received no prenatal care, and those with low birth weight infants, vaginal delivery, and who were White. The elevated transmission rate in White women was not a readily explainable phenomenon, as those women are usually more likely to receive prenatal care, antiretroviral therapy, and less likely to have a low birth weight infant (Wade et al.).

Pulver and colleagues found that the prevalence of HIV in childbearing women in New York City (NYC) declined from 1.22% in 1988–1989 to 0.62% in 1999–2000 (Pulver, Glebatis, Wade, Birkhead, & Smith, 2004). The decline was greatest for White women followed by Hispanic and African-American women. A less dramatic decline (24%) was noted among women living outside of NYC (Pulver et al.).

Illinois

In August 2003, Illinois passed the HIV Perinatal HIV Prevention Act (45) mandating prenatal care providers to counsel and offer HIV tests to all women as early in pregnancy as possible, and that maternal HIV status be documented in the maternal labor and delivery and newborn pediatric chart. Additionally, the Act specified that rapid testing must be offered to all laboring women (opt-in) and newborns (opt-out) with undocumented HIV status (Illinois, 2003).

By the spring of 2004, birthing hospitals in Illinois had not yet implemented the Perinatal HIV Prevention Act and hospitals in Illinois were not regularly using rapid HIV testing in the perinatal setting. A collaborative composed of public health workers and HIV clinicians called Perinatal Rapid Testing Implementation Initiative (PRTII) was funded by the Illinois Department of Public Health (IDPH) to assist in implementation of the 2003 law. Through a survey of all birthing hospitals in the state PRTII found that only 72% of hospitals routinely documented prenatal HIV results in the labor and delivery chart, 65% documented maternal

(continued)

HIV status in the newborn chart, 39% of hospitals routinely ordered HIV tests on labor and delivery for undocumented maternal HIV status in labor and only 61% of hospitals reported having intravenous Zidovudine (AZT) available (Bryant et al., 2007). These results showed that passage of legislation alone was not an effective strategy for implementation of statewide rapid testing in birthing hospitals and that a statewide implementation initiative with hospital-specific resources and education was necessary to overcome barriers to implementing this new law (Bryant et al.). Consequently, regional trainings were held around the state for all labor and delivery nurse managers to orient them to the steps for rapid HIV testing implementation.

To provide support to the birthing hospitals now providing rapid HIV testing, a 24/7 Perinatal HIV Hotline was established to help link hard-to-reach women to care and to provide real-time medical consultation on HIV related obstetric and pediatric issues. By September 2005 all 133 birthing hospitals in Illinois had implemented rapid HIV testing on labor and delivery (Borders, 2005).

In 2006, Illinois passed amendments to the Perinatal HIV Prevention Law requiring: (a) rapid HIV testing of all newborns whose maternal HIV status was undocumented, (b) all hospitals to call the Illinois 24/7 Perinatal HIV Hotline to report women and newborns identified as preliminarily HIV-positive within 24 h so that they could be linked to case management, (c) all hospitals to report monthly and annual data on rapid HIV testing during labor and delivery to the Illinois Department of Public Health, and (d) the Illinois Department of Public Health to fund the Illinois 24/7 Perinatal HIV Hotline and to fund the Perinatal HIV case management program.

As of September 2007, 94% of women in Illinois had a documented HIV test on arrival to labor and delivery, 98% of women who had an undocumented HIV test on arrival to L&D were counseled and accepted rapid HIV testing, and 99.9% of women/baby pairs had a documented HIV status prior to discharge home after delivery (Statton, Ayala, & Olszewski, 2007).

The approaches implemented in New York and Illinois demonstrate the importance and success of active implementation of state-wide systems to support and enforce laws addressing HIV testing and administration of ZVD. Although legislative mandates may seem to be the most powerful tools for ensuring uniform implementation of effective screening policies, mandates alone are seldom sufficient. Also necessary are statewide implementation strategies, surveillance efforts to monitor adherence to the mandates, as well as adequate funding for counseling, provider training, community outreach and coverage for the cost of antiretroviral medications (ARVs). Without active interventions, laws may merely remain on paper and not establish effective state-wide change.

Racial and Ethnic Disparities in STI and HIV Screening/ Testing During Pregnancy

The prevalence and burden of STIs and HIV for non-pregnant and pregnant women is disproportionately borne by women of color. The reasons for these disparities are many and varied. A 2002 article by Williams pointed out that, “although socioeconomic status is a central determinant of racial and ethnic disparities in health, other factors such as geographic location, migration and acculturation, and racism, also plays a role” (Williams, 2002). A final common pathway for many of these influences is inequity in access to and quality of health care. Women of color are in general at greater risk for insufficient access to primary care, and this is reflected in a reduced likelihood

that they will be tested for HIV and other STIs, despite the known fact that they suffer higher infection rates than White women (Anderson & Sansom, 2007).

Once pregnant, women of color are less likely to have access to early and adequate prenatal care (Martin et al., 2009). In part because of this disparity in access, pregnant women of color are less likely to be tested for HIV and other STIs (regardless of the screening approach utilized), less likely to have an infection recognized, and less likely to receive appropriate therapy (Hamilton, Martin, Ventura, Sutton,& Menacker, 2005). A national study showed that pregnant women without insurance coverage or without a medical care home are at increased risk of not being tested for HIV (Anderson & Sansom, 2007). In a retrospective study of over 1,800 births in California, women with prenatal care were seven times more likely to have been screened in pregnancy for a variety of infectious diseases (including syphilis, HIV, and hepatitis B) than women without prenatal care (Sheikh et al., 2009). In the same study, incarcerated women (many of whom are women of color) had reduced odds of being screened for HIV. This is particularly problematic, as rates of HIV/AIDS in U.S. prisons are 2.7–4.8 times higher than the general population (CDC, 2009b; Sheikh et al.).

Racial and ethnic disparities in screening rates for HIV and other STIs in pregnancy are a likely contributor to disparities in adverse pregnancy outcomes associated with these diseases. Disparities in screening for STIs may contribute to the elevated rates of prematurity, low birth weight and infant mortality found in populations of color, and disparities with respect to screening/testing for HIV during pregnancy are a likely cause of elevated perinatal HIV transmission among women of color. Though pregnancy provides an excellent opportunity for intervention as well as, screening women prior to pregnancy is the optimal method to improve outcomes. Attending to the health status of women prior to pregnancy is clearly important and there is increasing attention being paid to preconceptional care as an effective strategy for improving the health of both women and infants. Consequently, increased access for women of color to STI/HIV testing prior to as well as during pregnancy may have significant potential for reducing racial/ethnic disparities in pregnancy outcomes.

Conclusion

Effective detection and treatment of sexually transmitted infections, including HIV, during pregnancy and the entire perinatal period is clearly a powerful tool for improving the health status of women and infants; however, effective STI screening is also necessary prior to pregnancy in order to most effectively improve health outcomes. Regardless of pregnancy status, all women should have access to routine health services, including education, screening and treatment for sexually transmitted infections in order to maximally reduce preventable adverse pregnancy outcomes. However, while increased screening for STIs prior to and during pregnancy is likely to reduce maternal and infant morbidity and mortality for all women, it must be emphasized that reducing racial and ethnic disparities in STI screening, treatment and outcomes will require addressing barriers and inequities that contribute to population health differentials.

References

Aleixo, L. F., Goodenow, M. M., & Sleasman, J. W. (1997). Zidovudine administered to women infected with human immunodeficiency virus type 1 and to their neonates reduces pediatric infection independent of an effect on levels of maternal virus. *Journal of Pediatrics*, *130*(6), 906–914.

American College of Obstetricians and Gynecologists. (2004). Prenatal and perinatal human immunodeficiency virus testing: Expanded recommendations. ACOG Committee Opinion No. 304. *Obstetrics and Gynecology*, *104*, 1119–1124.

American College of Obstetricians and Gynecologists. (2007). Viral hepatitis in pregnancy. ACOG Practice Bulletin No. 86. *Obstetrics and Gynecology*, *110*, 941–955.

American College of Obstetricians and Gynecologists/American Academy of Pediatrics. (1999). *Joint Statement of ACOG/AAP on human immunodeficiency virus screening [Reaffirmed by AAP in 2005 and ACOG in 2006]*. Washington, DC: Author.

American College of Obstetricians and Gynecologists/American Academy of Pediatrics. (2007). *Guidelines for perinatal care* (6th ed.). Washington, DC: Author.

Anderson, J. E., & Sansom, S. (2007). HIV testing in a national sample of pregnant US women: Who is not getting tested? *AIDS Care*, *19*(3), 375–380.

Arevalo, J. A., & Washington, A. E. (1988). Cost-effectiveness of prenatal screening and immunization for hepatitis B virus. *JAMA: The Journal of the American Medical Association*, *259*(3), 365–369.

Barbacci, M., Repke, J. T., & Chaisson, R. E. (1991). Routine prenatal screening for HIV infection. *Lancet*, *337*(8743), 709–711.

Bath, S. K., Singleton, J. A., Strikas, R. A., Stevenson, J. M., McDonald, L. L., & Williams, W. W. (2000). Performance of U.S. hospitals on recommended screening and immunization practices for pregnant and postpartum women. *American Journal of Infection Control*, *28*, 327–332.

Beem, M. O., & Saxon, E. M. (1977). Respiratory-tract colonization and a distinctive pneumonia syndrome in infants infected with Chlamydia trachomatis. *New England Journal of Medicine*, *296*(6), 306–310.

Berenson, A. B., Hammill, H. A., Martens, M. G., & Faro, S. (1990). Bacteriologic findings of post-cesarean endometritis in adolescents. *Obstetrics and Gynecology*, *75*(4), 627–629.

Blas, M. M., Canchihuaman, F. A., Alva, I. E., & Hawes, S. E. (2007). Pregnancy outcomes in women infected with Chlamydia trachomatis: A population-based cohort study in Washington State. *Sexually Transmitted Infections*, *83*(4), 314–318.

Borders, A. (2005). Oral presentation: Illinois perinatal HIV rapid testing implementation Model. In *National HIV Prevention Conference, Atlanta*, June 2005.

Boyer, P. J., Dillon, M., Navaie, M., Deveikis, A., Keller, M., O'Rourke, S., et al. (1994). Factors predictive of maternal-fetal transmission of HIV-1. Preliminary analysis of zidovudine given during pregnancy and/or delivery. *JAMA: The Journal of the American Medical Association*, *271*(24), 1925–1930.

Brandt, A. M. (1987). *No magic bullet: A social history of venereal disease in the United States since 1880*. New York: Oxford University Press.

Branson, B. M., Handsfield, H. H., Lampe, M. A., Janssen, R. S., Taylor, A. W., Lyss, S. B., et al. (2006). Revised recommendations for HIV testing of adults, adolescents, and pregnant women in health-care settings. *Morbidity and Mortality Weekly Report. Recommendations and Reports*, *55*(RR-14), 1–17.

Breese, P., Burman, W., Shlay, J., & Guinn, D. (2004). The effectiveness of a verbal opt-out system for human immunodeficiency virus screening during pregnancy. *Obstetrics and Gynecology*, *104*(1), 134–137.

Bryant Borders, A. E., Eary, R. L., Olszewski, Y., Statton, A., Handler, A., Cohen, M. H., et al. (2007). Ready or not – intrapartum prevention of perinatal HIV transmission in Illinois. *Maternal Child Health Journal*, *11*(5), 485–493.

Bulterys, M., Jamieson, D. J., O'Sullivan, M. J., Cohen, M. H., Maupin, R., Nesheim, S., et al. (2004). Rapid HIV-1 testing during labor: A multicenter study. *JAMA: The Journal of the American Medical Association*, *292*(2), 219–223.

Centers for Disease Control and Prevention. (1994). Recommendations of the U.S. Public Health Service Task Force on the use of zidovudine to reduce perinatal transmission of human immunodeficiency virus. *Morbidity and Mortality Weekly Report. Recommendations and Reports*, *43*(RR-11), 1–20.

Centers for Disease Control and Prevention. (1995). U.S. Public Health Service recommendations for human immunodeficiency virus counseling and voluntary testing for pregnant women. *Morbidity and Mortality Weekly Report. Recommendations and Reports*, *44*(RR-7), 1–15.

Centers for Disease Control and Prevention. (1998). 1998 Guidelines for treatment of sexually transmitted diseases. *Morbidity and Mortality Weekly Report. Recommendations and Reports*, *47*(RR-1), 1–111.

Centers for Disease Control and Prevention. (2001). Revised recommendations for HIV screening of pregnant women. *Morbidity and Mortality Weekly Report. Recommendations and Reports*, *50*(RR-19), 63–85.

Centers for Disease Control and Prevention. (2002a). Sexually transmitted diseases treatment guidelines 2002. *Morbidity and Mortality Weekly Report. Recommendations and Reports*, *51*(RR-6), 1–78.

Centers for Disease Control and Prevention. (2002b). Screening tests to detect Chlamydia trachomatis and Neisseria gonorrhoeae infections. *Morbidity and Mortality Weekly Report. Recommendations and Reports*, *51*(RR-15), 1–38.

Centers for Disease Control and Prevention. (2002c). Notice to readers: Approval of new rapid test for HIV antibodies. *Morbidity and Mortality Weekly Report. Recommendations and Reports*, *51*, 1051–1052.

Centers for Disease Control and Prevention. (2002d). HIV testing among pregnant women – United States and Canada, 1998–2001. *Morbidity and Mortality Weekly Report. Recommendations and Reports*, *51*, 1013–1016.

Centers for Disease Control and Prevention. (2003). Advancing HIV prevention: New strategies for a changing epidemic – United States, 2003. *Morbidity and Mortality Weekly Report. Recommendations and Reports*, *52*(15), 329–332.

Centers for Disease Control and Prevention. (2004). Congenital syphilis – United States, 2002. *Morbidity and Mortality Weekly Report. Recommendations and Reports*, *53*(31), 716–719.
Centers for Disease Control and Prevention. (2006a). Sexually transmitted diseases treatment guidelines, 2006. *Morbidity and Mortality Weekly Report. Recommendations and Reports*, *55*(RR-11), 1–94.
Centers for Disease Control and Prevention. (2006b). Surveillance for acute viral hepatitis – United States, 2006. *Morbidity and Mortality Weekly Report. Recommendations and Reports*, *57*(2), 1–24.
Centers for Disease Control and Prevention. (2007). *HIV surveillance report, 2007*. Retrieved July 2, 2009, from http://www.cdc.gov/hiv/topics/surveillance/resources/reports/2007report/default.htm.
Centers for Disease Control and Prevention. (2008). *HIV statistics and surveillance*. Retrieved February 2, 2009, from http://www.cdc.gov/hiv/topics/surveillance/united_states.htm.
Centers for Disease Control and Prevention. (2009a). *Sexually transmitted disease surveillance, 2007*. Retrieved February 2, 2009, from http://www.cdc.gov/std/stats07/toc.htm.
Centers for Disease Control and Prevention. (2009b). HIV testing implementation guidance for correctional settings. Retrieved February 9, 2009. from http://www.cdc.gov/hiv/topics/testing/resources/guidelines/correctional-settings.
Chang, E. T., & So, S. K. S. (2007). Letter to the Editor: RE: Ten largest racial and ethnic health disparities in the United States based on Healthy People 2010 objectives. *American Journal of Epidemiology*, *166*(9), 1105–1106.
Clark, J., Sansom, S., Simpson, B. J., Walker, F., Wheeler, C., Yazdani, K., et al (2006). Promising strategies for preventing perinatal HIV transmission: Model programs from three states. *Maternal Child Health Journal*, *10*(4), 367–373.
Cohen, M. H., Olszewski, Y., Branson, B., Robey, M., Love, F., Jamieson, D. J., et al. (2003). Using point-of-care testing to make rapid HIV-1 tests in labor really rapid. *AIDS*, *17*(14), 2121–2124.
Connor, E. M., Sperling, R. S., Gelber, R., Kiselev, P., Scott, G., O'Sullivan, M. J., et al. (1994). Reduction of maternal-infant transmission of human immunodeficiency virus type 1 with zidovudine treatment. *New England Journal of Medicine*, *331*(18), 1173–1180.
Coonrod, D. V., Jack, B. W., BoggessK. A., Long, R., Conry, J. A., Cox, S. N., et al. (2008). The clinical content of preconception care: Infectious diseases in preconception care. *American Journal of Obstetrics and Gynecology*, *199*(6 Suppl 2), S296–S309
Dabis, F., Dequeae-Merchadou, L., Leroy, V., Mandelbrot, L., Welffens-Ekra, C., & DITRAME ANRS 049 Study Group. (1999). 15-Month efficacy of maternal oral zidovudine to decrease vertical transmission of HIV-1 in breastfed African children. *Lancet*, *354*(9195), 2050–2051.
Doyle, N. M., Levison, J. E., & Gardner, M. O. (2005). Rapid HIV versus enzyme-linked immunosorbent assay screening in a low-risk Mexican American population presenting in labor: A cost-effectiveness analysis. *American Journal of Obstetrics and Gynecology*, *193*(3 Pt 2), 1280–1285.
Eng, T. R., & Butler, W. T. (Eds.) (1997). *The hidden epidemic: Confronting sexually transmitted diseases*. Institute of Medicine, Washington, DC: National Academy Press.
Faden, R. R., Geller, G., & Powers, M. (1991). *AIDS, women, and the next generation: Towards a morally acceptable public policy for HIV testing of pregnant women and newborns*. New York: Oxford University Press.
Fowler, C. I., Gavin, N. I., Adams, E. K., Tao, G., & Chireau, M. (2008). Racial and ethnic disparities in prenatal syphilis screening among women with Medicaid-covered deliveries in Florida. *Maternal Child Health Journal*, *12*(3), 378–393.
Gorsky, R. D., Farnham, P. G., Straus, W. L., Caldwell, B., Holtgrave, D. R., Simonds, R. J., et al.. (1996). Preventing perinatal transmission of HIV-costs and effectiveness of a recommended intervention. *Public Health Reports*, *111*(4), 335–341.
Green, T., Talbot, M. D., & Morton, R. S. (2001). The control of syphilis, a contemporary problem: A historical perspective. *Sexually Transmitted Infections*, *77*(3), 214–217.
Greenspoon, J. S., Martin, J., Greenspoon, R. L., & McNamara, B. T. (1989). Necessity for routine obstetric screening for hepatitis B surface antigen. *Journal of Reproductive Medicine*, *34*(9), 655–658.
Grobman, W. A., & Garcia, P. M. (1999). The cost-effectiveness of voluntary intrapartum rapid human immunodeficiency virus testing for women without adequate prenatal care. *American Journal of Obstetrics and Gynecology*, *181*(5 Pt 1), 1062–1071.
Hallberg, A., Mårdh, P. A., Persson, K., & Ripa, T. (1979). Pneumonia associated with Chlamydia trachomatis infection in an infant. *Acta Paediatrica Scandinavica*, *68*(5), 765–767.
Hamilton, B. E., Martin, J. A., Ventura, S. J., Sutton, P. D., & Menacker, F. (2005). Births: Preliminary data for 2004. *National Vital Statistics Reports*, *54*(8) 1–17.
Hitti, J., Nugent, R., Boutain, D., Gardella, C., Hillier, S. L., & Eschenbach, D. A. (2007). Racial disparity in risk of preterm birth associated with lower genital tract infection. *Paediatric and Perinatal Epidemiology*, *21*, 330–337.
Hollier, L. M., Hill, J., Sheffield, J. S., & Wendel, G. D. Jr. (2003). State laws regarding prenatal syphilis screening in the United States. *American Journal of Obstetrics and Gynecology*, *189*(4), 1178–1183.
Illinois. (2003). HIV Perinatal HIV Prevention Act, Public Act 093-0566.
Institute of Medicine. (1999). *Reducing the odds: Preventing perinatal transmission of HIV in the United States*. Washington, DC: National Academy Press.

Jayaraman, G. C., Preiksaitis, J. K., & Larke, B. (2003). Mandatory reporting of HIV infection and opt-out prenatal screening for HIV infection: Effect on testing rates. *Canadian Medical Association Journal*, *168*(6), 679–682.

Kaiser Family Foundation. (2009). HIV testing for mothers and newborns, February 2007, Retrieved July 13, 2009, from http://www.statehealthfacts.org/comparetable.jsp?ind=563&cat=11.

Kirkham, C., Harris, S., & Grzybowski, S. (2005). Evidence-based prenatal care: Part II. Third-trimester care and prevention of infectious diseases. *American Family Physician*, 71(8), 1555–1560.

Kohl, K. S., Markowitz, L. E., & Koumans, E. H. (2003). Developments in the screening for Chlamydia trachomatis: A review. *Obstetrics and Gynecology Clinics of North America*, *30*(4), 637–658.

Kropp, R. Y., Sarnquist, C. C., Montgomery, E. T., Ruiz, J. D., & Maldonado, Y. A. (2006). A comparison of perinatal HIV prevention opportunities for Hispanic and non-Hispanic women in California. *AIDS Education and Prevention*, *18*(5), 430–443.

Locksmith, G., & Duff, P. (2004). Infection, antibiotics and preterm delivery. *Seminars in Perinatology*, *25*(5), 295–309.

López-Zambrano, M. A., Briceño, G., & Rodriguez-Morales, A. J. (2009). Trends in the prevalence of HIV and syphilis among pregnant women under antenatal care in central Venezuela. *International Journal of Infectious Diseases*, *13*(4), e189–e191.

Martin, J., Hamilton, B., Sutton, P., Ventura, S., Menacker, F., Kimeyer, S., et al. (2009). Births: Final data for 2006. *National Vital Statistics Reports*, *57*(7), 1–102.

Matheson, P. B., Abrams, E. J., Thomas, P. A., Hernán, M. A., Thea, D. M., Lambert, G., et al. (1995). Efficacy of antenatal zidovudine in reducing perinatal transmission of human immunodeficiency virus type 1. The New York City Perinatal HIV Transmission Collaborative Study Group. *Journal of Infectious Diseases*, *172*(2), 353–358.

Merkatz, I. R., & Thompson, J. B. (1990). *New perspectives on prenatal care*. New York: Elsevier.

Meyers, D. S., Halvorson, H., & Luckhaupt, S. (2007). Screening for chlamydial infection: An evidence update for the U.S. Preventive Services Task Force. *Annals of Internal Medicine*, *147*(2), 135–142.

Meyers, D., Wolff, T., Gregory, K., Marion, L., Moyer, V., Nelson, H., et al. (2008). USPSTF recommendations for STI screening. *American Family Physician*, *77*(6), 819–824.

Miller, W. C., Ford, C. A., Morris, M., Handcock, M. S., Schmitz, J. L., Hobbs, M. M., et al. (2004). Prevalence of chlamydial and gonococcal infections among young adults in the United States. *JAMA: The Journal of the American Medical Association*, *291*(18), 2229–2236.

Miller, J. M., Maupin, R. T., & Nsuami, M. (2005). Initial and repeat testing for chlamydia during pregnancy. *Journal of Maternal-Fetal and Neonatal Medicine*, *18*(4), 231–235.

Morabia, A., & Zhang, F. F. (2004). History of medical screening: From concepts to action. *Postgraduate Medical Journal*, *80*(946), 463–469.

Moyer, V. A., & Butler, M. (2004). Gaps in the evidence for well-child care: A challenge to our profession. *Pediatrics*, *114*(6), 1511–1521.

Mrus, J. M., & Tsevat, J. (2004). Cost-effectiveness of interventions to reduce vertical HIV transmission from pregnant women who have not received prenatal care. *Medical Decision Making*, *24*(1), 30–39.

Nesheim, S., Jamieson, D. J., Danner, S. P., Maupin, R., O'Sullivan, M. J., Cohen, M. H., et al. (2007). Primary human immunodeficiency virus infection during pregnancy detected by repeat testing. *American Journal of Obstetrics and Gynecology*, *197*(2), 149, e1–e5.

New Jersey Department of Health and Senior Services. (2001). *Health New Jersey 2010* (Vol. 2, 147). Trenton, NJ: New Jersey Department of Health and Senior Services.

New Jersey Department of Health and Senior Services. (2002). *Standard of care for women who present in labor with unknown HIV serostatus*. Retrieved from http://www.state.nj.us/health/aids/stdcare.pdf.

New Jersey Department of Health and Senior Services. (2005). *HIV/AIDS epidemiologic profile for the state of New Jersey 2005*. Retrieved from http://www.state.nj.us./health/aids/documents/epi2005.pdf.

New Jersey State Nurses Association. (2004). *HIV reproductive counseling and care*. Retrieved January 2009, from http://www.njsna.org/displaycommon.cfm?an=1&subarticlenbr=79.

New York State Department of Health (NYSDOH) (2002). *AIDS in New York State, 2001–02 Edition: Covering data reported through December 1999*. Albany, NY: Author.

New York State Department of Health (NYSDOH) (2004). Health alert: Steps to further reduce mother-to-child HIV transmission in New York State. Retrieved from http://www.health.state.ny.us/diseases/aids/testing/docs/healthalert0107.pdf.

Newell, M., & McIntyre, J. (2000). *Congenital and perinatal infections: Prevention, diagnosis and treatment*. Cambridge: Cambridge University Press.

Pavlin, N. L., Parker, R., Fairley, C. K., Gunn, J. M., & Hocking, J. (2008). Take the sex out of STI screening! Views of young women on implementing chlamydia screening in general practice. *BMC Infection Diseases, 8*, 62. Retrieved from http://www.biomedcentral.com/1471–2334/8/62.

Poirier, S. (1995). *Chicago's war on syphilis, 1937–1940: The Times, the Trib, and the Clap Doctor*. Chicago: University of Illinois Press.

Pulver, W. P., Glebatis, D., Wade, N., Birkhead, G. S., & Smith, P. (2004). Trends from an HIV seroprevalence study among childbearing women in New York State from 1988 through 2000: A valuable epidemiologic tool. *Archives of Pediatric Adolescent Medicine*, *158*(5), 443–448.

Rovira, M., Carreras, E., Bladé, J., Graus, F., Valls, J., Fernández-Avilés, F., et al. (2001). Dramatic improvement of POEMS syndrome following autologous haematopoietic cell transplantation. *British Journal of Haematology*, *115*(2), 373–375.

Sansom, S. L., Jamieson, D. J., Farnham, P. G., Bulterys, M., & Fowler, M. G. (2003). Human immunodeficiency virus retesting during pregnancy: Costs and effectiveness in preventing perinatal transmission. *Obstetrics and Gynecology*, *102*(4), 782–790.

Sarnquist, C. C., Cunningham, S. D., Sullivan, B., & Maldonado, Y. (2007). The effectiveness of state and national policy on the implementation of perinatal HIV prevention interventions. *American Journal of Public Health*, *97*(6), 1041–1046.

Schrag, S. J., Arnold, K. E., Mohle-Boetani, J. C., Lynfield, R., Zell, E. R., Stefonek, K., et al. (2003). Prenatal screening for infectious diseases and opportunities for prevention. *Obstetrics and Gynecology*, *102*(4), 753–760.

Shaffer, N., Chuachoowong, R., Mock, P. A., Bhadrakom, C., Siriwasin, W., Young, N. L., et al. (1999). Short-course zidovudine for perinatal HIV-1 transmission in Bangkok, Thailand: A randomised controlled trial. *Lancet*, *353*(9155), 773–780.

Sheikh, L. A., Sarnquist, C., Grieb, E. M., Sullivan, B., & Maldonado, Y. A. (2009). Prenatal screening for infectious diseases: An analysis of disparities and adherence to policy in California. *Maternal Child Health Journal*, *13*(2), 260–267.

Songok, E. M., Fujiyama, Y., Tukei, P. M., Vulule, J. M., Kiptoo, M. K., Adungo, N. O., et al. (2003). The use of short-course zidovudine to prevent perinatal transmission of human immunodeficiency virus in rural Kenya. *American Journal of Tropical Medicine and Hygiene*, *69*(1), 8–13.

Statton, A., Ayala, L., & Olszewski, Y. (2007). Creating a safety net for HIV exposed infants in Illinois. Perinatal HIV Elimination Project. In *National HIV Prevention Conference, Atlanta*.

St. Lawrence, J. S., Montaño, D. E., Kasprzyk, D., Phillips, W. R., Armstrong, K., & Leichliter, J. S. (2002). STD screening, testing, case reporting, and clinical and partner notification practices: A national survey of US physicians. *American Journal of Public Health*, *92*(11), 1784–1788.

Stringer, J. S., & Rouse, D. J. (1999). Rapid testing and zidovudine treatment to prevent vertical transmission of human immunodeficiency virus in unregistered parturients: A cost-effectiveness analysis. *Obstetrics and Gynecology*, *94*(1), 34–40.

Stringer, E. M., Stringer, J. S., Cliver, S. P., Goldenberg, R. L., & Goepfert, A. R. (2001). Evaluation of a new testing policy for human immunodeficiency virus to improve screening rates. *Obstetrics and Gynecology*, *98*(6), 1104–1108.

Trepka, M. J., & Bloom, S. A. (2006). Inadequate syphilis screening among women with prenatal care in a community with a high syphilis incidence. *Sexually Transmitted Diseases*, *33*(11), 670–674.

U.S. Department of Health and Human Services, Expert Panel on the Content of Prenatal Care. (1989). *Caring for our future: The content of prenatal care*. Washington, DC: Public Health Service.

U.S. Preventive Services Task Force. (2004a). Screening for syphilis infection: U.S. Preventive Services Task Force recommendation statement. *Annals of Family Medicine*, *2*(4), 362–365.

U.S. Preventive Services Task Force. (2004b). Screening for hepatitis B virus infection: U.S. Preventive Services Task Force recommendation statement. Rockville, MD: Agency for Healthcare Research and Quality. http://www.ahrq.gov/clinic/3rduspstf/hepbscr/hepbrs.htm.

U.S. Preventive Services Task Force. (2005). Screening for gonorrhea infection: U.S. Preventive Services Task Force recommendation statement. *Annals of Family Medicine*, *3*(3), 263–267.

U.S. Preventive Services Task Force. (2007). Screening for chlamydial infection: U.S. Preventive Services Task Force recommendation statement. *Annals of Internal Medicine*, *147*(2), 128–135.

Wade, N. A., Birkhead, G. S., Warren, B. L., Charbonneau, T. T., French, P. T., Wang, L., et al. (1998). Abbreviated regimens of zidovudine prophylaxis and perinatal transmission of the human immunodeficiency virus. *New England Journal of Medicine*, *339*(20), 1409–1414.

Wade, N. A., Zielinski, M. A., Butsashvili, M., McNutt, L. A., Warren, B. L., Glaros, R., et al. (2004). Decline in perinatal HIV transmission in New York State (1997–2000). *Journal of Acquired Immune Deficiency Syndrome*, *36*(5), 1075–1082.

Wiktor, S. Z., Ekpini, E., Karon, J. M., Nkengasong, J., Maurice, C., Severin, S. T., et al. (1999). Short-course oral zidovudine for prevention of mother-to-child transmission of HIV-1 in Abidjan, Cote d'Ivoire: A randomised trial. *Lancet*, *353*(9155), 781–785.

Williams, D. R. (2002). Racial/ethnic variations in women's health: The social embeddedness of health. *American Journal of Public Health*, *92*(4), 588–597.

Yudin, M. H., Moravac, C., & Shah, R. R. (2007). Influence of an "opt-out" test strategy and patient factors on human immunodeficiency virus screening in pregnancy. *Obstetrics and Gynecology*, *110*(1), 81–86.

Chapter 8
What is the Role of Prenatal Care in Reducing Racial and Ethnic Disparities in Pregnancy Outcomes?

Hannah Holcombe Walford, Sonya Trinh, Anna Wiencrot, and Michael C. Lu

> *After a comprehensive review of the literature on the value of prenatal care, the committee concluded that the overwhelming weight of the evidence is that prenatal care reduces low birthweight. This finding is strong enough to support a broad, national commitment to ensuring that all pregnant women in the United States, especially those at medical or socioeconomic risk, receive high-quality prenatal care.*
>
> Institute of Medicine (IOM), Preventing Low Birthweight (1985b, pp. 18–19)

According to the American Academy of Pediatrics (AAP) and the American College of Obstetricians and Gynecologists (ACOG), comprehensive prenatal care "involves a coordinated approach to medical care and psychosocial support that optimally begins before conception and extends throughout the antepartum period" (AAP/ACOG, 2007, p. 83). It consists of a series of clinical visits and ancillary services designed to promote the health and well-being of the mother, fetus, and family. Its three major components, as defined by the US Public Health Service Expert Panel on the Content of Prenatal Care [US Department of Health and Human Services (USDHHS), 1989] include: (1) early and continuing risk assessment; (2) health promotion; and, (3) medical and psychosocial interventions and follow-up.

Prenatal care has been offered to pregnant women in the U.S. for nearly 100 years, beginning with Mrs. William Lowell Putnam making home visits to pregnant women registered at the Boston Lying-In Hospital in 1909 (Alexander & Kotelchuck, 2001). Maternal morbidity and mortality, particularly related to complications of preeclampsia and eclampsia, were among the earliest targets of prenatal care. During the 1900s, support grew for the hypothesis that prenatal care could reduce the risk of infant mortality from LBW and preterm birth. In 1915, J. Withridge Williams of the Johns Hopkins Hospital, in championing the potential benefits of prenatal care, asserted that "prenatal care and instruction offer great possibilities for the diminution in the number of deaths [due to prematurity]" (p. 99). In 1947, Eastman described a marked reduction in risk for low birthweight among mothers who received "adequate care" (3+ visits, p. 347).

Several studies (Eisner, Brazie, Pratt, & Hexter, 1979; Greenberg, 1983; Taffel, 1978) published in the 1970s and early 1980s found a significant association between no prenatal care and the incidence of LBW, although none of these studies controlled for possible gestational age bias. In 1973,

M.C. Lu (✉)
Department of Community Health Sciences, David Geffen School of Medicine at UCLA,
UCLA School of Public Health, 650 Charles E. Young Drive S, CHS 36-070B, Box 951772,
Los Angeles, CA 90095-1772, USA
e-mail: mclu@ucla.edu

A. Handler et al. (eds.), *Reducing Racial/Ethnic Disparities in Reproductive and Perinatal Outcomes: The Evidence from Population-Based Interventions*, DOI 10.1007/978-1-4419-1499-6_8,

Kessner, Singer, Kalk, and Schlesinger demonstrated a systematic relationship between categories of adequate prenatal care utilization and LBW, developing the first prenatal care index. When Gortmaker reanalyzed the same data in 1979 using a modified Kessner Index, he found that the percentage of LBW births decreased with increasing adequacy of prenatal care use and that the relationship between prenatal care and infant mortality was attributed to the impact of prenatal care on LBW. Citing these and several other studies, the 1985 IOM report, *Preventing Low Birthweight*, concluded that the "overwhelming weight of the evidence is that prenatal care reduces low birthweight" (IOM, 1985b, p. 18).

Soon thereafter, in the mid- and late 1980s, the U.S. Congress enacted a series of legislative initiatives that incrementally expanded Medicaid eligibility to low-income pregnant women and children, independent of their welfare status (Hill, 1992). Many states followed by further expanding Medicaid eligibility and streamlining the process of enrollment into prenatal care. Arguments for expanding access to prenatal care were buttressed by cost-effectiveness analyses (IOM, 1985a) that suggested savings could be achieved by reducing LBW, although the cost-savings may have been overstated (Huntington & Connell, 1994). In 1986, the US Public Health Service assembled an expert panel to assess the content of prenatal care, which published its landmark report in 1989 (USDHHS, 1989). Following the report, several states expended considerable effort to enhance the content of prenatal care, motivated in part by the expectation that increases in early initiation and adequate utilization of high-quality prenatal care would lower the risk of LBW and, as a result, reduce infant mortality rates.

Partly as a result of these national and state policies, the use of early and adequate prenatal care has increased substantially over the past decade (Kogan et al., 1998; Martin et al., 2002, 2007). This increase, however, has not led to a significant decline in LBW and preterm births, as shown in Table 8.1.

While changes in maternal demographics, increases in multiple gestation, and advances in medical technology may have contributed to the rise in LBW (Alexander & Slay, 2002), some have begun to question the effectiveness of prenatal care in preventing LBW. As early as 1962, Schwartz suggested that gestational age may well be confounding the association between LBW and the trimester in which prenatal care began or the number of prenatal care visits. Terris and Glasser concluded from their life table analysis in 1974 that "early birth prevents the initiation of prenatal care instead of vice versa" (p. 870). Although the Kessner Index adjusts for gestational age at delivery, it fails to accurately reflect AAP-ACOG recommendations regarding the number of visits for "adequate" care, resulting in residual gestational age bias. Indeed, when other indices (Alexander & Kotelchuck, 1996; Kotelchuck, 1994) that better account for gestational age bias are used, the incremental relationship between less adequate use of prenatal care and LBW traditionally observed when using the Kessner Index disappears (Alexander & Kotelchuck, 2001).

Another important limitation of most observational studies examining the effectiveness of prenatal care has been their failure to adequately control for critical confounders and selection bias (Lu, Tache, Alexander, Kotelchuck, & Halfon, 2003). Women who seek prenatal care early may differ from those who seek prenatal care late or not at all. Similarly, women who attend all of their prenatal appointments may differ in many ways from those who miss most of their prenatal visits. Women who seek prenatal care early and attend all their prenatal appointments may be more likely to engage in other advantageous health-care-seeking and health-promoting behaviors, including planning their pregnancies, obtaining preconception and interconception care, maintaining a healthy diet, and abstaining from the use of tobacco, alcohol, and other drugs. They may also command more resources that promote good health before and during pregnancy. Because these advantageous behaviors and resources may contribute to reducing their risk of LBW deliveries, the adequacy of their prenatal care utilization could be conceptualized as a proxy indicator for a myriad of health-enhancing maternal behaviors and resources, rather than having a direct cause-effect relationship with LBW.

Randomization can help avoid some of the problems associated with potential confounding and selection bias. For ethical considerations, no study has examined the effectiveness of prenatal care

Table 8.1 Racial and ethnic disparities in prenatal care, low birthweight and preterm birth, U.S., 1990–2005

	First trimester prenatal care (%)			Low birth weight (%)			Preterm births (%)		
Year	All races	Non-Hispanic white	Non-Hispanic black	All races	Non-Hispanic white	Non-Hispanic black	All races	Non-Hispanic white	Non-Hispanic black
1990	75.8	83.3	61.0	7.0	5.6	13.3	10.6	8.5	18.9
1991	76.2	83.7	61.9	7.1	5.7	13.6	10.8	8.7	19.0
1992	77.7	84.9	64.0	7.1	5.7	13.4	10.7	8.7	18.5
1993	78.9	85.6	66.1	7.2	5.9	13.4	11.0	9.1	18.6
1994	80.2	86.5	68.3	7.3	6.1	13.3	11.0	9.3	18.2
1995	81.3	87.1	70.6	7.3	6.2	13.2	11.0	9.4	17.8
1996	81.9	87.4	71.5	7.4	6.4	13.1	11.0	9.5	17.5
1997	82.5	87.9	72.3	7.5	6.5	13.1	11.4	9.9	17.6
1998	82.8	87.9	73.3	7.6	6.6	13.2	11.6	10.2	17.6
1999	83.2	88.4	74.1	7.6	6.6	13.2	11.8	10.5	17.6
2000	83.2	88.6	74.3	7.6	6.6	13.1	11.6	10.4	17.4
2001	83.3	88.6	74.5	7.7	6.8	13.1	11.9	10.8	17.6
2002	83.7	88.7	75.3	7.8	6.9	13.4	12.1	11.0	17.7
2003	84.0	89.1	76.2	7.9	7.0	13.6	12.3	11.3	17.8
2004	84.2	89.0	76.3	8.1	7.2	13.7	12.5	11.5	17.9
2005	83.9	88.7	76.5	8.2	7.3	14.0	12.7	11.7	18.4

Source: Martin, J. A., Hamilton, B. E., Sutton, P. D., Ventura, S. J., Menacker, F., Kirmeyer, S., et al. (2007). Births: Final data for 2005. *National Vital Statistics Reports, 56*(6), 1–103. Hyattsville, MD: National Center for Health Statistics

Martin, J. A., Hamilton, B. E., Sutton, P. D., Ventura, S. J., Menacker, F., Kirmeyer, S. (2006). Births: Final data for 2004. *National Vital Statistics Reports, 55*(1), 1–101. Hyattsville, MD: National Center for Health Statistics

Martin, J. A., Hamilton, B. E., Sutton, P. D., Ventura, S. J., Menacker, F., Munson, M. L. (2005). Births: Final data for 2003. *National Vital Statistics Reports, 54*(2), 1–116. Hyattsville, MD: National Center for Health Statistics

by randomizing pregnant women to receiving prenatal care vs. no prenatal care. Several studies have randomized pregnant women to more vs. fewer visits and found no significant difference in birth outcomes (Villar, Carroli, Khan-Neelofur, Piaggio, & Gülmezoglu, 2002). A number of studies have randomized pregnant women to receiving standard vs. enhanced prenatal care with added components, such as preterm birth education. A systematic review by Fiscella in 1995 failed to demonstrate the effectiveness of these enhanced prenatal care programs in preventing preterm birth or LBW.

Both the 1995 Fiscella review and an additional 1995 review (Alexander & Korenbrot) raised other concerns regarding the validity of the evidence used to support the effectiveness of prenatal care. Citing problems with inconsistent results, insufficient adjustment for prematurity bias, and inadequate control for the effect of critical confounders and potential selection bias in earlier studies, Fiscella concluded that "current evidence does not satisfy the criteria necessary to establish that prenatal care definitely improves birth outcomes" (1995, p. 475). Alexander and Korenbrot (1995) also concluded from their systematic review that "[t]here is little done during the standard prenatal care visit that could be expected to reduce low birth weight" (p. 113), although they found prenatal care to have a positive effect on LBW at term. In a more recent review of the *content* of prenatal care in 2003, Lu et al. concluded that neither preterm birth nor intrauterine growth restriction (IUGR) – the twin constituents of LBW – can be effectively prevented by prenatal care in its present form.

Thus two decades following the 1985 IOM report, the effectiveness of prenatal care in preventing LBW remains a subject of great controversy. Furthermore, the claim that increasing access to and utilization of prenatal care can help reduce racial-ethnic disparities in LBW and related outcomes has yet to be validated. The primary aim of this chapter is to review the evidence of the effectiveness of prenatal care in preventing LBW, with an emphasis on its effectiveness in reducing racial-ethnic disparities in LBW.

Methods

Independent literature searches were conducted by two co-authors (HH and ST) to gather evidence on the effectiveness of prenatal care in preventing LBW. Since Fiscella (1995) had previously conducted an excellent systematic review of the literature between 1966 and 1994, we focused our review on studies published between January 1995 and August 2006. Studies were retrieved during July–August 2006 from PubMed and MDConsult using the search terms "prenatal care," "prenatal care adequacy," "prenatal care utilization," "enhanced prenatal care," "randomized," "low birth weight," and "preterm birth." In a second search, the references of retrieved articles were hand-searched for relevant studies (i.e., the snowball technique). No search software was used, no efforts were made to identify unpublished studies, and no contacts were made with the authors.

The intervention was defined broadly as prenatal care utilization; we did not review evidence of the effectiveness of any specific components of prenatal care. The studies fell into two broad categories: utilization studies and enhanced studies. Utilization studies are typically cohort or case-control (observational) studies aimed to establish whether adequate prenatal care is associated with better outcomes than inadequate care. Enhanced care studies as defined in this chapter are randomized controlled trials designed to determine if enhanced prenatal care provided to women at high risk produces better pregnancy outcomes than standard prenatal care. We excluded a third category of studies which use ecological designs to examine whether a change in availability of prenatal care affects population birth outcome statistics because such studies while producing results at the population level do not allow conclusions at the individual level. We focused on LBW, defined as birthweight of less than 2,500 g, or related outcomes including:

preterm delivery (PTD, delivery before 37 completed weeks' gestation), intrauterine growth restriction (IUGR, defined as birthweight below the tenth percentile in most studies), small-for-gestational age (SGA), very LBW (VLBW, birthweight less than 1,500 g), and very PTD (VPTD, delivery before 32 completed weeks' gestation). These birth outcomes were included in the review because they are related to LBW. Specifically, PTD and IUGR, two outcomes with overlapping but also divergent pathways, are two causes of LBW. Although the pathway for SGA is undefined in most studies, IUGR is one of several causes of SGA. VLBW and VPTD are subsets of LBW and PTD, respectively. Lastly, studies that treated birthweight and/or gestational age as continuous outcomes were also included.

Specific inclusion and exclusion criteria were established prior to the literature searches, modeled after the criteria used in the 1995 Fiscella study. Utilization studies were included if they statistically adjusted for potential confounders and used an adjustment factor for prenatal visits relative to gestational age (e.g., the Kessner Index, the Kotelchuck Index, R-Gindex, or a comparable factor) (Alexander & Kotelchuck, 1996; Kessner et al., 1973; Kotelchuck, 1994). Studies not primarily designed to assess the effects of prenatal care were included if they met all the aforementioned criteria. Studies were excluded if prenatal care was treated as a categorical variable (presence vs. absence of prenatal care). Enhanced care studies were included if the subjects were randomly assigned to either standard or enhanced prenatal care. Randomized controlled trials were excluded if there was evidence of contamination of treatment and control groups. Studies were excluded if LBW or a related outcome (VLBW, PTD, VPTD, IUGR, SGA) was not reported as a study outcome. Studies were also excluded if they were published solely in a foreign language or conducted outside of the U.S., since the experiences of racial-ethnic groups in the U.S. may differ from those of the same racial-ethnic groups living outside of the U.S. We searched all included studies for data related to racial-ethnic disparities in LBW.

Results

Our search identified 31 studies published between January 1995 and August 2006 that examined the basic relationship between prenatal care and LBW. Seven studies met both our inclusion and exclusion criteria: two utilization studies (Collins, Herman, & David, 1997; Krueger & Scholl, 2000) and five enhanced care studies (Brooten et al., 2001; Kitzman et al., 1997; Klerman et al., 2001; Little, Saul, Testa, & Gaziano, 2002; Moore et al., 1998). In Table 8.2, the utilization studies from this review have been added to those from the Fiscella review (Gortmaker, 1979; Kogan, Alexander, Kotelchuck, & Nagey, 1994; Malloy, Kao, & Lee, 1992; Murray & Bernfield, 1988; Mustard & Roos, 1994; Parker, McFarlane, & Soeken, 1994; Quick, Greenlick, & Roghmann, 1981; Raine, Powell, & Krohn, 1994; Scholl, Miller, Salmon, Cofsky, & Shearer, 1987; Schramm, 1992; Shiono, Kebanoff, Graubard, Berendes, & Rhoads, 1986; Showstack, Budetti, & Minkler, 1984; Terris & Glasser, 1974; Tyson et al., 1990). In Table 8.3, the enhanced care studies from this review have been added to those from the Fiscella review (Bryce, Stanley, & Garner, 1991; Collaborative Group on Preterm Birth Prevention, 1993; Goldenberg et al., 1990; Graham, Frank, Zyzanski, Kitson, & Reeb, 1992; Heins, Nance, McCathy, & Efird, 1990; Main, Gabbe, Richardson, & Strong, 1985; Main, Richardson, Hadley, & Gabbe, 1989; McLaughlin et al., 1992; Olds, Henderson, Tatelbaum, & Chamberlain, 1986; Spencer, Thomas, & Morris, 1989; Villar et al., 1992).

Table 8.4 lists all the excluded studies (Armson, Dodds, Haliburton, Cervin, & Rinaldo, 2003; Barnet, Duggan, & Devoe, 2003; Barros, Tavares, & Rodrigues, 1996; Binstock & Wolde-Tsadik, 1995; Blanchette, 1995; Boss & Timbrook, 2001; Dyson et al., 1998; Edwards et al., 1995; Gómez-Olmedo, Delgado-Rodriguez, Bueno-Cavanillas, Molina-Font, & Gálvez-Vargas, 1996; Helfand & Zimmer-Gembeck, 1997; Herman et al., 1996; Homan & Korenbrot, 1998; Hueston, 1995;

Table 8.2 Observational studies on the effect of prenatal care on birth outcomes

Authors (year)	Populations studied/ sample size	Adjustment for potential confounders	Outcome measures	Key findings related to intervention effectiveness	Address disparities	Findings support the intervention? For which populations?
Terris and Glasser (1974)[a]	NYC, 1991, black birth certificates/34,949	Age, parity, married, hospital, infant sex	% LBW vs. NBW	NS, 9.6 vs. 11.0%[b]	Yes	No Blacks
Gortmaker (1979)[a]	NYC, 1968, birth certificates/90,339	Age, education, parity, married, medical or OB risk, hospital	Odds ratio	0.71[c] (White), 0.56[c] (Black)	Yes	Yes
			LBW	NS, 0.94[b] (White), 0.83[c] (Black) (Protective OR calculated by chapter authors)		Whites, Blacks
			Neonatal mortality			
Quick et al. (1981)[a]	Portland, OR, 1973, white birth and death certificates/23,264	Age, education, HMO provider, birth order, married, medical risk	Improvement in mean birth weight	160 g[c]	No	Yes Whites
Showstack et al. (1984)[a]	CA, 1978, birth certificates/18,470	Age, race, education, multiple gestation, medical risk, hospital	Improvement in mean birth weight	207 g[d]	No	Yes
Shiono et al. (1986)[a]	Kaiser Birth Defects Study, 1974–1977/29,415	Age, race, education, married, employed,infant sex, parity, OB history, smoking, alcohol	Improvement in mean birth weight	29 g[c]	No	Yes
Scholl et al. (1987)[a]	Adolescent Program data 1983–1984/757	Age, race, parity, insurance, medical risk, alcohol, smoking	Odds ratio		No	No (LBW)
			LBW	NS, 1.35[b] (0.41–4.44) 0.34[c] (0.15–0.73)		Yes (PTB)
			Preterm delivery			
Murray and Bernfield (1988)[a]	CA, 1978, birth certificates/31,000	Age, education	Odds ratio	0.29[d] (Black), 0.62[d] (White)	Yes	Yes
			LBW			Whites, Blacks
Tyson et al. (1990)[a]	Hospital data 1977–1980/28,838	Age, race, education, married, education, multiple gestation, OB history	Odds ratio		No	No
			Short gestation	NS[b] for any cohort 0.80[e], 38 weeks cohort		
			SGA	0.67[e], 42 weeks cohort		
			Stillbirth	0.68[e], 34 weeks cohort 0.23[e], 42 weeks cohort		
			Infant death			
Schramm (1992)[a]	Birth certificates, Medicaid claims data 1988/12,023	Age, race, education, birth order, married, smoking	Relative risk		No	Yes
			LBW	0.63[e]		
			VLBW	0.66[e]		

Malloy et al. (1992)[a]	Birth or death certificates, 1980–1984/374,244	Age, race, education, parity, multiple gestation	% Mortality (stillbirth or infant death)	Reduced only among term infants[f]	No	Yes
Kogan et al. (1994)[a]	National Maternal and Infant Health Survey, 1988/9,394	Age, race, parity, married, employed, income, site of care, OB history, hypertension, smoking, received recommended advice	Odds ratio LBW	 0.48[c] (0.39–0.59)	No	Yes
Parker et al. (1994)[a]	Medical record retrospective review, 1990–1993/1,203	Age, married, education, parity, OB history, weight gain, interpregnancy interval, battering, infection, bleeding in first or second trimester, anemia, smoking, alcohol and drug use	Relative risk LBW	 NS, 1.1[b] (0.9–1.1)	No	No
Mustard and Roos (1994)[a]	Winnipeg hospital records 1987–1988/12,646	Age, married, income, parity, native, OB history, smoking, prenatal education, pregnancy complication	Improvement in mean birth weight	58 g[f]	No	Yes
Raine et al. (1994)[a]	WA, white women with two births, 1984–1990/3,334	Age, married, prior LBW birth, prior miscarriage, smoking, interpregnancy interval	Relative risk LBW	 NS, 0.77[b] (0.5–1.11)	No	No Whites
Collins et al. (1997)	Chicago, IL, Birth files, 1982–1983/81,427	Ethnicity, income	Odds ratio LBW, with inadequate PNC LBW, with adequate PNC	 4.8 (3.4–3.8) (Black vs. Mexican American) 2.0 (1.5–2.7) (Black vs. White) 3.8 (2.5–5.9) (Black vs. Mexican American) 2.5 (1.8–3.5) (Black vs. White)	Yes	LBW declines w/ increasing PNC, but PNC does not eliminate disparities in LBW

(continued)

Table 8.2 (continued)

Authors (year)	Populations studied/ sample size	Adjustment for potential confounders	Outcome measures	Key findings related to intervention effectiveness	Address disparities	Findings support the intervention? For which populations?
Krueger and Scholl (2000)	NJ, Camden Study, 1985–1995/1,771	Age, parity, ethnicity, pregravid body mass index, inadequate weight gain for gestation, cigarettes per day, prior preterm delivery, prior low birth weight	Odds Ratio (intermediate/ adequate vs. inadequate)		No	Yes, for PTD delivery
			LBW	1.46[b] (1.00–2.12) (Kessner, 1973) 1.14[b] (0.80–1.62) (Kotelchuck, 1994) 2.80 (2.07–3.78) (Kessner, 1973) 2.10 (1.58–2.81) (Kotelchuck, 1994)		
			Preterm delivery			

LBW low birth weight; *NBW* normal birth weight; *NS* not significant; *OB* obstetric; *SGA* small for gestational age; *HMO* health maintenance organization; *BMI* body mass index; *WIC* women, infant, and children; *PNC* prenatal care; *PTD* preterm delivery

[a]Studies included in Fiscella (1995)

Risk ratios for inadequate care have been inverted to risk rates for adequate care to facilitate comparison:

[b]$p > 0.05$

[c]$p < 0.01$

[d]$p < 0.0001$

[e]$p < 0.05$

[f]$p < 0.001$

Table 8.3 Randomized controlled trials of enhanced prenatal care (studies are listed within enhancement type)

Primary author	Sample description (group sizes)	Intervention	Outcome measure	Results	Caveats/biases	Supports intervention? For which populations?
Main (1985)[a]	Black, high risk (treatment: 64, control: 68)	Preterm birth prevention: weekly/biweekly visits and cervical examinations, education, hot line	Mean birth weight Mean GA % Preterm births	NS NS NS	Inadequate sample size	No
Main (1989)[a]	Indigent, black, high risk (treatment: 198, control: 178)	Preterm birth prevention, weekly or biweekly visits and cervical exams, education	Mean GA Mean birth weight % Preterm births	NS NS NS	90% Power to detect 50% reduction in preterm delivery	No
Goldenberg (1990)[a]	Predominantly black, indigent, high risk, (treatment: 474, control: 463)	Preterm birth prevention, weekly visits and cervical exams, education	% Preterm births % LBW	NS NS	Trend toward worse outcomes among treatment group	No
Collaborative Group (1993)[a]	Predominantly black, high risk (multicenter) (treatment: 1,200, control: 1,195)	Preterm birth prevention, weekly or biweekly visits, cervical exams, education	% Preterm births % LBW % VLBW	NS NS NS	Investigators disagreed about reasons for negative results	No
Heins (1990)[a]	High risk (treatment: 667, control 679)	Comprehensive care assessment and case management, weekly visits, education, follow up	% LBW	NS	90% Power to detect 38% reduction in LBW	No
McLaughlin (1992)[a]	Indigent, less than 28 weeks' gestation (treatment: 217, control 211)	Comprehensive care, multidisciplinary team approach, education, psychosocial support, follow up	% LBW	NS	Possible benefit for primiparous women ($p < 0.05$); higher dropout rate in control group (13 vs. 8%)	No
Klerman (2001)	African-American, high risk (treatment: 318 , control: 301)	Comprehensive care, multidisciplinary team, psychosocial support, education, smoking cessation, additional appointments	Mean birth weight Mean GA IUGR % Preterm birth	NS NS NS NS	Positive trends for intervention in outcome variables; Small sample size; Exclusion of women with high risk medical conditions	No

Table 8.3 (continued)

Primary author	Sample description (group sizes)	Intervention	Outcome measure	Results	Caveats/biases	Supports intervention? For which populations?
Moore (1998)	African-American, white or "other" ethnicity, high risk (treatment: 557 AA, 161 white, control: 556 AA, 159 white)	Telephone intervention by registered nurses	Mean weight Mean GA % LBW % Preterm	NS NS NS NS	1,573 women excluded because they could not be contacted by telephone	Total sample showed no treatment effect on outcomes. Greatest statistical significance in subgroup of AA women age 19 and older
Little (2002)	Predominantly African-American, Hispanic, high risk (treatment: 61, control: 50)	Telephonic nursing care and education	Mean gestation age Mean birthweight	NS SS SS	Small sample size	No significant difference in mean gestational age. Increased mean birth weight in treatment group
Olds (1986)[a]	White, high risk (treatment: 189, control: 165)	Home visitation by trained nurses	% LBW % Preterm births Mean birth weight	NS NS NS		Potential reduction in LBW and preterm birth in women ages 14–16
Spencer (1989)[a]	British, high risk (treatment: 271, control: 633)	Home visitation by trained lay workers	% LBW % Preterm birth % SGA	NS NS NS	Failed to analyze results by intention to treat	No
Bryce (1991)[a]	Australian, high risk (treatment: 983, control: 987)	Home visitation by trained midwives and telephone calls	% Preterm birth	NS	Analyzed by intention to treat	No overall effect of social support. Significant reduction in preterm birth in higher social class. Potential benefit for women with multiple gestations or prior preterm birth
Graham (1992)[a]	Black, high risk (treatment: 87, control: 58)	Home visitation byblack paraprofessionals	% LBW	NS	Inadequate sample size, failed to analyze results by intention to treat	No

Villar (1992)[a]	Latin American, high risk (mulitcenter) (treatment: 1,115, control: 1,120)	Home visitation by trained social workers	% LBW % Preterm births	NS NS	80% Power to detect 6% reduction in LBW	No
Kitzman (1997)	African-American, low income (treatment: 458, control: 681)	Home visitation by trained nurses	Mean birth weight Mean GA % LBW % IUGR % Preterm births	NS NS NS NS NS	High risk medical conditions excluded	No Other beneficial effects noted, but did not reduce LBW or preterm birth
Brooten (2001)	Predominantly African-American, high risk (treatment: 85 women, 95 infants, control: 88 women, 100 infants)	Home visitation by trained nurses, telephone outreach, postpartum visits	*Preterm infant*: Mean GA Mean birth weight *Term infant*: Mean GA Mean birth weight *Twin infants*: No. preterm Mean GA Mean birth weight	 NS SS NS NS SS SS SS	Small sample size	Treatment statistically significant for twin pregnancies for both PTD and LBW. Mean birth weight was greater in treatment group in preterm infant

[a]Studies included in Fiscella (1995)

NS not statistically significant; *SS* statistically significant, $p > 0.05$; *LBW* low birth weight; *GA* gestational age; *SGA* small for gestational age; *VLBW* very low birth weight

Hueston, Gilbert, Davis, & Sturgill, 2003; Ickovics et al., 2003; Jackson et al., 2003; Laditka, Laditka, Mastanduno, Lauria, & Foster, 2005; Lazariu-Bauer, Stratton, Pruzek, & Woelfel, 2004; McDuffie, Beck, Bischoff, Cross, & Orleans, 1996; Mvula & Miller, 1998; Partridge & Holman, 2005; Perkocha, Novotny, Bradley, & Swanson, 1995; Quinlivan & Evans, 2004; Reichman & Teitler, 2005; Sánchez-Nuncio, Pérez-Toga, Pérez-Rodriguez, & Vázquez-Nava, 2005; Tasnim, Mahmud, & Arif, 2005; Taylor, Alexander, & Hepworth, 2005; Vintzileos, Ananth, Smulian, & Scorza, 2003; Vintzileos, Ananth, Smulian, Scorza, & Knuppel, 2002; Visintainer et al., 2000; Zotti & Zahner, 1995) and rationale for exclusion. Table 8.5 presents quality ratings for the seven reviewed studies based on the Quality Checklist for RCTs and Observational Studies of Treatment Studies (see Chap. 2). Below we summarize the design and findings of each of the seven included studies.

Utilization Studies

Our search identified two prenatal care utilization studies (Collins, Herman, & David, 1997; Krueger & Scholl, 2000) that fit our inclusion/exclusion criteria since the last major review by Fiscella (1995). The characteristics and outcomes of the two included utilization studies are summarized in Table 8.2.

Collins et al. (1997) conducted a retrospective cohort study using 1982–1983 Chicago birth files to examine the relationship between prenatal care utilization, maternal ethnicity, and LBW. Of the 81,427 singleton birth files, 54% were African-American, 13% Mexican-American, and 34% white. Income data were obtained by linking birth files to census tract information. Prenatal care utilization was classified using the Kotelchuck Adequacy of Prenatal Care Utilization Index. The study found that adequacy of prenatal care utilization varied by race and place of residence. The African-American birthweight disadvantage persisted even among infants born to mothers in moderate-income areas (median family annual income of $20,001–$30,000) who received adequate and adequate-plus prenatal care. Similarly, although race-specific term (gestational age >37 weeks) LBW rates declined as prenatal care usage rose, the position of African-Americans relative to Mexican-Americans and whites was essentially unchanged. For example, among mothers with inadequate prenatal care residing in low-income areas, African-Americans had a 4.8 (95% CI 3.4, 3.8) times greater risk of LBW than Mexican-Americans and a 2.0 (95% CI 1.5, 2.7) times greater risk than whites. Among mothers with adequate prenatal care residing in low-income areas, African-Americans had a 3.8 (95% CI 2.5, 5.9) times greater risk of LBW than whites and a 2.5 (95% CI 1.8–3.5) times greater risk than Mexican-Americans. The authors concluded that maternal race or some factor closely related to it affects pregnancy outcome regardless of the adequacy of prenatal care utilization.

Krueger and Scholl (2000) conducted a retrospective cohort study in Camden, New Jersey to examine the association between prenatal care utilization and preterm birth. The study analyzed data from 1,771 pregnant women enrolled in a study of maternal growth among young gravidas. The study population included approximately 58% African-American, 33% Hispanic and 9% white women. Women were excluded from the study if they had serious medical complications. Prenatal care utilization was classified using both the Kessner and the Kotelchuck indices. PTD was measured using both the last menstrual period and the obstetric estimate for length of gestation. Logistic regression was used to control for potential confounding variables including black ethnicity, maternal age, pregravid body mass index, parity, inadequate weight gain for length of gestation, smoking, and previous delivery of low birth weight or preterm infant. The analysis compared women receiving inadequate care to women receiving adequate or intermediate care, instead of comparing each prenatal care utilization level separately. The study found that women who received inadequate care were at greater risk of having a PTD. The association between prenatal care

Table 8.4 Excluded studies and rationale for exclusion

Authors (year)	Study design	Rationale for exclusion
Edwards et al. (1995)	Retrospective cohort	Enhanced care study, but not randomized.
Perkocha et al. (1995)	Retrospective cohort	Enhanced care study, but not randomized. Compared two enhanced prenatal care programs (CPSP and CTAPPP).
Binstock and Wolde-Tsadik (1995)	Randomized control trial	Failed to use Kotelchuck, Kessner or comparable index: Counted number of visits.
Hueston (1995)	Retrospective cohort	Failed to use Kotelchuck, Kessner or comparable index. Compares prenatal care in urban and rural settings.
Zotti and Zahner (1995)	Retrospective cross-sectional study	Enhanced care study, but not randomized.
Blanchette (1995)	Retrospective cohort	Enhanced care study, but not randomized. Compared two different settings (private vs. public).
Gómez-Olmedo et al. (1996)	Case-control	Conducted outside of US
Barros et al. (1996)	Retrospective cohort	Conducted outside of US
McDuffie et al. (1996)	Randomized control trial	Failed to use Kotelchuck, Kessner or comparable index: Counted number of visits.
Herman et al. (1996)	Retrospective cohort	Enhanced care study, but not randomized.
Helfand and Zimmer-Gembeck (1997)	Retrospective cohort	Examined specific component of prenatal care.
Homan and Korenbrot (1998)	Retrospective cohort	Enhanced care study, but not randomized.
Dyson et al. (1998)	Randomized control trial	Compared different types of enhanced care.
Mvula and Miller (1998)	Prospective cohort	Enhanced care study, but not randomized.
Visintainer et al. (2000)	Prospective cohort	Enhanced care study, but not randomized.
Boss and Timbrook (2001)	Retrospective cohort	Examined continuity of care rather than utilization of care.
Vintzileos et al. (2002)	Retrospective cohort	Failed to use Kotelchuck, Kessner or comparable index: Treated PNC as a dichotomous variable (presence vs. absence of PNC).
Vintzileos et al. (2003)	Cohort	Failed to use Kotelchuck, Kessner or comparable index: Treated PNC as a dichotomous variable (presence vs. absence of PNC).
Jackson et al. (2003)	Cohort	Enhanced care study, but not randomized. Did not measure birth outcomes of interest.
Ickovics et al. (2003)	Prospective matched cohort	Enhanced care study, but not randomized.
Barnet et al. (2003)	Retrospective cohort	Enhanced PNC study, but not randomized. Compared PNC in different settings (school vs. hospital based).
Hueston et al. (2003)	Retrospective cross-sectional study	Failed to use Kotelchuck, Kessner or comparable index: Analyzed by trimester of PNC initiation.
Armson et al. (2003)	Case-control	Enhanced PNC study, but not randomized. Study conducted outside the US.
Quinlivan and Evans (2004)	Prospective cohort	Enhanced PNC study, but not randomized. Conducted outside of the US.

(continued)

Table 8.4 (continued)

Authors (year)	Study design	Rationale for exclusion
Lazariu-Bauer et al. (2004)	Retrospective cohort	Comparison of early vs. late PNC initiation.
Partridge and Holman (2005)	Retrospective cohort	Failed to use Kotelchuck, Kessner or comparable index: Counted number of visits.
Taylor et al. (2005)	Retrospective cohort (cluster analysis)	Failed to use Kotelchuck, Kessner or comparable index: Treated PNC as a dichotomous variable (presence vs. absence of PNC).
Tasnim et al. (2005)	Prospective cohort	Failed to use Kotelchuck, Kessner or comparable index: Counted number of visits. Conducted outside US.
Reichman and Teitler (2005)	Retrospective cohort	Failed to use Kotelchuck, Kessner or comparable index: Analyzed by trimester of PNC initiation.
Sánchez-Nuncio et al. (2005)	Case-control	Article in Spanish. Conducted outside US.
Laditka et al. (2005)	Retrospective cohort	Failed to report on primary outcome measures (LBW + preterm birth).

PNC prenatal care; *LBW* low birth weight; *CPSP* Comprehensive Perinatal Services Program; *CTAPPP* Comprehensive Teenage Pregnancy and Parenting Program

Table 8.5 Quality rating of PNC utilization and enhanced care studies included in this review

Primary author (year)	Study design	Reporting (max. 13)	External validity (max. 4)	Internal validity – bias (max. 7)	Internal validity – confounding (max. 6)	Power (max. 2)	Total quality score (max. 32)
Utilization studies							
Collins (1997)	Case-control	11	4	6	4	0	25
Krueger (2000)	Cohort	11	3	7	4	0	25
Enhanced studies							
Kitzman (1997)	RCT	12	3	6	6	2	29
Moore (1998)	RCT	12	3	6	5	1	27
Klerman (2001)	RCT	11	3	5	6	1	26
Brooten (2001)	RCT	10	4	5	6	0	25
Little (2002)	RCT	12	3	4	5	0	24

utilization and LBW was not statistically significant. The study did not examine how these associations might vary by race and ethnicity.

Enhanced Care Studies

Our search identified five randomized controlled trials of enhanced prenatal care (Brooten et al., 2001; Kitzman et al., 1997; Klerman et al., 2001; Little et al., 2002; Moore et al., 1998), since the last major review by Fiscella (1995) that met our inclusion/exclusion criteria. The characteristics and outcomes of the five included enhanced care studies are summarized in Table 8.3.

Kitzman et al. (1997) conducted a randomized controlled trial in Memphis, Tennessee to test the effect of nurse prenatal and infant home visits by nurses on birth and child health outcomes. The study enrolled 1,139 primarily African-American (92%) women who were at less than 29 weeks' gestation, had no previous live births, and met at least two sociodemographic risk characteristics (unmarried, <12 years of education, unemployed). Women with high-risk medical conditions were excluded. With computer-generated random assignment, women were randomized to an intervention group (n = 458) and a control group (n = 681). Intervention assignment was concealed from both study participants and intervention staff until recruitment was complete. Neither study participants nor those measuring main outcomes were blinded. In the control group, women received standard prenatal care plus free taxi transportation for prenatal appointments. In the treatment group, women received standard prenatal care, free transportation, and intensive nurse home visits. Nurses made an average of seven (range 0–18) home visits during pregnancy and followed a detailed visit-by-visit protocol to help women improve health related behavior. They helped women complete 24-h diet histories and track the weight gained over the course of the pregnancy in order to assess nutritional status. Nurses assessed cigarette smoking, alcohol, or illicit drug use and facilitated reduction in substance use through behavioral analysis. They also taught women how to identify the signs and symptoms of pregnancy complications, with particular attention to urinary tract infections, sexually transmitted diseases, and hypertensive disorders.

There were no significant differences in LBW, mean birthweight, preterm birth, mean gestational age, and intrauterine growth restriction between the intervention and control group. The incidence of LBW was 15% in the intervention group and 14% in the control group. The only birth outcome that differed between the two groups was the incidence of pregnancy-induced hypertension (13% vs. 20%; p = 0.009). While the intervention did not reduce LBW or related birth outcomes, it was found to reduce the number of subsequent pregnancies, close-spaced births, the use of welfare, negative beliefs about child-rearing, and criminal behavior among low-income unmarried mothers for up to 15 years after the birth of their first child (Olds et al., 1997, 1998). The study did not compare intervention effects by race and ethnicity.

Moore et al. (1998) conducted a randomized controlled trial in Winston-Salem, North Carolina to test the effect of a nursing telephone intervention on LBW and preterm births among low-income pregnant women. A total of 1,554 women receiving prenatal care in a public clinic who met study criteria were assigned randomly to intervention and control groups using a computer-generated randomization table, with a final analysis sample size of 1,433. Another 1,573 eligible women were not included in the study because they either refused or could not be contacted by telephone; their characteristics were not reported, raising questions about the generalizability of the study. Women in the intervention group received telephone calls from three registered nurses. Three calls were attempted weekly from 24 weeks' through 37 weeks' gestation, but only half of the calls (approximately 1.5 per week) were completed. Although no formal script was followed, each telephone call addressed three major areas: assessment of health status (perception of uterine contractions and other pregnancy changes, color of urine as an assessment of hydration, number of meals eaten, number of cigarettes smoked, alcohol and drug use, and ingestion of a prenatal vitamin capsule on the previous day); recommendations based on assessment; and, discussion of any additional issues important to the mother. Clinical personnel, including physicians, residents, and nurses, were blinded to group assignment during the study period. The nurse collecting data on the main outcomes was also blinded to group assignment.

LBW rates were 10.9% in the intervention group and 14.0% in the control group (RR 0.75; 95% CI 0.55, 1.03). Preterm birth rates were 9.7% in the intervention group and 11.0% in the control group (RR 0.87; 95% CI 0.62, 1.22). Neither main study analysis reached statistical significance. However, differences in the rates of LBW and preterm birth bordered on statistical significance for African-American women. A closer examination found the intervention to be

effective for a subgroup of African-American women aged 19 years and older. In this subgroup, LBW rates were 11.4% in the intervention group and 17.3% in the control group (RR 0.66; 95% CI 0.46, 0.94) and preterm birth rates were 8.7% in the intervention group and 15.4% in the control group (RR 0.56; 95% CI 0.38, 0.84). The authors attributed the intervention effect to enhanced education and support for a subgroup of women (African-American women aged 19 years and older) who often do not receive the same level of family and community support afforded to younger pregnant teens.

Klerman et al. (2001) conducted a randomized controlled trial in Jefferson County, Alabama, to test the effect of augmented prenatal care among high-risk African-American women on pregnancy outcomes as well as patients' knowledge of risks, satisfaction with care, and behavior. A total of 619 women (n = 318 in augmented care, n = 301 in regular care) were enrolled in the study. Nearly 8% of eligible women refused participation. The women enrolled were African-American, aged 16 years or older, and eligible for Medicaid. They had scored 10 or higher on a risk assessment scale but had no major medical complications. Augmented care was provided by a multidisciplinary team including an obstetrician, trained nurse practitioners, social workers, and behavioral medicine specialists and included educationally oriented peer groups, additional appointments, extended time with clinicians, and other supports. The control group received standard prenatal care from the county health department or the university's obstetric department. On-site child-care was provided, evening hours were available, and transportation was provided. Structured postpartum interviews were administered by interviewers blinded to the treatment group. Data were also gathered from clinic records, special forms prepared for the study, and a computerized database on Medicaid patients. Blinding of data collectors was not reported.

There were no significant differences in LBW, VLBW, mean birthweight, preterm birth, mean gestational age, IUGR, or any measured pregnancy outcomes between groups. LBW rates were 12.5% in the intervention group and 11.2% in the control group (p = 0.60). Both groups had lower than predicted rates of LBW. Preterm birth rates were 10.6% in the intervention group and 14.0% in the control group (p = 0.22). The authors concluded that high-quality augmented prenatal care that emphasized education, health promotion, and social support significantly increased women's satisfaction, knowledge of risk conditions, and perceived mastery in their lives; however, it did not reduce LBW.

Brooten et al. (2001) conducted a randomized controlled trial in Philadelphia, Pennsylvania to test the effect of prenatal nurse home visits on maternal and child health outcomes. A sample of 173 women (and 194 infants) with high-risk pregnancies (gestational or pregestational diabetes mellitus, chronic hypertension, preterm labor, or high risk of preterm labor) were enrolled in the study, of which approximately 94% (162 of 173) were African-American. The subjects were randomly assigned to the intervention group (85 women and 94 infants) or the control group (88 women and 100 infants) using a table of random numbers. Intervention assignment was concealed from both study participants and intervention staff until recruitment was complete. Women in the control group received standard prenatal care. Women in the intervention group received half of their prenatal care in their homes, in addition to education, counseling, telephone outreach, daily telephone availability, and a postpartum home visit. Blinding was not described in the paper.

LBW rates were 34% in the intervention group and 36% in the control group (RR = 0.95; 95% CI 0.65, 1.40). Preterm birth rates were 31% in the intervention group and 41% in the control group (RR = 0.76; 95% CI 0.1, 1.11). Neither main study analysis reached statistical significance. Mean birthweight among preterm infants was approximately 300 g greater in the intervention group (2,263.5 g ±711.0) compared to the control group (1,960 g ±748) (p < 0.05). Mean birthweight among term infants and mean gestational age did not differ significantly between the two groups. A large intervention effect was found among twins gestations; 4 of 18 twin gestations in

the intervention group (22%) and 16 of 24 twin gestations in the control group (67%) delivered preterm ($p < 0.05$). Mean birthweight was approximately 320 g greater and mean gestational age was 2.6 weeks greater among twin gestations in the intervention group compared to the control group. The intervention group also had fewer fetal/infant deaths among all infants (2 vs. 9; $p < 0.01$). Finally, the study reported preventing more than 750 total hospital days and saving $2,496,145 in hospital costs.

Little et al. (2002) conducted a randomized controlled trial in Minneapolis, Minnesuta to test the effects of telephonic nursing care on birth outcomes (mean gestational age and mean birth weight) and clinical resource utilization among low-income high-risk pregnant women. A total of 111 high-risk pregnant women who obtained prenatal care from two large obstetric clinics were enrolled in the study and randomly assigned to the case management group or the control group. Randomization was conducted by the study administrative assistant. No blinding by clinical personnel or data collectors was described. There were no significant differences between treatment and control groups; however, the treatment group had a larger proportion of patients with anemia, obesity, symptoms of preterm labor, and undiagnosed vaginal bleeding in pregnancy. The control group had a larger number of patients who reported problems with substance abuse. Another 64 women eligible for the study were eliminated because they could not be contacted, had a miscarriage, or refused to participate. Compared with participants, non-participants were more likely to be multiparous, single and white with less than a college-level education.

Nurse case managers contacted women in the intervention group every 7–14 days to assess their pregnancy status and offer support and teaching related to their pregnancy and diagnoses. The treatment group participants were encouraged to maintain good prenatal care and educated in the signs and symptoms of preterm labor, the importance of hydration, and the self-monitoring of fetal movement. Nurse case managers contacted the patients' health care providers as appropriate. A final contact was made after delivery to obtain delivery information and complete the postpartum mother/infant assessment. The control group completed the initial pregnancy risk screening and the postpartum mother/infant assessment.

A multiple analysis of variance with covariates was performed to examine the effect of the nursing telephone care on birth weight and gestational age, controlling for maternal obesity, maternal age, NICU admission, study group (treatment vs. control), gestational age at referral and number of preterm births. There was no effect of the intervention on preterm births; the mean gestational age at delivery was not significantly different between groups. After controlling for confounders, the study found a positive correlation between telephonic nurse case management and mean birthweight. Subgroup analysis by age and race-ethnicity was not performed due to small sample size.

Taken together, these studies overall found equivocal effects of enhanced prenatal care on LBW or preterm birth rates. However, some benefits in specific subgroups [e.g., for twin gestations in Brooten et al. (2001) or African-American women aged 19 years and older in Moore et al. (1998)] were noted.

Discussion

Consistent with the last major review, our review does not support the conclusion of the 1985 IOM report that "the overwhelming weight of the evidence is that prenatal care reduces low birthweight" (IOM, 1985b, pp. 18–19). Our review also suggests that prenatal care as currently delivered or in an enhanced form of the type discussed here is not effective in reducing racial-ethnic disparities in LBW.

Utilization Studies

Collins et al. (1997) found that race-specific term (gestational age >37 weeks) LBW rates declined as prenatal care usage rose. However, since the study only stratified on residential income within race-ethnic strata, it may not have adequately controlled for critical confounders and potential selection bias. Moreover, the study found that increased prenatal care utilization did not reduce disparities in the occurrence of LBW for blacks relative to whites and Hispanics. Krueger and Scholl (2000) did not find a significant association between prenatal care utilization and LBW; however, they did find inadequate prenatal care to be associated with increased risk for preterm birth. Their study controlled for a number of potential confounders, but several methodological limitations, including recruitment, exclusions, and combining categories of prenatal care utilization, raise concerns about the external and internal validity of the findings. Importantly, they did not examine the interaction between race/ethnicity and prenatal care utilization with respect to LBW or preterm birth.

Our review does not change the conclusion made in the last major review. Fiscella (1995) reviewed 14 observational studies on the association between prenatal care utilization and birth outcomes. Of the eight observational studies reporting LBW as an outcome, four found a significant benefit from prenatal care, while four did not. Although two studies found adequate prenatal care to be associated with greater reduction in the odds of LBW among African-American than among whites, Fiscella was critical of these studies for their failure to adequately control for potential confounding. He also raised concerns that while the Kessner, Kotelchuck, and similar indices were designed to minimize gestational age bias, they do not eliminate this bias altogether. Using Bradford-Hills criteria for evaluating evidence of a causal relationship, he concluded that the current evidence did not satisfy such criteria. We do not find evidence from the two included studies published subsequent to his review sufficiently strong to reverse this conclusion.

Enhanced Care Studies

None of the five randomized controlled trials was able to demonstrate a main effect of enhanced prenatal care in preventing LBW, though some studies suggested there may be specific subgroups that might benefit from enhanced prenatal care or that outcomes such as mean birthweight may be affected. The interventions included telephone calls, nurse home visits, and comprehensive prenatal care. Our findings are consistent with those of previous reviews. Fiscella (1995) reviewed 11 randomized controlled trials published between 1966 and 1994; none of the trials of enhanced care in his review showed positive main effects. Hueston, Knox, Eilers, Pauwels, and Lonsdorf (1995) also reviewed six randomized controlled trials of preterm birth prevention educational programs among high-risk women; using meta-analytic techniques, no significant benefits were found for preterm birth education programs in preventing neonatal death, LBW, or preterm birth. Hodnett and Fredericks (2003) conducted a Cochrane review of social support during pregnancy. Sixteen trials involving 13,651 women at-risk for preterm birth or LBW were included. Interventions included emotional or instrumental support, provided by professional or trained lay person, in-home or in clinical settings. Programs offering additional social support for at-risk pregnant women were not associated with improvements in any perinatal outcomes, including LBW and preterm birth. To date, available evidence does not support the effectiveness of enhanced prenatal care, in the forms of telephone calls, home visits, preterm birth education, comprehensive care, or social support, in preventing LBW and preterm birth for most populations although other benefits may be evident.

Racial-Ethnic Disparities in Birth Outcomes

A primary aim of this review is to examine the effectiveness of prenatal care in reducing racial-ethnic disparities in LBW and related birth outcomes. Of the two included utilization studies, only Collins et al. (1997) examined the association between prenatal care utilization and LBW by race-ethnicity. Collins et al. found that although race-specific term LBW rates declined as prenatal care usage rose, the position of African-Americans relative to Mexican-Americans and whites was essentially unchanged, raising serious doubts as to whether increasing access to and utilization of prenatal care in its current form can help reduce racial-ethnic disparities in LBW and related outcomes.

With respect to the randomized controlled trials of enhanced prenatal care, three studies enrolled predominantly African-American women. Kitzman et al. (1997) and Klerman et al. (2001) recruited sociodemographically at-risk African-American women, whereas Brooten et al. (2001) recruited medically at-risk African-American women. Since none of these three trials found a main effect of enhanced prenatal care on LBW and related outcomes, whether enhanced prenatal care (i.e., nurse home visits or comprehensive care model) can reduce racial-ethnic disparities in LBW and preterm birth remains questionable. Moore et al. (1998) found an intervention effect in a subgroup of African-American women 19 years and older. In this subgroup, a nursing intervention by telephone call reduced LBW rates by 34% and preterm birth rates by 44%. While these results appear promising, methodological concerns have been raised including the method of random allocation and the high rate of loss-to-follow-up. Furthermore, it remains unclear *why* the program succeeded when other trials involving more intensive nursing interventions failed. The authors argued that the program provided education and support to a subgroup with the greatest unmet needs for education and support, but offered no data to support this claim, such as a change in knowledge or perceived support pre- and post-intervention. Thus presently there is no conclusive evidence that enhanced prenatal care can reduce racial-ethnic disparities in LBW and related birth outcomes.

The Challenges of Studying the Effectiveness of Prenatal Care

The inconclusiveness of the evidence reviewed in this chapter reflects in part the challenges of studying the effectiveness of prenatal care. As discussed earlier, there are three major types of studies on the effectiveness of prenatal care: ecological studies, utilization studies and enhanced care studies. Ecological studies correlate prenatal care utilization to birth outcomes *at a population level*. For example, the study might show that following expansion of Medicaid eligibility in a state, there was an increase in the proportion of women who started prenatal care in the first trimester concomitant to a decline in LBW rate in the state. However, one cannot tell from population-level data correlating a rise in prenatal care utilization to a decline in LBW rate whether prenatal care is associated with favorable outcomes at the individual level.

The major challenge to the validity of utilization studies is the potential for selection bias. There are at least four different types of selection bias (Bell & Zimmerman, 2003). *Favorable selection* (low risk/high use) results when healthy women at low risk for poor outcomes are more likely to receive early and adequate prenatal care; such selection may overestimate the measured effects of prenatal care on birth outcomes. *Adverse selection* (high risk/high use) results when women at high risk for adverse outcomes are more likely to seek early and intensive prenatal care because of a preexisting medical condition, prior experience or family history; such selection may underestimate the measured effects of prenatal care on birth outcomes. *Estrangement selection* (high risk/low use) results when women at risk for adverse outcomes are more likely to receive inadequate or no prenatal care because of life circumstances such as homelessness, substance abuse or intimate partner

violence; such selection may overestimate prenatal care efficacy. Finally, *confidence selection* (low risk/low use) results when healthy women at low risk for adverse outcomes are less likely to use prenatal care because their general health or prior experience leads them to believe that they will have a healthy birth outcome with or without care; such selection may underestimate the measured effects of prenatal care on birth outcomes. It is possible that all four types of selection bias, operating in different directions, may be in play in most utilization studies.

Randomized controlled trials are typically considered the gold standard of study designs; however, randomizing women to receiving prenatal care vs. no prenatal care is neither feasible nor ethical. Enhanced care studies address a different question; instead of evaluating the effectiveness of prenatal care, these studies evaluate the effectiveness of added components of care such as health education or home visitation. Unfortunately, none of the 16 randomized controlled trials [11 reviewed by Fiscella, 1995 (Table 8.3) and five additional RCTs reviewed in this chapter] found a main effect of these added components on birth outcomes. Furthermore, in a review of psychosocial interventions to prevent LBW, Lu, Lu, and Schetter (2005) concluded that most such interventions were not driven by theory, did not have effective risk screening, did not match intervention to risk, and did not test process variables. Thus it is difficult to determine from these studies whether the failure of enhanced care in preventing LBW is due to ineffective interventions, poor study design, or both.

Most importantly, the effectiveness of prenatal care may well depend on how prenatal care is defined. In most studies, prenatal care is defined as a series of clinical visits based on a schedule recommended by ACOG and AAP (2007): "Generally, a woman with an uncomplicated pregnancy is examined every 4 weeks for the first 28 weeks of pregnancy, every 2–3 weeks until 36 weeks of gestation, and weekly thereafter" (p. 100). This schedule, which has been used to define the adequacy of prenatal care, was designed largely for early detection of preeclampsia and other pregnancy complications rather than for prevention of LBW or preterm birth. It is, therefore, not surprising that the adequacy of prenatal care is not definitively associated with LBW or preterm birth. In most utilization studies, only the timing and quantity of prenatal visits is considered; few studies have evaluated birth outcomes in relation to the content, quality, and mode of delivery of prenatal care. The effectiveness of prenatal care may also depend on the outcomes being studied. Prenatal care may not have been shown to be effective in preventing LBW for the index pregnancy, but little is known about its impact on a subsequent pregnancy or its long-term impact on the health and behaviors of the mother, child, and family. For example, less than adequate prenatal care has been associated with significantly fewer well-child visits and incomplete immunizations (Kogan, Alexander, Jack, & Allen, 1998). We caution against over-interpretation of our findings as a rejection of the importance of prenatal care; rather, our findings merely demonstrated the inconclusiveness of the evidence for its effectiveness in preventing LBW and potentially reducing disparities in LBW.

Rethinking Prenatal Care

More than two decades following the IOM report, the effectiveness of prenatal care for preventing LBW or reducing racial-ethnic disparities in LBW remains unproven. We suggest that some rethinking about the *content*, *timing*, and *context* of prenatal care is needed.

The Content of Prenatal Care. As several reviews had previously concluded, prenatal care in its present form is unlikely to reduce LBW because it does not address the underlying causes of LBW (Alexander & Kotelchuck, 2001; Lu et al., 2003). Recent advances in biomedical and social-behavioral research have improved our understanding of the etiologic mechanisms leading to LBW. Could the *content* of prenatal care be redesigned to address more effectively the underlying causes of LBW? For example, while the multiple pathways leading to PTD have not been clearly

elucidated, a growing body of evidence implicates: (1) activation of the maternal or fetal hypothalamic-pituitary-adrenal axis; (2) decidual-chorioamniotic or systemic inflammation; (3) decidual hemorrhage (i.e., abruption); and, (4) pathological distention of the uterus in the pathogenesis of PTD (Lockwood, 2003). Similarly, recent advances in research are beginning to elucidate some of the complex pathophysiological processes leading to IUGR, including the interaction between immunology and human placental implantation, the control and function of growth factors such as insulin-like growth factor and its binding proteins, and vasoactive agents such as prostacyclin, thromboxane A_2, endothelin-1, and nitric oxide, and genetic mutations.

Given these known pathways to preterm birth and IUGR, it is perhaps not surprising that prenatal care in its present form is ineffective in preventing LBW. The content of prenatal care, as recommended by current AAP-ACOG guidelines, was not designed to address these underlying mechanisms of LBW. For example, it is quite unlikely that checking blood pressure and urine protein, designed for early detection of preeclampsia, does anything to reverse premature activation of the maternal or fetal hypothalamic-pituitary-adrenal axis or decidual-chorioamniotic or systemic inflammation. Measuring fundal height, designed to screen for IUGR, has high interobserver variance and poor predictive values. Moreover, available interventions, such as bedrest or antenatal testing, may do little to improve placental blood flow that has been compromised by thromboses, atheroses, and other placental pathologies that may have resulted from aberrant placentation in early pregnancy (Lu et al., 2003). The challenge of rethinking the content of prenatal care to address racial-ethnic disparities in LBW is even more daunting because the underlying causes of the disparities are less well understood. Given known pathways to preterm birth and IUGR, chronic stress, inflammation and nutrition probably are major contributors to the disparities; yet presently these three concerns are poorly addressed by prenatal care. To prevent LBW and reduce disparities, there needs to be some rethinking about the content of prenatal care so that it can better address the underlying causes of LBW.

The Timing of Prenatal Care. Could the *timing* of prenatal care be improved to address more effectively the underlying causes of LBW? Many of the pathophysiologic processes leading to PTD or IUGR may have their onset early in pregnancy. For example, an infection potentially responsible for PTD may already be present in the urogenital tract in early pregnancy or even before conception (Goldenberg, Hauth, & Andrews, 2000). If it is not cleared by midgestation, preterm labor, or preterm premature rupture of membranes (PPROM) may ensue. Screening for and treating bacterial vaginosis with antibiotics in midgestation, weeks or perhaps even months after its onset, or giving antibiotics after preterm labor is already in progress, may prove to be ineffective in preventing preterm birth. Perhaps this explains the disappointing results of the antibiotic trials in pregnancy (King & Flenady, 2002). Even if the infection is treated, it may be too late to arrest the immune-inflammatory processes that have long been initiated. Similarly, the "uteroplacental insufficiency" responsible for IUGR may be traced to abnormal trophoblastic invasion during implantation early in pregnancy (Khong, De Wolf, Robertson, & Brosens, 1986). Implantation, in turn, is regulated by immunologic mechanisms involving predominantly decidual natural killer cells, which secrete certain cytokines to stimulate growth, differentiation, and migration of trophoblasts (Loke & King, 1997). Immunologic dysregulation of implantation could lead the pregnancy, shortly after conception, down the pathophysiogic pathway toward IUGR which may be difficult for prenatal care to reverse. The timing of these events underscores the potential contributions of preconception and interconception care to preventing LBW. While current research has focused primarily on its benefit in preventing congenital anomalies through dietary control of pregestational diabetes mellitus or hyperphenylalaninemia or nutrition supplementation (e.g., folic acid) (Korenbrot, Steinberg, Bender, & Newberry, 2002), future research needs to investigate the effectiveness of preconception interventions in preventing PTD or IUGR, and that of interconception care in preventing their recurrence (Johnson et al., 2006). Given significant racial-ethnic disparities in healthcare access for women before and between pregnancies, increasing access to preconception and interconception

care may hold greater promise for reducing racial-ethnic disparities in LBW than prenatal care has demonstrated.

But even preconceptional care may do too little too late for preventing LBW or reducing disparities in LBW. Lu and Halfon (2003) recently proposed using a life-course perspective to reexamine racial-ethnic disparities in birth outcomes. Vulnerability to PTD or IUGR may be traced to not only risk factors before and during pregnancy, but to experiences and exposures that occur early in life and accumulate throughout the life course of the woman. A growing body of research on life course health development has suggested that the functional capacity of many organ systems begins in-utero and continues to develop over the life course. A woman's reproductive capacity is no exception. Early life experiences become embedded into her reproductive biology and may influence her future potential to conceive and carry a healthy pregnancy to term. For example, it has been shown that maternal stress is associated with higher stress reactivity in her offspring that persists well into adulthood (Hertzman, 1999; Seckl, 1998; Wadhwa, 1998), which may be related to feedback resistance as a result of decreased expression of glucocorticoid receptors in the brain during critical period of neuroendocrine development (Meaney, Aitken, Sharma, Viau, & Sarrieau, 1989). Early life exposures to stress hormones during critical periods of immune maturation may also alter immune function, leading to increased susceptibility to infectious or inflammatory diseases over the life course (Coe, 1999). Hypothetically, maternal stress could thus prime the neuroendocrine axes and immune system of her developing fetus with stress hormones, leading to higher stress reactivity and immune-inflammatory dysregulation that could increase her female offspring's vulnerability to PTD or IUGR later on in life (Lu & Halfon, 2003). This might help explain the observed intergenerational clustering of preterm birth and LBW (Emanuel, 1997).

Beyond early life, cumulative exposures to chronic stress results in wear and tear, what Bruce McEwen refers to as "allostatic load," on the body's adaptive systems (1998). Studies have found in animals and humans subjected to chronic and repeated stress, elevated basal cortisol levels and exaggerated ACTH and cortisol responses to natural or experimental stressors (Kristenson et al., 1998; Sapolsky, 1995; Sapolsky, Krey, & McEwen, 1984). This HPA hyperactivity may reflect the inability of a worn-out HPA axis for self-regulation, possibly due to the loss of feedback inhibition via down-regulation of glucocorticoid receptors in the brain (Sapolsky; Sapolsky et al.). Similarly, chronically elevated levels of cortisol may also lead to not only relative immune suppression, but also immune-inflammatory dysregulation due to the loss of counter-regulation by the HPA axis, resulting in part from down-regulation of glucocorticoid receptors in the immune cells (Chrousos, 2000). HPA hyperactivity and immune-inflammatory dysregulation are two of several possible mechanisms by which accommodation to chronic and repeated stress over the life-course may lead to increased vulnerability to PTD and IUGR during pregnancy. Evidence supporting the cumulative pathway mechanism comes from research on the weathering hypothesis (Geronimus, 1996).

From a life-course perspective, it is perhaps not surprising that the effectiveness of prenatal care for preventing LBW or reducing racial-ethnic disparities in LBW has not been conclusively demonstrated. To expect prenatal care, in less than 9 months, to reverse the impacts of early life programming and cumulative allostatic load on a woman's reproductive health may be expecting too much of prenatal care. Even preconceptional care may do too little too late if it is provided in a single visit shortly before a planned pregnancy, rather than as an integral part of women's health care continuum for all women of reproductive age. Ultimately, preventing LBW will take a fundamental reconceptualization of prenatal care as part of a longitudinally integrated strategy that promotes optimal development of women's reproductive health not only during pregnancy, but over their entire life course.

The Context of Prenatal Care. Could the *context* of prenatal care be expanded to address more effectively the multilevel, multiple determinants of racial-ethnic disparities in LBW? Presently, prenatal care is still delivered primarily through the obstetrical visit, with links to public health ancillary services such as WIC services or social support services for low-income women (Alexander &

Kotelchuck, 2001). These clinical and ancillary services, while necessary, are hardly sufficient to address the multiple causes of LBW. For example, Collins et al. (1998) found a two to threefold increase in the risk of VLBW births (most of which were preterm) among African-American women who rated their neighborhoods unfavorably in terms of police protection, protection of property, personal safety, friendliness, delivery of municipal services, cleanliness, quietness, and schools. A more recent case-control study (Collins et al., 2000) found that among low-income African-American women in Chicago, the adjusted odds of giving birth to a VLBW infant was 3.3 times greater among women who reported having experienced racial discrimination than among those who did not. A greater African-American-white gap in infant mortality has also been found in cities that are more segregated (LaVeist, 1993; Polednak, 1996). A growing body of literature also links air and water pollution to preterm birth and IUGR (Sram, Binkova, Dejmek, & Bobak, 2005). In many disadvantaged communities, there are more liquor stores than grocery stores, and more fast food restaurants than healthy restaurants. It has been shown that the typical cost of food is approximately 15–20% higher in poor neighborhoods, while the quality of food available is poorer (Emmons, 2000). For individuals growing up and living in those communities, the relative unavailability of healthy, nutritious food may pattern a lifelong habit of making unhealthy food choices that becomes difficult to change during pregnancy. Currently, little is done during the standard prenatal visit, or through its public health ancillary services, to address neighborhood factors, racial discrimination and residential segregation, air and water pollution, unavailability of healthy food choices, or other contextual determinants of LBW.

Health care providers and public health professionals are not exempt from addressing causes of health disparities outside of the clinical domain (Hogan, Njorge, Durant, & Ferre, 2001). They may not be able to solve all the problems, but it is imperative that they reach out to those who could. These may include the partner, family, and peers who could provide the pregnant woman with consistent daily support between prenatal visits. These may also include leaders of business, community or faith-based organizations who could reinforce health promotion messages outside of clinical settings. Prenatal care should not cease once the pregnant woman walks out of her doctor's office; it should continue at home, at work, in neighborhood parks and grocery stores, and in every aspect of her everyday life.

What is needed is a contextually integrated model of prenatal care. Risk assessment, health promotion, and medical and psychosocial interventions need to address causes of LBW not only at the individual level, but also at the interpersonal, institutional, community, and policy levels. A contextually integrated model of prenatal care will require cross-sectoral collaboration; health care providers and public health professionals need to engage other Maternal and Child Health (MCH) and non-MCH service providers, as well as leaders from business, civic, and faith-based sectors, in a collaborative effort to prevent LBW. It will take building stronger and healthier communities that promote not only healthy pregnancy, but the life-course health development of women and families. This will require investments in infrastructure, such as affordable and decent housing, safe neighborhood, accessible parks and recreation, clean air and water, and competent health care. These investments ought to be decided with full community participation. A contextually integrated model of PNC will also require social investments, with the goal of reducing cumulative allostatic load over the life-course of women. This requires policymakers to pay attention to issues that disproportionately impact on women's lives, such as domestic violence and child care. Men (especially fathers) play an important role, positive or negative, in the lives of women and children, and yet they are often treated as an afterthought in MCH. Current policies provide little support, and in some cases great disincentives, for male involvement in pregnancy and parenting, leaving women to bear greater burdens of childbearing and childrearing (Lu et al., 2007). The impact of social legislation (e.g., maternity leave policies, laws prohibiting employment discrimination, or safeguards for work safety and working conditions) on pregnancy and parenting also merits greater attention. Public policies and social movements to combat racism and gender inequality may be one of the most effective components of

prenatal care. As Alexander and Korenbrot observe, the "ultimate success of prenatal care in reducing current low birth weight percentages in the United States may hinge on the development of a much broader and more unified conception of prenatal care than currently prevails" (1995, p. 114).

Conclusion

Our review should not be interpreted as a rejection of prenatal care, which may benefit pregnancy outcomes other than LBW, such as reduced maternal, fetal, and infant morbidities and mortality, or improved maternal health status and parenting behaviors (e.g., well-baby care or vaccinations) (Grimes, 1994; Kogan et al., 1998). From a life-course perspective, the benefits of prenatal care may accrue over the maternal life course, from one pregnancy to the next or even across generations, rather than in immediate birth outcomes. Our specific aims were to review the evidence of effectiveness of prenatal care for preventing LBW or reducing racial-ethnic disparities in LBW in the current pregnancy, and to catalyze some rethinking about its content, timing, and delivery. We conclude that preventing LBW and reducing racial-ethnic disparities will take much more than prenatal care in its present form; it will require a fundamental reconceptualization of prenatal care as part of longitudinally and contextually integrated strategy to promote optimal development of women's reproductive health not only during pregnancy, but over the life course.

References

Alexander, G. R., & Korenbrot, G. (1995). The role of prenatal care in preventing low birth weight. *Future of Children*, *5*, 103–120.

Alexander, G. R., & Kotelchuck, M. (1996). Quantifying the adequacy of prenatal care: A comparison of indices. *Public Health Reports, 111*, 408–409.

Alexander, G. R., & Kotelchuck, M. (2001). Assessing the role and effectiveness of prenatal care: History, challenges, and directions for future research. *Public Health Reports*, *116*, 306–316.

Alexander, G. R., & Slay, M. (2002). Prematurity at birth: Trends, racial disparities, and epidemiology. *Mental Retardation and Developmental Disabilities Research Reviews*, *8*, 215–220.

American Academy of Pediatrics & The American College of Obstetricians and Gynecologists. (2007). *Guidelines for perinatal care* (6th ed.). Washington, DC: Author.

Armson, B. A., Dodds, L., Haliburton, S. C., Cervin, C., & Rinaldo, K. (2003). Impact of participation in the Halifax County Preterm Birth Prevention Project. *Journal of Obstetrics and Gynecology Canada*, *25*(3), 209–217.

Barnet, B., Duggan, A. K., & Devoe, M. (2003). Reduced low birth weight for teenagers receiving prenatal care at a school-based health center: Effect of access and comprehensive care. *Journal of Adolescent Health*, *33*(5), 349–358.

Barros, H., Tavares, M., & Rodrigues, T. (1996). Role of prenatal care in preterm birth and low birthweight in Portugal. *Journal of Public Health Medicine*, *18*(3), 321–328.

Bell, J. F., & Zimmerman, F. J. (2003). Selection bias in prenatal care use by Medicaid recipients. *Maternal and Child Health Journal*, *7*, 239–252.

Binstock, M. A., & Wolde-Tsadik, G. (1995). Alternative prenatal care. Impact of reduced visit frequency, focused visits and continuity of care. *Journal of Reproductive Medicine*, *40*(7), 507–512.

Blanchette, H. (1995). Comparison of obstetric outcome of a primary-care access clinic staffed by certified nurse-midwives and a private practice group of obstetricians in the same community. *American Journal of Obstetrics and Gynecology*, *172*(6), 1864–1868.

Boss, D. J., & Timbrook, R. E. (2001). Clinical obstetric outcomes related to continuity in prenatal care. *Journal of the American Board of Family Practice*, *14*(6), 418–423.

Brooten, D., Youngblut, J. M., Brown, L., Finkler, S. A., Neff, D. F., & Madigan, E. (2001). A randomized trial of nurse specialist home care for women with high-risk pregnancies: Outcomes and costs. *American Journal of Managed Care*, *7*(8), 793–803.

Bryce, R. L., Stanley, F. J., & Garner, J. B. (1991). Randomized controlled trial of antenatal social support to prevent preterm birth. *British Journal of Obstetrics and Gynecology*, *98*, 1001–1008.

Chrousos, G. P. (2000). Stress response and immune function: Clinical implications. *Annals of the New York Academy of Sciences*, *917*, 38–67.

Coe, C. L. (1999). Psychosocial factors and psychoneuroimmunology within a lifespan perspective. In D. P. Keating & C. Hertzman (Eds.), *Developmental health and the wealth of nations: Social, biological and educational dynamics* (pp. 201–219). New York: Guilford.

Collaborative Group on Preterm Birth Prevention. (1993). Multicenter randomized, controlled trial of a preterm birth prevention program. *American Journal of Obstetrics and Gynecology*, *169*, 352–366.

Collins, J. W. J., David, R. J., Symons, R., Handler, A., Wall, S., & Andes, S. (1998). African-American mothers' perception of their residential environment, stressful life events, and very low birthweight. *Epidemiology*, *9*, 286–289.

Collins, J. W. J., David, R. J., Symons, R., Handler, A., Wall, S. N., & Dwyer, L. (2000). Low-income African-American mothers' perception of exposure to racial discrimination and infant birth weight. *Epidemiology*, *11*, 337–339.

Collins, J. W. J., Herman, S. N., & David, R. J. (1997). Adequacy of prenatal care utilization, maternal ethnicity, and infant birthweight in Chicago. *Journal of the National Medical Association*, *89*(3), 198–203.

Dyson, D. C., Danbe, K. H., Bamber, J. A., Crites, Y. M., Field, D. R., Maier, J. A., et al. (1998). Monitoring women at risk for preterm labor. *The New England Journal of Medicine*, *338*(1), 15–19.

Eastman, N. J. (1947). Prematurity from the viewpoint of the obstetrician. *American Practitioner*, *I*(7), 343–352.

Edwards, C. F., Chazotte, C., Freda, M. C., Shah, L., Girz, B., Damus, K., et al. (1995). Impact of an inner-city, hospital-based preterm prevention program on preterm births in twin gestation. *Journal of the Association for Academic Minority Physicians*, *6*(2), 78–81.

Eisner, V., Brazie, J. V., Pratt, M. W., & Hexter, A. C. (1979). The risk of low birthweight. *American Journal of Public Health*, *69*, 887–893.

Emanuel, I. (1997). Invited commentary: An assessment of maternal intergenerational factors in pregnancy outcome. *American Journal of Epidemiology*, *146*, 820–825.

Emmons, K. M. (2000). Health behaviors in a social context. In L. F. Berkman & I. Kawachi (Eds.), *Social epidemiology* (pp. 242–266). Oxford: Oxford University Press.

Fiscella, K. (1995). Does prenatal care improve birth outcomes? A critical review. *Obstetrics and Gynecology*, *85*(3), 468–479.

Geronimus, A. T. (1996). Black/white differences in the relationship of maternal age to birthweight: A population-based test of the weathering hypothesis. *Social Science & Medicine*, *42*, 589–597.

Goldenberg, R. L., Davis, R. O., Copper, R. L., Corliss, D. K., Andrews, J. B., & Carpenter, A. H. (1990). The Alabama preterm birth prevention project. *Obstetrics and Gynecology*, *5*, 933–939.

Goldenberg, R. L., Hauth, J. C., & Andrews, W. W. (2000). Intrauterine infection and preterm delivery. *New England Journal of Medicine*, *342*, 1500–1507.

Gómez-Olmedo, M., Delgado-Rodriguez, M., Bueno-Cavanillas, A., Molina-Font, J. A., & Gálvez-Vargas, R. (1996). Prenatal care and prevention of preterm birth. A case-control study in southern Spain. *European Journal of Epidemiology*, *12*(1), 37–44.

Gortmaker, S. L. (1979). The effects of prenatal care upon the health of the newborn. *American Journal of Public Health*, *69*, 653–660.

Graham, A. V., Frank, S. H., Zyzanski, S. J., Kitson, G. C., & Reeb, K. G. (1992). A clinical trial to reduce the rate of low birth weight in an inner-city black population. *Family Medicine*, *24*, 439–446.

Greenberg, R. S. (1983). The impact of prenatal care in different social groups. *American Journal of Obstetrics and Gynecology*, *145*, 797–801.

Grimes, D. A. (1994). The morbidity and mortality of pregnancy: Still risky business. *American Journal of Obstetrics and Gynecology*,*170*, 1489–1494.

Heins, H. C. Jr., Nance, N. W., McCathy, B. J., & Efird, C. M. (1990). A randomized trial of nurse-midwifery prenatal care to reduce low birth weight. *Obstetrics and Gynecology*, *75*, 341–345.

Helfand, M., & Zimmer-Gembeck, M. J. (1997). Practice variation and the risk of low birth weight in a public prenatal care program. *Medical Care*, *35*(1), 16–31.

Herman, A. A., Berendes, H. W., Yu, K. F., Cooper, L. C., Overpeck, M. D., Rhoads, G., et al. (1996). Evaluation of the effectiveness of a community-based enriched model prenatal intervention project in the District of Columbia. *Health Services Research*, *31*(5), 609–621.

Hertzman, C. (1999). The biological embedding of early experience and its effects on health in adulthood. *Annals of the New York Academy of Sciences*, *896*, 85–95.

Hill, I. T. (1992). The role of Medicaid and other government programs in providing medical care for children and pregnant women. *Future of Children*, *2*(2), 134–153.

Hodnett, E. D., & Fredericks, S. (2003). Support during pregnancy for women at increased risk of low birthweight babies. *Cochrane Database of Systematic Reviews*, (3), CD000198.

Hogan, V. K., Njorge, T., Durant, T. M., & Ferre, C. D. (2001). Eliminating disparities in perinatal outcomes – Lessons learned. *Maternal and Child Health Journal*, *5*, 135–140.

Homan, R. K., & Korenbrot, C. C. (1998). Explaining variation in birth outcomes of Medicaid-eligible women with variation in the adequacy of prenatal support services. *Medical Care*, *36*(2), 190–201.

Hueston, W. J. (1995). Prenatal care and low-birth-weight rates in urban and rural Wisconsin. *Journal of the American Board of Family Practice*, *8*(1), 17–21.

Hueston, W. J., Gilbert, G. E., Davis, L., & Sturgill, V. (2003). Delayed prenatal care and the risk of low birth weight delivery. *Journal of Community Health*, *28*(3), 199–208.

Hueston, W. J., Knox, M. A., Eilers, G., Pauwels, J., & Lonsdorf, D. (1995). The effectiveness of preterm-birth prevention educational programs for high-risk women: A meta-analysis. *Obstetrics and Gynecology*, *86*, 705–712.

Huntington, J., & Connell, F. A. (1994). For every dollar spent – The cost-savings argument for prenatal care. *New England Journal of Medicine*, *331*, 1303–1307.

Ickovics, J. R., Kershaw, T. S., Westdahl, C., Rising, S. S., Klima, C., Reynolds, H., et al. (2003). Group prenatal care and preterm birth weight: Results from a matched cohort study at public clinics. *Obstetrics and Gynecology*, *102*(5 Pt 1), 1051–1057.

Institute of Medicine. Committee to Study the Prevention of Low Birth Weight. (1985a). *Preventing low birth weight*. Washington, DC: National Academy Press.

Institute of Medicine. Committee to Study the Prevention of Low Birth Weight. (1985b). *Preventing low birth weight*. Summary. Washington, DC: National Academy Press.

Jackson, D. J., Lang, J. M., Swartz, W. H., Ganiats, T. G., Fullerton, J., Ecker, J., et al. (2003). Outcomes, safety, and resource utilization in a collaborative care birth center program compared with traditional physician-based perinatal care. *American Journal of Public Health*, *93*(6), 999–1006.

Johnson, K., Posner, S. F., Biermann, J., Cordero, J. F., Atrash, H. K., Parker, C. S., et al. (2006). Recommendations to improve preconception health and health care – United States. A report of the CDC/ATSDR Preconception Care Work Group and the Select Panel on Preconception Care. *Morbidity and Mortality Weekly Report. Recommendations and Reports*, *55*(RR-6), 1–23.

Kessner, D. M., Singer, J., Kalk, C. W., & Schlesinger, E. R. (1973). Infant death: An analysis by maternal risk and health care. In Institute of Medicine (Ed.), *Contrasts in health status* (Vol. I). Washington, DC: National Academy of Sciences.

Khong, T. Y., De Wolf, F., Robertson, W. B., & Brosens, I. (1986). Inadequate maternal vascular response to placentation in pregnancies complicated by preeclampsia and by small-for-gestational-age infants. *British Journal of Obstetrics and Gynaecology*, *93*, 1049–1059.

King, J., & Flenady, V. (2002). Antibiotics for preterm labour with intact membranes. *Cochrane Database of Systematic Reviews*, (2), CD000246.

Kitzman, H., Olds, D. L., Henderson, C. R. Jr., Hanks, C., Cole, R., Tatelbaum, R., et al. (1997). Effect of prenatal and infancy home visitation by nurses on pregnancy outcomes, childhood injuries, and repeated childbearing. A randomized controlled trial. *JAMA: The Journal of the American Medical Association*, *278*(8), 644.

Klerman, L. V., Ramey, S. L., Goldenberg, R. L., Marbury, S., Hou, J., & Cliver, S. P. (2001). A randomized trial of augmented prenatal care for multiple-risk, Medicaid-eligible African-American women. *American Journal of Public Health*, *91*(1), 105–111.

Kogan, M. D., Alexander, G. R., Jack, B. W., & Allen, M. C. (1998). The association between adequacy of prenatal care utilization and subsequent pediatric care utilization in the United States. *Pediatrics*, *102*, 25–30.

Kogan, M. D., Alexander, G. R., Kotelchuck, M., & Nagey, D. A. (1994). Relation of the content of prenatal care to the risk of low birth weight. *The Journal of the American Medical Association*, *271*, 1340–1345.

Kogan, M. D., Martin, J. A., Alexander, G. R., Kotelchuck, M., Ventura, S. S., & Frigoletto, F. D. (1998). The changing pattern of prenatal care utilization in the United States, 1981–1995, using different prenatal care indices. *The Journal of the American Medical Association*, *279*, 1623–1628.

Korenbrot, C. C., Steinberg, A., Bender, C., & Newberry, S. (2002). Preconception care: A systematic review. *Maternal and Child Health Journal*, *6*(2), 75–88.

Kotelchuck, M. (1994). An evaluation of the Kessner Adequacy of Prenatal Care Index and a proposed Adequacy of Prenatal Care Utilization Index. *American Journal of Public Health*, *84*, 1414–1420.

Kristenson, M., Kucinskien, Z., Bergdahl, B., Calkauskas, H., Urmonas, V., & Orth-Gomer, K. (1998). Increased psychosocial strain in Lithuanian versus Swedish men: The LiVicorida study. *Psychosomatic Medicine*, *60*, 277–282.

Krueger, P. M., & Scholl, T. O. (2000). Adequacy of prenatal care and pregnancy outcome. *Journal of the American Osteopathic Association*, *100*(8), 485–492.

Laditka, S. B., Laditka, J. N., Mastanduno, M. P., Lauria, M. R., & Foster, T. C. (2005). Potentially avoidable maternity complications: An indicator of access to prenatal and primary care during pregnancy. *Women & Health*, *41*(3), 1–26.

LaVeist, T. A. (1993). Segregation, poverty, and empowerment: Health consequences for African-Americans. *Milbank Quarterly*, *71*, 41–64.

Lazariu-Bauer, V., Stratton, H., Pruzek, R., & Woelfel, M. L. (2004). A comparative analysis of effects of early versus late prenatal WIC participation on birth weight: NYS, 1995. *Maternal and Child Health Journal*, *8*(2), 77–86.

Little, M., Saul, G. D., Testa, K., & Gaziano, C. (2002). Improving pregnancy outcome and reducing avoidable clinical resource utilization through telephonic perinatal care coordination. *Lippincott's Case Management*, *7*(3), 103–112.

Lockwood, C. J. (2003). Testing for risk of preterm delivery. *Clinical Laboratory Medicine*, *23*, 345–360.

Loke, Y. W., & King, A. (1997). Immunology of human placental implantation: Clinical implications of our current understanding. *Molecular Medicine Today*, *4*, 153–159.

Lu, M. C., & Halfon, N. (2003). Racial and ethnic disparities in birth outcomes: A life-course perspective. *Maternal and Child Health Journal*, *7*(1), 13–30.

Lu, M. C., Jones, L., Bond, M., Pumpuang, M., Maidenberg, M., Jones, D., et al. (2007). Where is the F in MCH: Father involvement in African-American families. *Ethnicity and Disease*, *20*(1 Suppl 2), S2–S49–S61.

Lu, Q., Lu, M. C., & Schetter, C. D. (2005). Learning from success and failure in psychosocial intervention: An evaluation of low birth weight prevention trials. *Journal of Health Psychology*, *10*, 185–195.

Lu, M. C., Tache, V., Alexander, G. R., Kotelchuck, M., & Halfon, N. (2003). Preventing low birth weight: Is prenatal care the answer? *Journal of Maternal-Fetal & Neonatal Medicine*, *13*, 362–380.

Main, D. M., Gabbe, S. G., Richardson, D., & Strong, S. (1985). Can preterm deliveries be prevented? *American Journal of Obstetrics and Gynecology*, *151*, 892–898.

Main, D. M., Richardson, D. K., Hadley, C. B., & Gabbe, S. G. (1989). Controlled trial of a preterm labor detection program: Efficacy and costs. *Obstetrics and Gynecology*, *74*, 873–877.

Malloy, M. H., Kao, T. C., & Lee, Y. J. (1992). Analyzing the effect of prenatal care on pregnancy outcome: A conditional approach. *American Journal of Public Health*, *82*, 448–450.

Martin, J. A., Hamilton, B. E., Sutton, P. D., Ventura, S. J., Menacker, F., Kirmeyer, S., et al. (2007). Births: Final data for 2005. *National Vital Statistics Reports*, *56*(6), 1–103.

Martin, J. A., Hamilton, B. E., Ventura, S. J., Menacher, F., Park, M. M., & Sutton, P. D. (2002). Births: Final data for 2001. *National Vital Statistics Reports*, *51*(2), 1–102.

McDuffie, R. S., Jr., Beck, A., Bischoff, K., Cross, J., & Orleans, M. (1996). Effect of frequency of prenatal care visits on perinatal outcome among low-risk women. A randomized controlled trial. *The Journal of the American Medical Association*, *275*(11), 847–851.

McEwen, B. S. (1998). Protective and damaging effects of stress mediators. *New England Journal of Medicine*, *338*, 171–179.

McLaughlin, F. J., Altemeier, W. A., Christensen, M. J., Sherrod, K. B., Dietrich, M. S., & Stern, D. T. (1992). Randomized trial of comprehensive prenatal care for low-income women: Effect on infant birth weight. *Pediatrics*, *89*, 128–132.

Meaney, M. J., Aitken, S., Sharma, S., Viau, V., & Sarrieau, A. (1989). Postnatal handling increases hippocampal type II glucocorticoid receptors and enhances adrenocortical negative-feedback efficacy in the rat. *Journal of Neuroendocrinology*, *5*, 597–604.

Moore, M. L., Meis, P. J., Ernest, J. M., Wells, H. B., Zaccaro, D. J., & Terrell, T. (1998). A randomized trial of nurse intervention to reduce preterm and low birth weight births. *Obstetrics and Gynecology*, *91*(5. Pt 1), 656–661.

Murray, J. L., & Bernfield, M. (1988). The differential effect of prenatal care on the incidence of low birth weight among blacks and whites in a prepaid health care plan. *New England Journal of Medicine*, *319*, 1385–1391.

Mustard, A. C., & Roos, N. P. (1994). The relationship of prenatal care and pregnancy complications in birthweight in Winnipeg, Canada. *American Journal of Public Health*, *84*, 1450–1457.

Mvula, M. M., & Miller, J. M. (1998). A comparative evaluation of collaborative prenatal care. *Obstetrics and Gynecology*, *91*(2), 169–173.

Olds, D. L., Eckenrode, J., Henderson, C. R., Jr., Kitzman, H., Powers, J., Cole, R., et al. (1997). Long-term effects of home visitation on maternal life course and child abuse and neglect: 15-year follow-up of a randomized trial. *The Journal of the American Medical Association*, *278*, 637–643.

Olds, D. L., Henderson, C. R., Jr., Cole, R., Eckenrode, J., Kitzman, H., Luckey, D., et al. (1998). Long-term effects of nurse home visitation on children's criminal and antisocial behavior: 15-year follow-up of a randomized trial. *The Journal of the American Medical Association*, *280*, 1238–1244.

Olds, D. L., Henderson, C. R. Jr., Tatelbaum, R., & Chamberlain, R. (1986). Improving the delivery of prenatal care and outcomes of pregnancy. A randomized control trial of nurse home visitation. *Pediatrics*, *77*, 16–28.

Parker, B., McFarlane, J., & Soeken, K. (1994). Abuse during pregnancy: Effects on maternal complications and birth weight in adult and teenage women. *Obstetrics and Gynecology*, *84*, 323–328.

Partridge, C. A., & Holman, J. R. (2005). Effects of a reduced-visit prenatal care clinical practice guideline. *The Journal of the American Board of Family Practice*, *18*(6), 555–560.

Perkocha, V. A., Novotny, T. E., Bradley, J. C., & Swanson, J. (1995). The efficacy of two comprehensive perinatal programs on reducing adverse perinatal outcomes. *American Journal of Preventive Medicine*, *11*(3 Suppl), 21–29.

Polednak, A. P. (1996). Trends in urban black infant mortality, by degree of residential segregation. *American Journal of Public Health*, *86*, 723–726.

Quick, J. D., Greenlick, M. R., & Roghmann, K. J. (1981). Prenatal care and pregnancy outcome in an HMO and general population: A multivariate cohort analysis. *American Journal of Public Health*, *71*, 381–390.

Quinlivan, J. A., & Evans, S. F. (2004). Teenage antenatal clinics may reduce the rate of preterm birth: A prospective study. *BJOG*, *111*(6), 571–578.

Raine, T., Powell, S., & Krohn, M. (1994). The risk of repeating low birth weight and the role of prenatal care. *Obstetrics and Gynecology*, *84*, 485–489.

Reichman, N. E., & Teitler, J. O. (2005). Timing of enhanced prenatal care and birth outcomes in New Jersey's HealthStart program. *Maternal and Child Health Journal*, *9*(2), 151–158.

Sánchez-Nuncio, H. R., Pérez-Toga, G., Pérez-Rodriguez, P., & Vázquez-Nava, F. (2005). Impact of the prenatal care in the neonatal morbidity and mortality. *Revista Médica del Instituto Mexicana del Seguro Social*, *43*(5), 377–380.

Sapolsky, R. M. (1995). Social subordinance as a marker of hypercortisolism: Some unexpected subtleties. *Annals of the New York Academy of Sciences*, *771*, 626–639.

Sapolsky, R. M., Krey, P., & McEwen, B. (1984). Glucocorticoid-sensitive hippocampal neurons are involved in terminating the adrenocortical stress-response. *Proceedings of the National Academy of Sciences of the United States of America*, *81*, 6174.

Scholl, T. O., Miller, L. K., Salmon, R. W., Cofsky, M. C., & Shearer, J. (1987). Prenatal care adequacy and the outcome of adolescent pregnancy: Effects on weight gain, preterm delivery, and birth weight. *Obstetrics and Gynecology*, *69*, 312–316.

Schramm, W. F. (1992). Weighing costs and benefits of adequate prenatal care for 12,023 births in Missouri's Medicaid program, 1988. *Public Health Reports*, *107*, 647–652.

Schwartz, S. (1962). Prenatal care, prematurity, and neonatal mortality. *American Journal of Obstetrics and Gynecology*, *83*, 591–598.

Seckl, J. R. (1998). Physiologic programming of the fetus. *Emerging Concepts in Perinatal Endocrinology*, *25*, 939–962.

Shiono, P. H., Kebanoff, M. A., Graubard, B. I., Berendes, H. W., & Rhoads, G. G. (1986). Birth weight among women of different ethnic groups. *The Journal of the American Medical Association*, *255*, 48–52.

Showstack, J. A., Budetti, P. P., & Minkler, D. (1984). Factors associated with birthweight: An exploration of the roles of prenatal care and length of gestation. *American Journal of Public Health*, *74*, 1003–1008.

Spencer, B., Thomas, H., & Morris, J. (1989). A randomized controlled trial of the provision of a social support service during pregnancy: The South Manchester Family Worker Project. *British Journal of Obstetrics and Gynecology*, *74*, 873–877.

Sram, R. J., Binkova, B., Dejmek, J., & Bobak, M. (2005). Ambient air pollution and pregnancy outcomes: A review of the literature. *Environmental Health Perspectives*, *113*(4), 375–382.

Taffel, S. (1978). Prenatal care: United States, 1969–1975. *Vital and Health Statistics, Series 21, no 33*. Washington, DC: National Center for Health Statistics.

Tasnim, N., Mahmud, G., & Arif, M. S. (2005). Impact of reduced prenatal visit frequency on obstetric outcome in low-risk mothers. *Journal of the College of Physicians and Surgeons – Pakistan*, *15*(1), 26–29.

Taylor, C. R., Alexander, G. R., & Hepworth, J. T. (2005). Clustering of U.S. women receiving no prenatal care: Differences in pregnancy outcomes and implications for targeting interventions. *Maternal and Child Health Journal*, *9*(2), 125–133.

Terris, M., & Glasser, M. (1974). A life table analysis of the relation of prenatal care to prematurity. *American Journal of Public Health*, *64*, 869–875.

Tyson, J., Guzick, D., Rosenfeld, C. R., Lasky, R., Gant, N., Jiminez, J., et al. (1990). Prenatal care evaluation and cohort analyses. *Pediatrics*, *85*, 195–204.

US Department of Health and Human Services. US Public Health Service. (1989). *Caring for our future: The content of prenatal care*. Washington, DC: Author.

Villar, J., Carroli, G., Khan-Neelofur, D., Piaggio, G., & Gülmezoglu, M. (2002). Patterns of routine antenatal care for low-risk pregnancy (Cochrane Review). *The Cochrane Library, Issue 3*. Oxford: Update Software.

Villar, J., Farnot, U., Barros, F., Victoria, C., Langer, A., & Belizan, J. M. (1992). A randomized trial of psychosocial support during high risk pregnancies. The Latin American Network for Perinatal and Reproductive Research. *New England Journal of Medicine*, *327*, 1266–1271.

Vintzileos, A. M., Ananth, C. V., Smulian, J. C., & Scorza, W. E. (2003). The impact of prenatal care on preterm births among twin gestations in the United States, 1989–2000. *American Journal of Obstetrics and Gynecology*, *189*(3), 818–823.

Vintzileos, A. M., Ananth, C. V., Smulian, J. C., Scorza, W. E., & Knuppel, R. A. (2002). The impact of prenatal care in the United States on preterm births in the presence and absence of antenatal high-risk conditions. *American Journal of Obstetrics and Gynecology*, *187*(5), 1254–1257.

Visintainer, P. F., Uman, J., Horgan, K., Ibald, A., Verma, U., & Tejani, N. (2000). Reduced risk of low weight births among indigent women receiving care from nurse-midwives. *Journal of Epidemiology and Community Health*, *54*(3), 233–238.

Wadhwa, P. D. (1998). Prenatal stress and life-span development. In H. S. Friedman (Ed.), *Encyclopedia of mental health* (pp. 265–280), San Diego: Academic.

Williams, J. W. (1915). The limitations and possibilities of prenatal care. *The Journal of the American Medical Association*, *LXIV*(2), 95–101.

Zotti, M. E., & Zahner, S. J. (1995). Evaluation of public health nursing home visits to pregnant women on WIC. *Public Health Nursing*, *12*(5), 294–304.

Chapter 9
Current Approaches to Reducing Premature Births and Implications for Disparity Elimination

Vijaya K. Hogan, Meghan E. Shanahan, and Diane L. Rowley

Background

In 2006, over 500,000 babies, or one of every eight babies born in the United States were born premature (Martin et al., 2009). Preterm birth (PTB) is associated with more than one third of all infant deaths (MacDorman & Mathews, 2008) and preterm births are at significantly increased risk of adverse neurodevelopment sequelae (IOM, 2007). Although the vast majority of preterm newborns survive, studies of short- and long-term outcomes find significantly higher rates of neurodevelopmental morbidity, sensory-neural impairments, and other disabilities (e.g., cerebral palsy, and visual, auditory and intellectual impairments), and higher rates of complications of the respiratory, gastrointestinal and renal systems (Barker, Osmaond, Golding, Kuh, & Wadsworth, 1998; Escobar, Littenberg, & Petitti, 1991; IOM, 2007; Kuban and Leviton, 1994). PTB is also a major economic burden with associated costs totaling over $26 billion in 2005 (IOM, 2007).

The US preterm birth rate rose 20% between 1990 and 2006 to 12.8 preterm births per 100 live births (Martin et al., 2009). This long term increase in the overall preterm birth rate is evident in both late preterm (34–36 weeks) and moderately preterm (32–36 weeks) births during this time period (1990–2006). Further, African-American women have almost a twofold increased risk of preterm birth compared to non-Hispanic White women (18.5% vs. 11.7%, respectively) (Martin et al.). This race/ethnic disparity in PTB remains largely unchanged and unexplained, and contributes to a lifelong cycle of reproductive disadvantage with far-reaching social and medical consequences (Frankel, Elwood, Sweetnam, Yarnell, & Davey-Smith, 1996; Hales et al., 1991; Law and Shiell, 1996; Leon, Lithell, Vagero, McKeigue, & Koupilova, 1997).

While there are many explanations for the rising preterm birth rates (i.e., *increased medical intervention to prevent infant or maternal death, more accurate gestational age dating, increase in multiple gestation pregnancies, increase in ART pregnancies, changes in the classification between live births and fetal deaths*) (Ananth and Vintzileos, 2008; IOM, 2007; MacDorman & Mathews, 2008; PHAC, 1999), none of these factors individually or collectively are primarily responsible for the increasing rates. The fact remains that decades of research and intervention has not resulted in any success in bringing these rates down, making this issue of major public health concern. The Healthy People 2010 objective is to reduce all population group-specific preterm birth rates to 7.6 per 100 live births (USDHHS, 2000). It is highly unlikely that this goal will be met given the current state of the science for both preterm birth prevention and disparity elimination.

V.K. Hogan (✉)
Department of Maternal and Child Health, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, 425 Rosenau Hall, CB# 7445, 421 Pittsboro Street, Chapel Hill, NC 27599, USA
e-mail: vhogan@email.unc.edu

A. Handler et al. (eds.), *Reducing Racial/Ethnic Disparities in Reproductive and Perinatal Outcomes: The Evidence from Population-Based Interventions*, DOI 10.1007/978-1-4419-1499-6_9,

State of the Science of Preterm Birth Prevention

Many of the challenges in preventing preterm birth rest in the nature of the known risk factors. The known risk factors and markers include race/ethnicity, history of a previous preterm birth, some medical conditions, smoking, stress, multifetal pregnancy, uterine or cervical anomalies, drug or alcohol use, urogenital tract infection, as well as late or no prenatal care (Goldenberg, Culhane, Iams, & Romero, 2008). A risk-based prevention approach generally relies on the ability to identify intervenable risk factors and then deliver specific interventions to address those identified risks. Eliminating the risk factors would theoretically lower the risk of the disease, in this case, preterm birth. However, the strongest risk factors and risk markers for preterm birth are typically not those amenable to intervention. Likewise, we have yet to identify ways to address the special needs of women in the highest risk groups. The failure of any clinical or public health intervention to reduce population rates of preterm birth is evident as the rate has not improved over time.

This chapter deals with evidence for two specific clinical approaches that have potential population-based effects. Over the past decade, hopes for successful prevention of PTB have been separately raised by two promising interventions: treatment for bacterial vaginosis and progesterone prophylactic therapy for the prevention of preterm birth. Based on its presumed promise, one intervention (progesterone) is becoming a part of standard prenatal clinical practice even in light of a limited evidence base and inconclusive clinical recommendations. The other intervention, addressing the risks of PTB clearly posed by BV, has remained mired in inconsistent study results, inconclusive practice recommendations, and thus non-action on a known risk that contributes substantially to disparities in PTB. As such, it is instructive to examine the processes of development, interpretation and application of the evidence for each intervention. Such an examination can inform the future development of clinical or population-based approaches for both prevention of preterm births and disparity elimination.

Development of Evidence: The Ideal Case

The prevailing framework that guides most clinical and population research is the primary prevention model. Within this framework, we generally navigate a very specific scientific process to systematically generate knowledge about a disease and its treatment. The systematic construction of knowledge occurs most effectively when answers to the following sequence of questions are sought: *What is the natural history of the disease/health status outcome? What factors cause the disease/health status outcome? What are the biologic or physiologic mechanisms? What interventions reduce these risk factors and thus reduce the risk of disease/unhealthy outcome? How do we apply these interventions effectively in clinical practice*?

While this process has been effective for disease reduction, we have learned that for disparity reduction, this is not enough. For example, let us examine the development of knowledge for addressing Sudden Infant Death Syndrome (SIDS), one of the top three leading causes of infant mortality. Traversing the first three questions of the sequence noted above, eventually led to the discovery that placing an infant to sleep in the prone position increased the risk of SIDS, leading to the introduction, testing and subsequent evidence base supporting the *Back to Sleep* intervention – placing an infant to sleep on their back rather than in the prone position. In the U.S., the *Back to Sleep* intervention was introduced through a national educational campaign fueled by and implemented through Title V agencies. Examination of the rates of SIDS death before and after the national introduction of the *Back to Sleep* campaign revealed a discontinuity in the rates as they shifted downward, heralding declines in the national rates of SIDS. These declines were seen

equally among whites, blacks, Asian, Hispanic, and Native American populations, attesting to the efficacy of the intervention. However, when examining the disparities that exist among these groups, particularly between blacks and whites, it became clear that the efficacy of the intervention did not extend to the elimination of the excess risk among blacks. While the evidence-based intervention was successful at decreasing each of the population-specific disease rates, it was not sufficient to also impact the *disparity*. The gap between black and white SIDS rates has persisted over time. Thus, in addition to the knowledge that needs to be generated about the cause of 'disease', we need to further generate knowledge about the populations that experience these adverse conditions. Specifically, we need to identify *what factors make some populations more likely to experience health disadvantages compared to other populations*, and then identify and implement ways to ameliorate these disadvantages. While the generation of knowledge relevant to successful disease reduction should theoretically drive disease prevalence rates down in all population subgroups, a knowledge base relevant to *population vulnerability* will be needed to change the *slope* of decline in the most vulnerable groups. Both a downward trend in disease rates as well as a change of slope, representing an *accelerated* rate of decline in the most vulnerable population groups, will be necessary to achieve disparity reduction.

Development of Evidence: The Reality

Despite the promise of two compelling clinical interventions to reduce rates of preterm birth and to close the existing disparity between populations, we argue that there have been several deviations from the ideal scientific process intended to systematically lead to the unfolding of knowledge and generate evidence. We also argue that these deviations have adversely impacted the likelihood of achieving declines in either preterm birth rates or in decreasing the disparity between groups. Case studies of these promising interventions and the evidence associated with their promulgation are presented below.

Case Study I: Screening and Treatment of Bacterial Vaginosis to Reduce Preterm Birth

Bacterial vaginosis (BV) is a genital tract infection characterized by a disruption in the normal balance of bacteria in the vagina. The normal vaginal ecosystem that is dominated by lactobacilli is changed by an overgrowth of harmful bacteria. BV is sometimes accompanied by discharge, odor, pain, itching, or burning, but in many cases is asymptomatic. BV is associated with a number of adverse maternal health conditions including pelvic inflammatory disease, post abortion infections, endometritis and HIV (Persson et al., 1996; Sewankambo et al., 1997; Soper, Bump, & Hurt, 1990; Taha et al., 1998).

BV is not considered to be a sexually transmitted infection, but the triggers that initiate the adverse microbial environment of BV are unknown. Risk factors sometimes associated with BV include douching, high number of sexual partners, and early age at sexual debut (Holzman et al., 2001). The prevalence of BV tends to be higher among African-American women, and is thought to be partly related to higher rates of douching among African-American women (Zhang, Thomas, & Leybovich, 1997; Zhang et al., 2004). While the excess risk for African-American women is also not explained by differences in other behaviors, an increase in BV has been correlated with stress (Culhane et al., 2001; Culhane, Rauh, McCollum, Elo, & Hogan, 2002; Holzman et al. 2001).

BV is usually diagnosed in research settings by use of a gram stain and Nugent's score (Nugent, Krohn, & Hillier, 1991) or in clinical settings by use of Amsel's criteria (Amsel et al., 1983). The Amsel criteria are the universally accepted clinical criteria for making a diagnosis of bacterial

vaginosis, although it has been shown that clinicians may selectively use subsets of these criteria in making a diagnosis in actual practice (Hogan et al., 2007; Keane, Maw, Protchard, & Ison, 2005; Ness, Kip, et al., 2006; Wiesenfeld & Macio, 1999). According to Amsel's criteria, bacterial vaginosis is present if any three of the following four conditions are present: (1) presence of a vaginal discharge that is of a milky consistency; (2) vaginal acidity, or pH, above 4.5; (3) a positive "whiff" test for amine odor; or, (4) the identification of at least 20% of the cells on a microscopic wet mount as "clue cells". The reliability of Amsel clinical criteria for diagnosing BV in community clinical practice settings is unknown (USPSTF, 2008).

The standard of care for pregnant women with symptomatic BV infection has always been to treat with antibiotics. Bacterial vaginosis has been treated with either systemic or vaginal application of antibiotics (clindamycin, erythromycin, metronidazole), but has a high rate of both spontaneous resolution (Ness, Kip, et al., 2006) and recurrence (Hay, 2009).

Because asymptomatic BV infection has been linked to adverse pregnancy outcomes, guidelines for screening symptomatic and asymptomatic pregnant women have been disseminated by three US scientific groups, the Centers for Disease Control and Prevention (CDC), the American College of Obstetricians and Gynecologists (ACOG), and the U.S. Preventive Services Task Force (USPSTF). In the early 1990s CDC recommended screening and treating only pregnant women who were symptomatic (CDC, 1993). In 1998, the guidelines were modified to recommend screening for asymptomatic patients at high risk for preterm labor (e.g., those with a history of preterm delivery) in the second trimester (CDC, 1998). In 2002, the guidelines were once again modified to recommend screening high-risk patients for asymptomatic BV at the first prenatal care visit, with treatment and follow-up for women who test positive. These guidelines have remained unchanged since 2002 [CDC, 2002, 2006a, 2006b, 2006c, (http://www.cdc.gov/std/treatment/2006/vaginal-discharge.htm, last accessed August 14, 2009)]. ACOG also recommended screening and treating high-risk pregnant women in 1998, but modified its recommendations in 2001 to recommend not screening high-risk asymptomatic women (ACOG, 1998, 2001). Clinical management guidelines for vaginitis published in 2006 noted the existence of conflicting study results for treatment of asymptomatic high-risk women and made no recommendation on screening and treatment of BV for this group (ACOG, 2006). The U.S. Preventive Services Task Force (USPSTF) reported insufficient evidence to recommend screening asymptomatic high risk women in 2001, but noted that the magnitude of benefit exceeded the risk in several studies (USPSTF, 2001). A 2008 review of evidence on the benefits and harms of screening and treating asymptomatic pregnant women found no studies that compared pregnancy outcomes for women in screened vs. non-screened populations; reported that clinical trials to treat women at high-risk for preterm birth (history of preterm labor or midtrimester miscarriage) had conflicting results due to methodological differences; and, concluded that a sizeable group of asymptomatic high-risk women would receive no benefit but may experience harm (Nygren et al., 2008b; USPSTF, 2008).

Early in the history of exploring BV as a potential ameliorable risk factor for preterm birth, several strong epidemiologic studies established that BV is associated with preterm birth (Gibbs, Romero, Hillier, Eschenbach, & Sweet, 1992; Gravett et al., 1986; Hay et al., 1994; Hillier, Nugent, et al., 1995; Watts, Krohn, Hillier, & Eschenbach, 1992). Consistency across studies, strength of evidence, strength and direction of the risk, temporality, and biologic plausibility firmly established the relationship between BV and preterm birth (Koumans, Markowitz, & Hogan, 2002). The proposed biologic mechanism of the effect of BV on preterm birth is that the vaginal infection ascends into the uterine cavity and stimulates an inflammatory response in either the mother or the fetus (Hillier, Krohn, et al., 1995; Hillier et al., 1998). In general, preterm birth is highly correlated with evidence of maternal infection of the decidua, placenta or amniotic fluid, and at earlier gestations, higher levels of infection are detected (Goldenberg et al., 2008; Pararas, Skevaki, & Kafetzis, 2006). Also, inflammation associated cytokines are higher in women who have delivered preterm, suggesting a connection between early labor and the action of these cytokines (Goldenberg, Hauth, & Andrews, 2000).

Based on the understanding of the role of infection in the initiation of early labor, and the excitement over the existence of a potentially treatable risk factor for preterm birth, several treatment trials ensued. The assumption underlying these trials was that if a drug could be identified that destroyed the BV-associated microbes, it could reduce the risk of preterm birth. It was assumed that the action of these "bad bugs" was the biologic mechanism leading to premature birth; therefore, the major clinical objective was to "kill" these microorganisms. It was unknown at the time whether vaginal or systemic treatment would be most effective, which drug would have the greatest impact, when treatment should occur, or whether it was critical to replace the "good bacteria" (*Lactobacilli*) that not only had been overrun by the bad microbes, but which would be eliminated along with the "bad" BV-associated microbes. As a result, the different clinical trials conducted included experimentation with varying aspects of the disease – vaginal or systemic therapy, types of antibiotics, timing of therapy, dosages, etc. However, once a clear association with preterm birth was established and treatment guidelines were published based on the earliest clinical trials (CDC, 1993), it became ethically impossible to continue to conduct trials where new experimental treatment regimens would be withheld from a control group. As the early treatment guidelines were targeted at symptomatic women, while up to 50% of BV positive women are asymptomatic, many later studies restricted their study populations to asymptomatic women who were BV positive.

With the above information as background, to assess the current state of the evidence of the effectiveness of the treatment of BV for the reduction of preterm birth, we conducted a review in 2008 of the Pubmed, Medline, CINAHL and Cochrane databases with the search terms "*BV, or vaginitis or Bacterial vaginosis and preterm birth or prematurity*" and set the limits to include only randomized controlled trials, reviews and meta-analyses between 1985 and midyear 2008. Three recent papers on this topic were comprehensive reviews: a U.S. Preventive Services Task Force Systematic review (Nygren et al., 2008a), a Cochrane Review paper (McDonald, Brocklehurst, & Gordon, 2007) and an evidence based review for ACOG (Okun, Gronau, & Hannah, 2005). All three reviews (Table 9.1) included all of the individual studies that emerged from our search. We refer in the discussion below to the summary findings from these reviews and cite the results as obtained from the individual studies in the tables.

Nygren et al. (2008a) included 14 studies (Table 9.1) and McDonald et al. (2007) included 15 studies (Table 9.1) in their comprehensive reviews, of which 10 were the same. Nygren et al. (2008a) selected studies that screened and treated asymptomatic pregnant women classified as either low-risk for preterm delivery (N = 3 studies), average risk for preterm delivery (N = 8 studies), or high-risk for preterm delivery (N = 6). Several studies included women in multiple categories as study classification according to risk for preterm delivery was not mutually exclusive. Nygren and colleagues (2008a) found no evidence of clinical benefit for preterm delivery among low-risk women who were screened and treated with antibiotics as compared to low-risk women who were screened for bacterial vaginosis but were not given antibiotics (absolute risk reduction: –0.019, 95% confidence interval (CI) –0.056 to 0.018). Pooled studies of average risk women showed no reduction in preterm delivery before 37 weeks (absolute risk reduction 0.006, CI –0.009 to 0.022), or before 32 weeks (absolute risk reduction 0.001, CI 0.008–0.010). For high-risk women, studies were not pooled for outcomes at <37 weeks because three studies reported benefits, one reported no benefit, and one reported an increased chance of preterm delivery. For delivery <34 weeks pooled data indicated no significant treatment effect (absolute risk reduction 0.006 CI –0.067 to 0.079). Furthermore, Nygren et al. (2008a) noted that a main reason for the difference in treatment response across studies of high-risk women was the baseline rates of preterm delivery in the placebo group (no treatment group). Studies in which the baseline risk for preterm delivery was greater than 30%, favored treatment, whereas studies with lower baseline risks favored the placebo group (Nygren et al., 2008a).

McDonald et al. (2007) reviewed 15 randomized clinical trials (Table 9.1) (N = 5,888 women), but expanded the inclusion of screening studies to include more than bacterial vaginosis

Table 9.1 Meta-analyses for the relationship between treatment for bacterial vaginosis and preterm birth

Source	Number of studies, N	Findings	Contextual factors	Disparities/ comments
Guise et al. (2001)	Seven RCTs; four average risk (McGregor, Joessof, McDonald, Carey) and five high risk (Morales, Hauth, McDonald, Carey,Vermeulen)	No benefit of treatment for average risk women; among high risk – one study showed no benefit, three showed some benefit of treatment	Outcomes PTB <37 weeks	Treatment protocols and timing vary across studies. No data presented by Race/ ethnicity
McDonald et al. (2007)	Cochrane Review. RCT that compared antibiotics with a placebo, no treatment, or alternative antibiotic regimens in women with BV; 15 studies, 2,989 in treatment group, 2,899 in control group	Antibiotics were effective in treating BVduring pregnancy [OR 0.17 (95% CI 0.15–0.20)] Treatment did not reduce risk of PTB birth before 37 weeks when compared to placebo or no treatment [OR 0.91 (95% CI 0.78–1.06)] Treatment before 20 weeks gestation may reduce risk of PTB before 37 weeks [OR 0.63 (95% CI 0.48–0.84)]	– Trial protocols varied in a number of ways: – How BV was diagnosed – Timing of the screening – Timing of treatment – Period between screening and treatment – Does not mention study settings – Only included studies that enrolled women in the second or third trimester	No studies present data by race/ ethnicity or SES
Okun et al. (2005)	RCT that compared antibiotics with a placebo, or no treatment, in women with BV; 14 studies, 3,146 in treatment group, 2,906 in control group	Antibiotics reduced the risk of persistent BV during pregnancy [RR 0.32 (95% CI 0.20–0.52) $p < 0.001$] Antibiotic treatment did not reduce the risk of preterm birth at <37 weeks [RR 0.93 (95% CI 0.70–1.22) $p = 0.6$]	– 12 studies evaluated the effect of antibiotics mainly on women with BV; in two of these studies some women also had Trichomonas vaginalis – Two studies evaluated the effect of antibiotics mainly on women with Trichomonas vaginalis; in one of these studies some woman had BV as well – There was variability among the studies in how BV was diagnosed – Doesn't mention where studies were conducted Outcome among high risk PTB <34 weeks	No studies present data by race/ ethnicity or SES
Nygren et al. (2008a)	Seven new studies added to 2001 (Guise) meta analysis for a total of 14 studies Included three low risk, four average risk and one high risk population	No benefit of treatment in average and low risk groups. Two studies showed a positive but non significant risk reduction, no effect shown in two studies and potential harm of treatment in one study. Combined results demonstrate no evidence of treatment benefit		Treatment protocols and timing vary across studies

(e.g., asymptomatic intermediate flora and Trichomonas vaginalis). With the combined results across studies, antibiotic therapy was effective at eradicating bacterial vaginosis during pregnancy (Peto odds ratio (OR) 0.17, CI 0.15–0.20, ten trials, N = 4,357 women) but treatment did not significantly reduce the risk of preterm birth <37 weeks (OR 0.91, CI 0.78–1.06, 15 trials, 5,888 women), preterm birth <34 weeks (OR 1.22, CI 0.67–2.19), five trials, N = 851 women), or, preterm birth <32 weeks (OR 1.14, CI 0.76–1.70, four trials, N = 3,565 women).

In women with a previous preterm birth, antibiotic treatment did not reduce the risk of subsequent preterm birth (OR 0.83, CI 0.59–1.17, five trials, N = 622 women); however, antibiotic treatment was found to have some effect on decreasing the risk of premature rupture of membranes (OR 0.14, CI 0.05–0.38, two trials, N = 114 women), and low birthweight (OR 0.31, CI 0.13–0.75, two trials, N = 114 women). McDonald et al. (2007) concluded that there was little evidence that screening and treating all pregnant women with asymptomatic bacterial vaginosis would prevent preterm birth and its sequelae. However, for some women with a previous preterm birth, the authors noted that treatment of bacterial vaginosis might reduce the risk of preterm rupture of membranes and low birthweight. McDonald further noted that five trials of early treatment (13–20 weeks gestation) of bacterial vaginosis or abnormal flora when analyzed separately demonstrated a significant decrease in preterm births at 37 weeks (OR 0.72, CI 0.55–0.95).

Okun et al.'s review (2005) which included 14 studies concluded that there was no evidence to support treatment for either symptomatic or asymptomatic BV for women at any level of risk during pregnancy. However, they also called attention to two studies (Lamont, Duncan, Mandal, & Bassett, 2003; Ugwumadu, Manyonda, Reid, & Hay, 2003) which showed positive effects with treatment at an earlier gestational period than the other studies.

It is not surprising that the comprehensive reviews reached conclusions that leaned toward no benefit to screening asymptomatic women, particularly those who are low or medium risk. Of the individual studies we identified in our review (Table 9.2), most of which were included in one or more of the meta-analyses described above, five showed positive effects (Hauth, Goldenberg, Andrews, Dubard, & Copper, 1995; Kiss, Petricevic, & Hussein, 2004; Lamont et al., 2003; Morales, Schorr, & Albritton, 1994; Ugwumadu et al., 2003) seven showed no effect (Carey et al., 2000; Guaschino et al., 2003; Joesoef et al., 1995; Kekki et al., 2001; McDonald et al., 1997; McGregor et al., 1994; Vermeulen & Bruinse, 1999) and, two appeared to increase preterm birth among treated women (Kurkinen-Raty, 2000; Odendaal, Popov, Schoeman, Smith, & Grove, 2002).

It is from this mixed evidence-base that the most recent guidelines from both ACOG and the USPSTF shifted to a more conservative conclusion about the benefits of screening and treating bacterial vaginosis among high-risk pregnant women (ACOG, 1998, 2001, 2006; USPSTF, 2001, 2008). The latest ACOG practice guideline on treatment of vaginosis during pregnancy makes no recommendations (ACOG, 2006) and the latest USPSTF review recommends against screening and treatment for low-risk pregnant women and concludes that the current evidence is insufficient to assess the balance of benefits and harms of screening high-risk women (USPSTF, 2008).

Critique of the Development of the Evidence Base for the Effectiveness of BV Treatment for Preventing PTB

While several studies consistently identified BV as a risk factor for PTB, no thorough natural history studies were conducted prior to the clinical trials, creating gaps in the basic scientific knowledge base about this syndrome and how to treat it. Specifically, we had little knowledge about the proper timing of treatment, vaginal vs. systemic effects, and the type of antibiotics and/or anti-inflammatory agents that would be most effective prior to the initiation of the major BV clinical trials. Some information had been identified about the recurrence rates of BV (Hay et al., 1994)

Table 9.2 Studies examining the evidence between treatment for bacterial vaginosis and preterm birth

Author (year)	Study design	Description of intervention	Population studied	Address disparities (Y/N)	Key findings *(PTB in Trt vs. control)*	Conclusions
Morales et al. (1994)	RCT	Oral metronidazole or placebo for 7 days	80 high risk women, 44 in treatment group and 36 in placebo group	No	18% of Trt group vs. 39% of placebo group ($p < 0.05$)	Reduction in PTB
McGregor (1994)	RCT	Vaginal clindamycin or placebo for 7 days	271 high and low risk women; 60 in treatment group, 69 in placebo group, and 122 with no BV	No	15% of Trt group vs. 7.2% of placebo group vs. 3.3% of non-BVgroup (RR 3.3; 95% CI 1.2–9.1)	Treatment group had higher rate of PTB
Hauth et al. (1995)	RCT	Oral metronidazole or erythromycin or placebo for 7 days	624 high risk women; 433 in treatment group and 191 in placebo group	No	Among women with BV: 31% of Trt group vs. 49% of placebo group ($p = 0.006$)	Reduction in PTB among women with BV
Joesoef et al. (1995)	RCT	Vaginal clindamycin or placebo for 7 days	681 high and low risk women; 340 in treatment group and 341 in placebo group	No	15% of Trt group vs. 13.5% of placebo group; (OR 1.11, 95% CI 0.77–1.61)	No difference between treatment and control
McDonald et al. (1997)	RCT	Oral metronidazole or placebo twice daily for 2 days	879 high and low risk women; 439 in treatment group and 440 in placebo group	No	7.2% of Trt group vs. 7.5% of placebo group (NS)	No difference between treatment and control
Vermeulen and Bruinse (1999)	RCT	Vaginal clindamycin of placebo for 7 days	168 high risk women; 83 women in treatment group and 85 women in control group	No	21.2% of Trt group vs. 27.7% of control group (NS)	No difference between treatment and control
Carey et al. (2000)	RCT	Two doses of 2 g oral metronidazole or two doses of a placebo	1,953 high and low risk women; 966 women in treatment group and 987 women in placebo group	No	12.2% of Trt group vs. 12.5% of placebo group (RR 1.0; 95% CI 0.8–1.2)	No difference between treatment and control
Kurkinen-Raty (2000)	RCT	2% clindamycin vaginal or placebo for 7 days	101 low risk women; 62 women in treatment group and 61 women in placebo group	No	13.7% of Trt group vs. 6% of placebo group (OR 2.5; 95% CI 0.6–10)	No difference between treatment and control
Kekki et al. (2001)	RCT	Vaginal clindamycin or placebo for 7 days	375 low risk women; 187 in treatment group and 188 in placebo group	No	5% of Trt group vs. 4% of placebo group (OR 1.3; 95% CI 0.5–3.9)	No difference between treatment and control

Odendaal et al. (2002)	RCT	Oral metronidazole 2 days	464 Primigravidae women; 67 women in treatment group, 83 women in placebo group, and 314 non-BV group 491 Multigravidae women; 74 women in treatment group, 53 women in placebo group, and 364 non-BV women	No	Primigravidae: 18% of Trt group vs. 16% of placebo group vs. 21% of non-BV group (NS) Multigravidae: 43% of Trt group vs. 29% of non-BV group (p = 0.0231); 43% of Trt group vs. 24% of placebo group (p = 0.0274)	Primigravidae: No difference between treatment and controls Multigravidae: Treatment group had worse PTB than placebo group and non-BV group
Guaschino et al. (2003)	RCT	2% vaginal clindamycin 7 days or no treatment	100 high and low risk women; 49 women in treatment group and 51 women in control group	No	12.2% of Trt group vs. 15.7% of control (p = 0.78)	Reduction in PTB
Lamont et al. (2003)	RCT	Vaginal clindamycin or placebo for 3 days; administered again if abnormal persisted 3 weeks later	409 high and low risk women; 208 in treatment group and 201 in placebo group	No	4% of Trt group vs. 10% of placebo group (p = 0.03)	Reduction in PTB
Ugwumadu et al. (2003)	RCT	Clindamycin 300 mg or placebo orally twice daily for 5 days	485 high and low risk women; 244 in treatment group and 241 in placebo group	No	Spontaneous preterm delivery 5% of Trt group vs. 12% of control group p = 0.001	Reduction of PTB
Kiss et al. (2004)	RCT	Women received standard treatment and follow up for any detected infection. In the control group, the results of the vaginal smears were not revealed to the caregivers	4,429 low and high risk women; 2,058 women in the intervention group and 2,097 women in the control group	No	3.0% of Trt vs. 5.3% of control (95% CI 1.2–3.6; p = 0.0001)	Kiss et al. (2004)

Table 9.3 Quality rating of studies associated with bacterial vaginosis treatment to reduce preterm birth

Author (year)	Type of study	Reporting	External validity	Internal validity – bias	Internal Validity – confounding	Power	Total quality score ≤14 = poor 15–19 = fair ≥20 = good
Morales et al. (1994)	RCT	7	1	5	4	1	18
McGregor et al. (1994)	RCT	10	1	5	4	0	20
Hauth et al. (1995)	RCT	10	1	4	5	0	20
Joesoef et al. (1995)	RCT	9	2	6	4	0	21
McDonald et al. (1997)	RCT	13	2	5	5	1	26
Vermeulen and Bruinse (1999)	RCT	7	2	6	4	1	20
Carey et al. (2000)	RCT	12	2	6	5	0	25
Kurkinen-Raty (2000)	RCT	9	1	5	4	2	21
Kekki et al. (2001)	RCT	8	2	4	5	1	20
Odendaal et al. (2002)	RCT	9	1	4	4	0	18
Guaschino et al. (2003)	RCT	9	2	3	4	0	18
Lamont et al. (2003)	RCT	10	2	5	5	2	24
Ugwumadu et al. (2003)	RCT	11	2	6	5	2	26
Kiss et al. (2004)	RCT	11	2	4	4	1	22

and the fact that later acquisition of BV did not seem to put women at risk for PTB (Riduan et al., 1993). Given the dearth of knowledge about this syndrome, each of the clinical trials used different types of drugs, and various routes of administration (oral vs. topical), and included interventions of different duration, frequency and timing, all of which may explain some of the variation observed in the results (Guise et al., 2001). Likewise, some reviewers raised concerns about variable definitions of bacterial vaginosis, as well as the timing of the previous preterm birth (any previous pregnancy vs. the most recent pregnancy). All of these issues clearly limit the generalizability of the study findings and call into question the clinical recommendations made based on these studies.

An additional consideration is that the largest clinical trial (Carey et al., 2000) may have some limitations that are not captured by the standards of quality applied for reviewing the evidence in this book (Table 9.3). One of these is that the timing of treatment may have occurred too late to have an impact on the inflammatory pathways creating risk for PTB. Secondly, the control groups included in each *meta-analysis* had variable baseline preterm birth rates across the studies being examined. The PTB rates ranged from 23 to 35–57% raising questions about how representative the combined study populations were of the general public.

As a result, even though BV is considered to be a strong risk factor for PTB and is a major contributor to the racial and ethnic disparity in this birth outcome (Hitti et al., 2007; Fiscella 1996), studies to date have not lead to consistent and clear professional recommendations, either for the interconceptional (Coonrod et al., 2008) or prenatal periods (ACOG, 2006; USPSTF, 2008), for its diagnosis or treatment. There has also been very little examination of the strengths and limitations of each study conducted in order to conceptually redefine the best strategy for screening and treatment with respect to timing, dosage, type of drug, population characteristics, and to conduct a new trial that takes into consideration lessons learned from these prior studies. Research on the effect of treatment for BV during pregnancy has essentially come to a standstill since investigators appear to have agreed that the definitive study has already been conducted (Carey et al., 2000). Consequently, no clear clinical strategy for screening and treatment of this major risk factor and contributor to disparities in preterm birth has been universally promulgated or consistently implemented in practice.

Case Study 2: Review of Development of the Evidence Base for Progesterone Effects on PTB

Another intervention to reduce the risk of preterm birth that has been the focus of a great deal of attention in the last decade involves the use of the hormone progesterone as a prophylactic agent during pregnancy. Three forms of progesterone have been administered in clinical trials: 17 alpha-hydroxyprogesterone caproate (17p), a synthetic agent administered by weekly injection, a naturally occurring progesterone administered by vaginal suppository, and a commercially produced vaginal progesterone gel. Despite the lack of a clear biologic model, the recent impetus to test the effectiveness of 17p for reducing premature birth was stimulated by studies conducted a generation ago that observed a reduction in uterine activity following 17p treatment. The disappointing results of the BV trials, and a dire need to see some progress in preterm birth prevention may have also fueled excitement over the initial results of 17p trials and eventual uptake of 17p into clinical practice. Recent randomized clinical trials of vaginal progesterone do demonstrate preterm delivery reduction among women experiencing a singleton pregnancy with a previous preterm delivery or with shortened cervical length (Da Fonseca, Celik, Parra, Singh, & Nicolaides, 2007; O'Brien et al., 2007).

To assess the state of the science with respect to the effectiveness of progesterone in reducing preterm birth, we conducted a review in 2008 of the Pubmed, Medline, CINAHL and Cochrane databases with the search terms "*progesterone or progestogens or 17 alpha hydroxyprogesterone caproate, and preterm birth or prematurity*", and set the limits to include only randomized controlled trials (RCTs), reviews, and meta-analyses for the years 1985–2008. We updated the search in 2009 with the same search terms. We found eight RCTs (Caritis et al., 2009; Da Fonseca, Bittar, Carvalho, & Zugaib, 2003; Da Fonseca et al., 2007; Meis et al., 2003; Norman et al., 2009; O'Brien et al., 2007; Rouse et al., 2007; Yemini et al., 1985) (Table 9.4) and seven meta-analyses (Coomarasamy, Thangaratinam, Gee, & Khan, 2006; Dodd, Flenady, Cincotta, & Crowther, 2006, updated in 2008; Goldstein, Berrier, Rosen, Sacks, & Chalmers, 1989; Keirse, 1990; Mackenzie, Walker, Armson, & Hannah, 2006; Sanchez-Ramos, Kaunitz, & Delke, 2005) published in this time period (Table 9.5). One meta-analysis was excluded because the outcome measure was miscarriage and not preterm birth (Daya, 1989). Although all of the meta-analyses included basically the same individual studies, most RCTs prior to 2003 consisted of small samples, and the interpretation of the pooled findings in two of the meta-analyses (Goldstein et al.; Keirse), did not include the results of large scale RCTs conducted in 2003 and 2006. Therefore, this review considers the individual studies and meta-analyses conducted pre-2003, and those conducted in 2003 and after, separately.

Pre 2003 Individual RCT studies (Table 9.4)*:* Yemini et al. (1985) randomized 80 high risk women into a treatment vs. control group; women in the treatment group received 17p intramuscular injections. The preterm birth rate was lower in the treatment group (16%) compared to the control group (38%) who received a placebo ($p < 0.05$), supporting the efficacy of 17p in the prevention of PTB.

Pre-2003 Meta-analyses (Table 9.5)*:* Goldstein et al. (1989) published a meta-analysis consisting of data from 15 studies on assorted progestogens and assessed the combined effect on preterm birth. They found only marginal effects of treatment on PTB and concluded that there was not enough evidence to support routine treatment except for women with defined hormonal deficiencies. Keirse (1990) published a meta-analysis using only studies ($n = 17$) that specifically included 17p as the treatment and concluded that while 17p had a modest effect on preterm birth (odds ratio, 0.5, 95% CI 0.3–0.85), there was no impact on perinatal morbidity or mortality.

2003–2009 Individual RCTs (Table 9.4)*:* In 2003, the National Institute of Child Health and Human Development (NICHD), Maternal Fetal Medicine Units (MFMU) Network published results of an RCT that assessed the effect of weekly *intramuscular administration* of progesterone at 15–20 weeks of gestation through 36 weeks of pregnancy on the risk of preterm birth in women who had previously experienced a spontaneous singleton preterm delivery (Meis et al., 2003). Of the 1,039

Table 9.4 Studies examining the evidence between progesterone use and preterm birth

Author (year)	Study design	Description of intervention	Population studied	Address disparities (Y/N)	Key findings *(PTB trt vs. untreated)*	Conclusions
Yemini et al. (1985)	RCT	17p 250 mg intramuscular weekly and unspecified placebo	80 high risk women; 39 in treatment group and 40 in placebo group	No	16% of Trt group vs. 38% of placebo group $p < 0.05$	Reduction in PTB
Da Fonseca et al. (2003)	RCT	Prophylactic vaginal progesterone or placebo applied nightly for 10 weeks	142 high-risk singleton pregnancies; 72 women in treatment group and 70 women in placebo group Ethnicity (Placebo vs. Progesterone) White 71.4% 68.0% Nonwhite 28.6% 32.0%	No	13.8% of Trt group vs. 28.5%, of placebo group ($p = 0.03$)	Reduction in PTB
Meis et al. (2003)	RCT	Weekly injections of 250 mg of 17p or weekly injections of an inert oil placebo	463 high and low risk women; 310 women in the progesterone group and the 153 women in the placebo group	No	36.3% of Trt group vs. 54.9% of the placebo group (RR 0.66; 95% CI 0.54–0.81)	Reduction in PTB
O'Brien (2007)	RCT	Daily vaginal gel containing 90 mg progesterone or placebo of an identical delivery system without progesterone to women with a history of spontaneous singleton preterm birth in the preceding pregnancy	332 in progesterone group and 327 in control group	See DeFranco, O'Brien, Adair, Lewis, Hall, Fusey, et al. (2007), below	10.0% in the progesterone group and 11.3% in placebo group, not significant	No reduction in early PTB
DeFranco, O'Brien, Adair, Lewis, Hall, Phillips, et al. (2007)	RCT Subanalysis of O'Brien (2007)	See O'Brien et al. (2007) above	19 of 313 who received progesterone and 27 of 307 who received the placebo with cervical length <28 mm at enrollment	No	0% in progesterone group vs. 29.6% in placebo group, $p = 0.014$	Vaginal progesterone may reduce early PTB in women with a short cervical length

DeFranco, O'Brien, Adair, Lewis, Hall, Fusey, et al. (2007)	RCT Sub-analysis of O'Brien (2007)	See O'Brien et al. (2007) above	Subpopulation of 116 of 620 women in the study population described above with midtrimester cervical shortening of </=30 mm (30 black, 26 white, 32 Asian/Pacific Islander, 28 Hispanic/other women)	Yes, subanalysis b.	An improvement in gestational age: 35.2 weeks placebo vs. 36.6 weeks progesterone, p = 0.038, which was significant in black women, 34.3 weeks (placebo) vs. 37.6 weeks (progesterone), p = 0.032; and white women, p = 0.098, but not in other racial/ethnic groups	Progesterone gel is associated with later gestational age at delivery in black women.
Da Fonseca et al. (2007)	RCT	200 mg capsule of micronized progesterone rightly from by weekly gestation to 33 weeks 6 days	250 women with a cervical length of < = 15 mm; 125 in treatment groupal in placebo group	Relative risk did not vary by race	19.2% of Trt group vs. 34–4% of placebo gtoup delivered before 34 weeks gestation hazard rato, 0.57; 95% CI 0.35 to 0.92; P = 0.02	Reduction in PTB among women with a short anix
Rouse et al. (2007)	RCT	Weekly injections of 250 mg of 17p or weekly injections of castor oil placebo	327 women with twin gestations in the progesterone group and 334 women with twin gestations in the placebo group	No	Fetal death or delivery before 35 weeks gestation in 41.5%of 17p group and 37.3% ofplacebo group, RR = 1.1; 95% CI, 0.9–1.3)	No improvement in delivery or fetal death before 35weeks gestation.
Norman et al. (2009)	RCT	90 mg of progesterone gel daily or placebo	250 women with twin pregnancy in progesterone group and 250 women with twin pregnancy in placebo group	No	Fetal death or delivery before 34 weeks gestation in 24.7% of progesterone group and 19.4% of placebo group, OR 1.36, 95% CI 0.89–2.09, p = 0.1	No improvement in delivery or fetal death before 34 weeks gestation.
Caritis et al. (2009)		Weekly injections of 250 mg 17p in progesterone group or castor oil in placebo group	71 women with triple pregnancy in progesterone group and 63 women with triple pregnancy in placebo group	No	Fetal death or delivery before 35 weeks gestation in 83% of pregnancies in 17p group, and 84% in the placebo group, RR 1.0, 95% CI 0.9–1	No improvement in delivery or fetal death before 35 weeks gestation.

Table 9.5 Meta-analyses for the relationship between progesterone use and preterm birth

Source	Number of studies/N	Findings	Contextual factors	Disparities/comments
Goldstein et al. (1989)	RCT involving the use of progesterone and other progestational agents; 15 studies, 819 in treatment group, 841 in placebo group	Trt vs. placebo OR 0.71, 95% CI 0.49–1.02. Progesterone did not show a benefit in the reduction of PTB	– Two studies did not use a placebo – Doesn't mention where studies were conducted	No studies present data by race/ethnicity or SES
Keirse (1990)	Reviewed seven RCT, double blinded of which five looked at PTB as outcome. N = 104	PTB <37 weeks-OR 0.50 95% CI (0.30–0.80) Preterm labor OR 0.43 95% CI (0.20–0.89) LBW OR 0.46 95% CI (0.27–0.80)	– Included only studies which used a single agent, 17p, range of 250–1,000 mg weekly dose. Entry to study GA from 33 weeks or entry to PNC to 37 weeks…	No studies present data by race/ethnicity or SES While outcomes favored treatment, lower mortality and morbidity (RDS, hyperbilirubinemia) were not detected.
Sanchez-Ramos et al. (2005)	RCT that compared progesterone agents with a placebo for patients at risk for PTB; 10 studies, 562 in treatment group, 421 in placebo group	26.2% of Trt group vs. 35.9% of control group [OR 0.45 (95% CI 0.25–0.80)]	– Eight studies assess 17α-hydroxyprogesterone caproate and two assessed other progestational agents – Settings: United States, France, Brazil, Italy, Israel, and Finland	No studies present data by race/ethnicity or SES
Coomarasamy et al. (2006)	Meta-analysis included nine RCTs that recruited women at risk for PTB.	Protective effect of treatment for births <37 weeks GA: OR 0.42 95% CI (0.31–0.57); protective effect for births <34 weeks GA: OR 0.51 95% CI (0.34–0.77)	Included all progestational agents or metabolites, but excluded synthetic agents. Of nine studies, five used 17p, two used vaginal suppositories, one used oral progesterone and one used intramuscular pellets	No studies present data by race/ethnicity or SES
Dodd et al. (2006)	Cochrane Review. All published and unpublished randomized controlled trials, in which progesterone was given by any route for preventing PTB; 6 studies, 565 in treatment group, 423 in placebo group	Trt group vs. control group RR 0.65; 95% CI 0.54–0.79 Progesterone appears to lower the risk of PTB, but most results weighted by one trial. Also insufficient information about potential harms. More research is needed.	– Considered all routes of administering progesterone – Doesn't mention where studies were conducted	No studies present data by race/ethnicity or SES

Mackenzie et al. (2006)	Meta-analysis of randomized controlled trial to determine whether progesterone initiated in the 2nd trimester decreased risk of pattern delivery; Three studies, 399 women in the treatment groups and 249 in the placebo grouped.	Summary protective effect of treatment for spontaneous preterm birth, RR 0.57, 95% CI 0.36–0.90; for LBW, 0.66, 95% CI 0.51–0.87	Only RCTs that enrolled women by the second trimester were included. Studies were excluded it women enrolled had signs or symptoms of threatened, preterm lasor, or ruptured membranes	No studies present data by race/ethnicity or SES
Dodd et al. (2008)	Published randomized controlled trials in which progesterone was administered for the prevention of preterm birth; 11 randomized controlled trials (2,425 women and 3,187 infants) Primary outcome preterm birth at less than 34 weeks gestation	Progesterone was associated with reduction in preterm birth before 34 weeks, 142 women, RR 0.15, 95% CI 0.04–0.64, for women with a history of spontaneous preterm birth; progesterone was associated with reduction in preterm birth before 34 weeks, 250 women, RR 0.58, 95% CI 0.38–0.87 for women with a short cervix;	Considered all routes of administering progesterone – Doesn't mention where studies were conducted	No studies present data by race/ethnicity or SES

women eligible to participate in the trial, 463 (44%) enrolled and 59% were African-American. Women randomized to progesterone treatment had a significantly lower rate of PTB compared to the placebo controls at less than 37 weeks, less than 35 weeks, and less than 32 weeks of gestation; however, the rates of preterm delivery were extraordinarily high in both the treatment (36.3%) and placebo groups (54.9%) and when women randomized to progesterone were compared to a non-randomized comparable population, PTB rates (<37 weeks) were similar. Reductions in the rate of preterm delivery did not differ between black and non-black women.

The high rates of PTB in both the treatment and placebo groups in the Meis et al. (2003) study raise questions about the interpretation of the effect and measurement of effect sizes in this trial. One of the major issues centers on the use of castor oil as the placebo treatment which has been used by midwives for years to stimulate sluggish labor, and is known to stimulate uterine activity. Use of this placebo may have biased the results in favor of the treatment group. In any case, the higher than predicted rates of PTB in both the placebo and treatment groups need to be explained to allow appropriate assessment of the results of this study.

A second double-blind RCT (Rouse et al., 2007) tested the efficacy of 17p to prevent preterm birth among women with twin gestations. Delivery or fetal death before 35 weeks gestation occurred in 41.5% of pregnancies in the 17p group and in 37.3% of the placebo group [relative risk (RR) 1.1; CI 0.9–1.1]. Side effects, consisting mostly of injection site reactions, were frequent in both groups: 65.9% in the 17p group and 64.4% in the placebo group.

Da Fonseca et al. (2003) conducted a randomized clinical trial to examine the efficacy of daily progesterone in suppository form given to high-risk women between 24 and 34 weeks gestation, and found that PTB was 28.5% in the placebo group compared to 13.8% in the treated group for births before 37 weeks, and 18.6% (placebo) and 2.8% (progesterone treated women) for births before 34 weeks. A multicenter randomized clinical trial (Da Fonseca et al., 2007) tested the effect of double the dose of progesterone in suppository form on the incidence of spontaneous preterm delivery in asymptomatic women with shortened cervical length (15 mm or less) at 20–24 weeks gestation [the latter is an identified risk factor for spontaneous preterm birth (Iams et al., 1996)]. A total of 24,620 women undergoing routine ultrasonography at 20–24 weeks gestation agreed to transvaginal ultrasonographic measurement; 413 had cervical length of 15 mm or less and 250 participated in the clinical trial. The adjusted relative risk for spontaneous preterm delivery before 34 weeks gestation was 0.54, CI 0.34–0.88 with a 32.1% rate of spontaneous preterm delivery in the placebo group compared to 17.5% in the progesterone treatment group. Results were similar in the subgroup of women with no previous history of delivery before 34 weeks, 31.2% for the placebo group vs. 17.9% for the group treated with progesterone (RR 0.57, 95% CI 0.35–0.93). The findings from these two studies suggest that high-risk women, defined either by previous preterm delivery or short cervical length, benefit from treatment with progesterone suppositories between 24 and 34 weeks gestation.

O'Brien et al. (2007) conducted a multi-national double-blind randomized trial of daily progesterone vaginal gel (90 mg of progesterone gel considered equivalent to 600 mg of suppositories) involving pregnant women (N = 669) with a history of spontaneous singleton preterm birth between 20 and 35 weeks in the most recent previous pregnancy. Women were enrolled between 18 and 22 weeks +6 days gestation and treated until 37 completed weeks or delivery. The rate of preterm birth at less than or equal to 32 weeks gestation was not significantly different between the study groups. A secondary analysis limited to women with a cervical length less than 28 mm at the time of enrollment (N = 46) demonstrated a statistically significant reduction in deliveries at 32 weeks or less (0/19, or 0% vs. 3/27, or 29.6%, p = 0.014), but not with women with cervical lengths less than or equal to 30 mm (DeFranco et al., 2007).

Two 2009 RCTs assessed the use of progesterone to reduce PTB among women with multiple gestation and found no clinical effect (Caritis et al., 2009; Norman et al., 2009). Caritis et al. tested intramuscular 17p on women carrying triplets and examined a composite outcome of delivery or

fetal loss before 35 completed weeks gestation. Norman et al. examined a composite outcome of delivery or fetal loss before 34 weeks with the use of progesterone gel on women carrying twins.

Meta-analyses 2003 and after (Table 9.5)*:* Coomarasamy et al. (2006), Dodd et al. (2006), Dodd, Flenady, Cincotta, and Crowther (2008), Mackenzie et al. (2006), and Sanchez-Ramos et al. (2005), conducted meta-analyses published after 2003. Sanchez-Ramos included ten studies in her meta-analysis (including Da Fonseca et al., 2003; Meis et al., 2003; Yemini et al., 1985) and concluded that the preterm birth rate was lower in the progestogen treated group (OR 0.45, 95% CI 0.25–0.80) as well as in the subgroup who specifically received 17p (OR 0.45, CI 0.22–0.93). Perinatal mortality rates were not significantly different between the progesterone and placebo study groups. MacKenzie et al. limited her meta-analysis to studies with low loss to follow up and included three RCTs (Da Fonseca et al.; Johnson, Austin, Jones, Davis, & King, 1975; Meis et al.). She found a significant reduction in risk of preterm delivery among the progesterone group (RR 0.57, CI 0.36–0.90), but no significant effect on perinatal mortality or neonatal morbidity. Dodd et al. published a meta-analysis of eight studies that met their inclusion criteria (including Da Fonseca et al.; Meis et al.). For all women administered progesterone, there was a reduction in the risk of preterm birth at less than 37 weeks (six studies, N = 988, RR 0.65, CI 0.54–0.79) and at less than 34 weeks (one study, N = 142, RR 0.15, CI 0.04–0.64). Coomarasamy's meta-analysis reviewed studies using natural progesterone or its metabolites (including Meis). Pooled results showed a significant reduction in preterm delivery before 37 weeks gestation (OR 0.42, 95% CI 0.34–0.77), and before 34 weeks (OR 0.51, 95% CI 0.34–0.77). In addition, they found a 45% reduction in respiratory distress syndrome among infants born to mothers in the progesterone group. Cumulative meta-analysis by year of study and study quality showed similar results. The magnitude of benefit was similar across studies with variable baseline rates of preterm delivery in the control group.

Dodd et al. published a Cochrane Library systematic review of *Prenatal administration of progesterone for preventing preterm birth in women considered to be at risk of preterm birth* in 2006 and later updated it in 2008 (Dodd et al., 2006). The preliminary findings of the first meta-analysis of seven RCTs that met the inclusion criteria was that regardless of route of administration, women receiving progesterone were significantly less like to give birth before 37 weeks gestation, RR = 0.58 (95% CI 0.48–0.70), and sensitivity analysis using the highest quality trials indicated significant differences in intraventricular hemorrhage and risk of infant death (Dodd et al.). For the 2008 review, studies were subdivided based on the reasons the women were considered to be at risk for preterm birth (e.g., history of previous spontaneous preterm birth, multiple pregnancy, ultrasound-identified short cervical length, presentation with symptoms or signs of threatened preterm labor); preterm birth outcomes were restricted to less than 34 weeks gestation (Dodd et al., 2008). Eleven of 22 studies met the quality criteria: Borna and Sahabi (2008), Da Fonseca et al. (2003, 2007), Hartikainen-Sorri, Kauppila, & Tuimala (1980), Hauth et al. (1983), Johnson et al. (1975), Meis et al. (2003), O'Brien et al. (2007), Papiernik-Berkhauer (1970), Facchinetti, Paganelli, Comitini, Dante, & Volpe (2007), two of which were excluded from this analysis because they studied progesterone use only in women who were already in threatened labor (Borna & Sahabi; Facchinetti et al., 2007) (Table 9.6). Women with a history of spontaneous preterm birth or with a short cervix identified by ultrasound were significantly less likely to have a preterm birth before 34 weeks gestation (one study, RR 0.15, 95% CI 0.04–0.64; one study, 250 women, RR 0.58, 95% CI 0.38–0.87, respectively), but not women with multiple gestation.

Critique of the Development of the Evidence Base for the Effectiveness of the Use of Progesterone for Preventing PTB

The publication of the Meis et al. (2003) study generated a tremendous amount of excitement in the medical and public health communities because the findings suggested that there might be a potential "treatment" for preterm birth. However, this enthusiasm emerged despite the fact that the mechanism of action by which progesterone affects labor is not known although it is hypothesized

Table 9.6 Quality rating of studies of the association of progestational agents and preterm birth

Author (year)	Type of study	Reporting	External validity	Internal validity – bias	Internal validity – confounding	Power	Total quality score ≤14 = poor 15–19 = fair ≥20 = good
Yemini et al.(1985)	RCT	9	2	5	5	0	21
Da Fonseca et al. (2003)	RCT	10	2	6	4	1	23
Meis et al. (2003)	RCT	9	1	6	6	1	23
Da Fonseca et al. (2007)	RCT	11	2	7	6	1	27
O'Brien (2007)	RCT	11	2	6	4	1	24
DeFranco, O'brien, Adair, Lewis, Hall, Phillips, et al. (2007)	Secondary analysis of O'Brien et al. (2007)	7	1	5	4	0	17
DeFranco, O'Brien, Adair, Lewis, Hall, Fusey, et al. (2007)	Secondary analysis of O'Brien et al. (2007)	5	1	4	4	0	14
Rouse et al. (2007)	RCT	11	2	7	5	2	27
Norman et al. (2009)	RCT	11	1	6	5	1	24
Caritis et al. (2009)	RCT	11	1	7	5	2	26

to act by either relaxing uterine muscle or by exerting some anti-inflammatory effects. Progesterone may suppress myometrial activity, reduce the rate of cervical shortening, and down regulate the production of interleukin 8 by the cervix; importantly, vaginal administration may enhance bioavailability (Da Fonseca et al., 2007; DeFranco et al., 2007; O'Brien et al., 2007). In fact, there is some evidence for a biologic gradient for *vaginally administered* progesterone for women at risk of PTB due to cervical shortening. In the 2007 trial using progesterone suppositories for women with cervical lengths of 15 mm or less, the effect of progesterone appeared to be beneficial from the commencement of the drug (Nicolaides, Celik, & Fonseca, 2007); however, for 17p, a synthetic agent administered intramuscularly used in the Meis et al. study, there is no agreed upon bioequivalent to vaginal suppositories, and intramuscular administration may not provide endometrial concentrations as high as vaginal administration (O'Brien et al., 2007).

In addition to concerns about the biological mechanism for the action of 17p, Keirse (2004) identified methodologic problems in the Meis et al. (2003) study. The PTB rate in the treatment group (36.3%) was similar to the population-based rate of women with a previous PTB (37%), while the treated group would have been expected to have a lower rate had the treatment been effective. Also, the PTB rate in the control group was much higher than expected (54.9%). This result has not been explained, but there is speculation that the placebo used (castor oil), a known uterine stimulant, may have increased the rate in the placebo group.

Even though the evidence-base to support the use of 17p in clinical practice was not particularly deep, Petrini and others (2005) conducted an analysis of the effect of the nationwide use of 17p for all women with previous spontaneous preterm births and estimated only a modest impact of 17p (a 2% decrease in the national rate of PTB if eligible women, women with recurrent preterm birth) were treated but conclude that 17p is likely to play an important role in the future in reducing the risk of recurrent preterm birth. While a small potential impact was found, this may be an overestimate since the effect estimates were based on the Meis et al. (2003) study, in which the baseline rates of preterm delivery were extremely high for both treatment and control groups and in which

women received weekly injections of 17p, results that may not be replicable in 'real world' conditions. Additionally, Petrini et al.'s approach to estimating the potential national effect of 17p preventive therapy on preterm birth rates did not use methods that assessed both benefits and harms. The USPSTF recommends examining both benefits and harms when evaluating the magnitude of the effect of a potential preventive service. They also note that some harms occur in routine practice that are not completely measured and reported in randomized clinical trials (Sawaya, Guirguis-Blake, LeFevre, Harris, & Petitti, 2007). It is likely that the majority of women taking progesterone will not obtain any benefit because they would not develop the outcome regardless of treatment (Penston, 2005). Nearly 45% of women in the control group in the Meis et al. study did not experience preterm birth, the outcome of interest. If all of these women received progesterone, there would have been substantial overtreatment. In fact, acute and long term drug safety of both 17p and vaginal progesterone has been a concern of the U.S. Food and Drug Administration (FDA) and other clinicians and researchers. In particular, there is trepidation about miscarriages, stillbirths, perinatal morbidity and mortality, pregnancy complications, and long term childhood outcomes, as well as adverse events associated with administration of the treatment (Dodd et al., 2008; FDA, 2008; Rebarber et al., 2007; Thornton, 2007).

Although there was immediate enthusiasm in many quarters when the Meis et al. (2003) study was published, ACOG's clinical guidelines for progesterone supplementation issued in Nov 2003 appeared to be equivocal with respect to treatment, similar to BV guidelines: "*Recent studies* (Da Fonseca et al., 2003; Meis et al., 2003) *support the hypothesis that progesterone supplementation reduces the risk of preterm birth in a select group of women. Despite the apparent benefits of progesterone in this high risk population (referring to women with a prior preterm birth) ACOG believes that further studies are needed to evaluate the use of progesterone in patients with other high risk factors……When progesterone is used, it is important to restrict its use only to women with a documented history of a previous spontaneous birth at less than 37 weeks….because unresolved issues remain such as optimal route of drug delivery and long term safety of the drug*." More recently, the Society of Obstetricians and Gynecologists of Canada issued a statement (January 2008) indicating that women "should be aware that a previous spontaneous preterm labour and/or short cervix (<15 mm at 22–26 weeks' gestation) could be used as an indication for prophylactic progesterone treatment." In addition to recommending dose, the Society also suggested that women be informed about the lack of data for many neonatal outcomes (Doyle, 2009; Farine et al., 2008). Authors of some studies (Da Fonseca et al., 2007; Meis et al.; Tita & Rouse, 2009) and others (Lamont & Jayasooriya, 2009) are encouraging the use of 17p or vaginal progesterone in pregnant women with a previous spontaneous PTB or short cervical length as measured by transvaginal ultrasound, while others (How & Sibai, 2009) have suggested that more data are needed on the safety of progesterone in pregnant women and their children as well as comparative data on tolerance, pharmacokinetics of injectables, suppositories and gels. National surveys indicate that progesterone use by maternal fetal medicine specialists has increased by 76% between 2003 and 2005, from 38 to 67% (Ness, Dias, et al., 2006) and in 2007 over 70% of non maternal fetal medicine obstetricians in the US reported recommending progesterone (Henderson, Power, Berghella, Lackritz, & Schulkin, 2009).

Discussion: Impact of the Limitations in Evidence Development on the Elimination of Disparities in PTB

In the case of the two interventions discussed above, the ways in which the evidence bases have been developed and supported should give pause to the scientific community and encourage a reevaluation of our scientific and research priorities and directions with respect to preterm birth and disparity reduction. The concerns associated with the development of our current evidence base for PTB can be viewed as process, scientific, and conceptual.

Process issues. In both of the cases discussed here, we note some aspect of putting the proverbial cart before the horse. Treatment trials in the case of BV, ensued before a thorough understanding of the natural history of BV was developed; likewise, in the case of 17p, trials began before there was an adequate understanding of the underlying biologic mechanisms through which 17p exerts its effects. On the whole, the public health and medical communities often neglect to look critically and synthesize existing knowledge to inform the best approach for a next generation treatment trial. While there has been considerable research on BV, variation in treatment types, modality and timing in the studies conducted has led to difficulty in interpretation of the findings and a somewhat ambiguous and inconsistent message for clinical practice. While there has been less of a commitment to scientific inquiry on the effects of progesterone, translation into practice has occurred. As such, in both cases, the current situation has resulted in ambiguous guidelines for providers, and in interpreting the evidence for clinical practice. Likewise, in both cases, one large multi-center trial (Carey/BV; Meis/17p) significantly influenced the direction of the evidence and practice guidelines. However, while after the publication of the Carey BV trial, interest in and available grants to support BV research waned, after the publication of the Meis et al. (2003) progesterone trial, a variety of grants from institutions such as the CDC and March of Dimes ensued to support the development of population-based implementation strategies. In the case of 17p, one could question the logic of defining a population based strategy for implementation when there was not a strong clinical evidence base on which to draw conclusions about the efficacy of progesterone as a preterm birth prevention treatment, when there is a lack of understanding of the basic biological action of progesterone, and when there was some suggestion of possible adverse effects. In addition, the progesterone agent (17p) used in the Meis et al. study is not available commercially.

Scientific Issues. One major shortcoming of conducting the BV trials before the completion of a high quality natural history study is that questions remain unanswered about the pathway between BV and preterm birth. The Carey trial results suggests that BV is a systemic problem since the successful eradication of the BV microbes does not eliminate a recurrence during pregnancy and even without recurrence does not result in a reduction in preterm birth. Additionally, the demonstrated association of a non-reproductive tract infection such as periodontal disease (Sacco et al., 2008) with preterm birth further supports the notion of a systemic effect. However, is the problem of preterm birth primarily due to infection or is it related to inflammation? Bacterial vaginosis may be a marker of an inflammatory process or an immune response associated with a physiologic pathway that includes an altered immune response and the production of cytokines (Ruiz, Fullerton, & Dudley, 2003). The response may be triggered by psychological factors, social factors, biologic predisposition, or all three (Hobel, Goldstein, & Barrett, 2008; Pretorius, Jagatt, & Lamont, 2007; Ruiz et al.). Still unanswered is the role of lactobacilli and whether active re-colonization is needed to ensure a return to normal vaginal flora and to resolve the inflammatory process. Further, without more specificity than identifying "BV positive vs. negative" or symptomatic vs. asymptomatic women, and keeping in mind that many women do not recognize or report common BV symptoms, it has not been determined which women among the high-risk might benefit from treatment and which might be harmed.

In the case of progesterone, a reduction in preterm births is associated with some forms of treatment, however, questions about the safety of treatment, the optimal form of progesterone (synthetic vs. natural), the route of administration (intramuscular vs. vaginal), and commercial availability need to be answered. Several ongoing clinical trials may answer some of these questions (Farine et al., 2008; Thornton, 2007).

Conceptual Issues. When we attempt to address the risks for PTB and the potential for reducing disparities, it is important to remember that PTB has clear social correlates. These social factors impact initial risk and vulnerability, treatment access and acceptability, and the efficacy of treatment. Even assuming a clinical treatment can be identified, and assuming perfect clinical application, we would continue to see a disparity in PTB rates between blacks and whites unless we also address the social contributors to the

disparity as well as those that contribute to the disease. As such, reducing preterm birth and eliminating disparities requires more than a clinical strategy. A public health approach that takes into consideration population characteristics is essential. The current state of *clinical* evidence based medicine is therefore necessary but insufficient by itself to eliminate disparities in preterm birth. We suggest the need to develop a fuller *public health evidence base* that would define the specific ways to both reduce the rates of disease and change the slope of decline for those populations with excess risk.

Despite numerous studies over the last two decades, we are at a standstill with respect to BV research. In fact, we remain relatively close to the beginning of the process of evidence generation. BV continues to be a looming risk factor for PTB and black women continue to have higher rates of BV. While medical care may not be the prime factor in disparity causation, we cannot minimize the role it does play. As a known risk factor with higher prevalence among black women, the lack of a proven treatment modality for BV allocates a larger proportion of black women to a cycle of risk that can span across several pregnancies, perhaps over the her life course and may even have intergenerational effects. The high prevalence of BV, estimated to be upwards of 50% in African-American pregnant women, necessitates a stronger national push to find a treatment or a prevention modality that is effective this group.

In terms of progesterone use in reducing preterm birth, ongoing clinical trials will likely confirm its efficacy and answer practical concerns about treatment strategies. However, administration of progesterone may potentially increase the disparity because of lower rates of entry into prenatal care by the fourth month of gestation for black, non-Hispanic women compared to white, non-Hispanic women (Martin et al., 2009), and variation in subcategories of preterm birth (Zhang & Savitz, 1992). In 2003–2004, 86% of Non-Hispanic black women with a known prenatal care history entered care by 4 months gestation, compared to 94.1% of non-Hispanic white women (MacDorman & Mathews, 2008); if these data are typical, then many more black women than white women will be ineligible for screening and treatment with progesterone. Thus, it is imperative that the current clinical trials for efficacy be followed by effectiveness trials that determine how to provide the greatest number of black women with preventive care.

Our recommendations for improving the state of the evidence for PTB reduction and disparity elimination are:

1. Resume more carefully designed BV treatment research after carefully synthesizing current knowledge to establish a more refined sense of direction of what to test, who to treat and when to treat them with what therapy, bearing in mind that this might include consideration of interconceptional timing of treatment and some therapy for vaginal microbial normalization.
2. Clearly separate recommendations for symptomatic vs. asymptomatic BV positive high risk women to remove some degree of confusion among providers who may be over-cautious in interpreting the currently ambiguous guidelines.
3. Develop clear biologic models for the effects of progesterone on PTB, conduct more studies to add to the weight of evidence, and conduct long term follow-up of treated women and their infants to assess long term impacts of this treatment.
4. Increase the focus on population studies to address both the social contributors to PTB and those factors which mediate access to care and contribute to the disparity at every stage of risk and intervention.

Conclusion

In this chapter, we have identified some limitations in the development of the evidence base for treatment of PTB and have made a distinction between a clinical evidence base and a broader public health evidence base focused on social determinants. Reliance on the development of a clinical

evidence base to address individual risks is necessary but not sufficient, as the development of a public health evidence base is required to address and ultimately eliminate disparities in preterm birth.

In so far as a clinical intervention is conceptually justified, we need rigorous study and subsequent evaluation of the evidence before defining clinical recommendations that are unambiguous. In addition, it is no longer feasible to simply define an "effective" clinical intervention, we must also understand and define through research, the contextual conditions in which the intervention is most likely to be effective (including the appropriate timing of intervention, biologic plausibility, etc.). Further, we need improved strategies for the *design* and *delivery of* these interventions so as not to inadvertently increase disparities because of unequal access to prevention resources and/or treatment. Finally, we need to invest substantial research dollars and effort into decoding the complexities of the intersecting social, medical care, and medical risk contexts affecting the health and well being of black women. The fact that pertinent risk factors exist in greater prevalence in some populations is the first signal that social inequities exist and should stimulate a research strategy that aims to decipher this enigma as a primary prevention strategy. Although it may appear to be too complicated a task to unravel and address the complex world of historical and social influences on medical conditions, it cannot possibly be any more challenging than trying to find the magic clinical bullet – a challenge, we must add, that we have been failing to accomplish after decades and millions of dollars of investment devoted to this research.

References

American College of Obstetricians and Gynecologists. (1998). ACOG committee opinion. Bacterial vaginosis screening for prevention of preterm delivery. *International Journal of Gynecology and Obstetrics*, *61*(3), 311–312.

American College of Obstetricians and Gynecologists. (2001). ACOG Practice Bulletin. Assessment of risk factors for preterm birth. Clinical management guidelines for obstetrician-gynecologists. *Obstetrics and Gynecology*, *98*(4), 709–716.

American College of Obstetricians and Gynecologists. (2006). ACOG Practice Bulletin. Clinical management guidelines for obstetrician-gynecologists, Vaginitis. *Obstetrics and Gynecology*, *107*(5), 1195–1206.

Amsel, R., Totten, P. A., Spiegel, C. A., Chen, K. C., Eschenbach, D., & Holmes, K. K. (1983). Nonspecific vaginitis. Diagnostic criteria and microbial and epidemiologic associations. *American Journal of Medicine*, *74*(1), 14–22.

Ananth, C. V., & Vintzileos, A. M. (2008). Medically indicated preterm birth: Recognizing the importance of the problem. *Clinics in Perinatology*, *35*, 53–67.

Barker, D. J. P., Osmaond, C., Golding, J., Kuh, D., & Wadsworth, M. E. J. (1998). Growth in utero, blood pressure in childhood and adult life, and mortality for cardiovascular disease. *British Medical Journal*, *298*, 564–567.

Borna, S., & Sahabi, N. (2008). Progesterone for maintenance tocolytic therapy after threatened preterm labour: A randomized controlled trial. *Australian and New Zealand Journal of Obstetrics and Gynaecology*, 48(1), 58–63.

Carey, J. C., Klebanoff, M. A., Hauth, J. C., Hillier, S. L., Thom, E. A., Ernest, J. M., et al. (2000). Metronidazole to prevent preterm delivery in pregnant women with asymptomatic bacterial vaginosis. *New England Journal of Medicine*, *42*(8), 534–540.

Caritis, S. N., Rouse, D. J., Peaceman, A. M., Sciscione, A., Momirova, V., Spong, C. Y., et al. (2009). Prevention of preterm birth in triplets using 17 alpha-hydroxyprogesterone caproate: A randomized controlled trial. *Obstetrics and Gynecology*, *113*(2 Pt 1), 285–292.

Centers for Disease Control and Prevention. (1993). 1993 Guidelines for treatment of sexually transmitted diseases. *Morbidity and Mortality Weekly Report. Recommendations and Reports*, *42*(RR-14), 1–102.

Centers for Disease Control and Prevention. (1998). 1998 Guidelines for treatment of sexually transmitted diseases. *Morbidity and Mortality Weekly Report. Recommendations and Reports*, 47(RR-01), 1–118.

Centers for Disease Control and Prevention. (2002). Sexually transmitted diseases treatment guidelines – 2002. *Morbidity and Mortality Weekly Report. Recommendations and Reports*, *51*(RR-6), 1–80.

Centers for Disease Control and Prevention. (2006a). Recommendations to improve preconception health and health care – United States. A report of the CDC/ATSDR Preconception Care Workgroup and the select panel on preconception care. *Morbidity and Mortality Weekly Report. Recommendations and Reports*, *55*(RR-06), 1–23.

Centers for Disease Control and Prevention. (2006b). STD Treatment Guidelines 2006. Diseases characterized by vaginal discharge. *Morbidity and Mortality Weekly Report. Recommendations and Reports*, *55*(RR-11), 50–52.

Centers for Disease Control and Prevention. (2006c). *Racial and ethnic approaches to community health*. (December 12, 2006); http://www.cdc.gov/reach2010/goals.htm.

Coomarasamy, A., Thangaratinam, S., Gee, H., & Khan, K. S. (2006). Progesterone for the prevention of preterm birth: A critical evaluation of evidence. *European Journal of Obstetrics, Gynecology, and Reproductive Biology*,*129*(2), 111–118.

Coonrod, D. V., Jack, B. W., Stubblefield, P. G., Hollier, L. M., Boggess, K. A., Cefalo, R., et al. (2008). The clinical content of preconception care: Infectious diseases in preconception care. *American Journal of Obstetrics and Gynecology*, *199*(6), s296–s309.

Culhane, J. F., Rauh, V. A., McCollum, K., Elo, I., & Hogan, V. (2002). Exposure to chronic stress and ethnic differences in rates of bacterial vaginosis among pregnant women. *American Journal of Obstetrics and Gynecology*, *187*(5), 1272–1276.

Culhane, J. F., Rauh, V. A., McCollum, K., Hogan, V. K., Agnew, K., & Wadhwa, P. (2001). Maternal stress is associated with bacterial vaginosis in human pregnancy. *Maternal and Child Health Journal*, *5*(2), 127–134.

Da Fonseca, E. B., Bittar, R. E., Carvalho, M. H., & Zugaib, M. (2003). Prophylactic administration of progesterone by vaginal suppository to reduce the incidence of spontaneous preterm birth in women at increased risk: A randomized placebo controlled double blind study. *American Journal of Obstetrics and Gynecology*, *188*(2), 419–424.

Da Fonseca, E. B., Bittar, R. E., Damiao, R., & Zugaib, M. (2009). Prematurity prevention: The role of progesterone. *Current Opinion in Obstetrics & Gynecology*, *21*(2), 142–147.

Da Fonseca, E. B., Celik, E., Parra, M., Singh, M., & Nicolaides, K. H. (2007). Progesterone and the risk of preterm birth among women with a short cervix. *New England Journal of Medicine*, *357*(5), 462–469.

Daya S. (1989). Efficacy of progesterone support for pregnancy in women with recurrent miscarriage: A meta-analysis of controlled trials. *BJOG*, *96*, 275–280.

DeFranco, E., O'Brien, J., Adair, D., Lewis, D., Hall, D., Philips, J., & Creasy, G., et al. (2007). The PROVAGGEL study group. Is there a racial disparity in the efficacy of progesterone to prevent preterm birth. American Journal of Obstetrics and Gynecology, *197*(6), S200.

Dodd, J. M., Flenady, V., Cincotta, R., & Crowther, C. A. (2006). Prenatal administration of progesterone for preventing preterm birth. *Cochrane Database of Systematic Reviews*, *25*(1), CD004974.

Dodd, J. M., Flenady, V. J., Cincotta, R., & Crowther, C. A. (2008). Progesterone for the prevention of preterm birth: A systematic review. *Obstetrics and Gynecology*, *112*(1), 127–134.

Doyle, L. W. (2009). Antenatal progesterone to prevent preterm birth. *Lancet*, *373*(9680), 2000–2002.

Escobar, G. J., Littenberg, B., & Petitti, D. B. (1991). Outcome among surviving very low birthweight infants: A meta-analysis. *Archives of Disease in Childhood*, *66*(2), 204–211.

Facchinetti, F., Paganelli, S., Comitini, G., Dante, G., & Volpe, A. (2007). Cervical length changes during preterm cervical ripening: Effects of 17-alpha-hydroxyprogesterone caproate. *American Journal Obstetrics and Gynecology*, *196*(5), 453.

Farine, D., Mundle, W. R., Dodd, J., Basso, M., Delisle, M. F., Farine, D., et al. (2008). The use of progesterone for prevention of preterm birth. *Journal of Obstetrics and Gynaecology Canada*, *30*(1), 67–77.

Fiscella, K. (1996). Racial disparities in preterm births. The role of urogenital infections. *Public Health Reports*, *111*(2), 104–113.

Food and Drug Administration. (2008). 17-Alpha hydroxyprogesterone caproate for prevention of preterm birth: Overview FDA background document, Rockville, MD (July 29, 2008); http://www.fda.gov/ohrms/dockets/ac/06/briefing/2006-4227B1-02-01-FDA-Background.pdf.

Frankel, S., Elwood, P., Sweetnam, P., Yarnell, J., & Davey-Smith, G. (1996). Birthweight, body-mass index in middle age, and incident coronary heart disease. *Lancet*, *348*, 1478–1480.

Gibbs, R. S., Romero, R., Hillier, S. L., Eschenbach, D. A., & Sweet, R. L. (1992). A review of premature birth and subclinical infection. *American Journal of Obstetrics and Gynecology*, *166*(5), 1515–1528.

Goldenberg, R. L., Culhane, J. F., Iams, J. D., & Romero, R. (2008). Epidemiology and causes of preterm birth. *Lancet*, *371*(9606), 75–84.

Goldenberg, R. L., Hauth, J. C., & Andrews, W. W. (2000). Intrauterine infection and preterm delivery. *New England Journal of Medicine*, *342*(20), 1500–1507.

Goldstein, P., Berrier, J., Rosen, S., Sacks, H. S., & Chalmers, T. C. (1989). A meta analysis of randomized controlled trials of pregestational agents in pregnancy. *BJOG*, *96*, 265–274.

Gravett, M. G., Nelson, H. P., DeRouen, T., Critchlow, C., Eschenbach, D. A., & Holmes, H. H. (1986). Independent associations of bacterial vaginosis and Chlamydia trachomatis infection with adverse pregnancy outcome. *Journal of the American Medical Association*, *256*, 1899–1903.

Guaschino, S., Ricci, E., Franchi, M., Frate, G. D., Tibaldi, C., Santo, D. D., et al. (2003). Treatment of asymptomatic bacterial vaginosis to prevent preterm delivery: A randomized trial. *European Journal of Obstetrics, Gynecology, and Reproductive Biology*, *110*(2), 149–152.

Guise, J. M., Mahon, S. M., Aickin, M., Helfand, M., Peipert, J. F., & Westhoff, C. (2001). Screening for Bacterial vaginosis in pregnancy. *American Journal of Preventive Medicine*, *20*(3, Suppl 1) 62–72.

Hales, C. N., Barker, D. J. P., Clark, P. M. S., Cox, L. J., Fall, C., Osmond, C., et al. (1991). Fetal and infant growth and impaired glucose tolerance at age 64. *British Medical Journal*, *303*, 1019–1022.

Hartikainen-Sorri, A. L., Kauppila, A., & Tuimala, R. (1980). Inefficacy of 17 alpha-hydroxyprogesterone caproate in the prevention of prematurity in twin pregnancy. *Obstetrics and Gynecology*, *56*(6), 692–695.

Hauth, J. C., Gilstrap, L. C 3rd., Brekken, A. L., & Hauth, J. M. (1983). The effect of 17 alpha-hydroxyprogesterone caproate on pregnancy outcome in an active-duty military population. *Am J Obstet Gynecol*, *146*(2), 187–190

Hauth, J. C., Goldenberg, R. L., Andrews, W. W., Dubard, M. B., & Copper, R. L. (1995). Reduced incidence of preterm delivery with metronidazole and erythromycin in women with bacterial vaginosis: A placebo controlled doubly blind study. *American Journal of Obstetrics and Gynecology*, *171*(2), 345–347.

Hay, P. (2009). Recurrent bacterial vaginosis. *Current Opinion in Infectious Diseases*, *22*, 82–86.

Hay, P. E., Lamont, R. F., Taylor-Robinson, D., Morgan, D. J., Ison, C., & Pearson, J. (1994). Abnormal bacterial colonization of the genital tract and subsequent preterm delivery and late miscarriage. *British Medical Journal*, *308*(6924), 295–298.

Henderson, Z. T., Power, M. L., Berghella, V., Lackritz, E. M., & Schulkin, J. (2009). Attitudes and practices regarding use of progesterone to prevent preterm births. *American Journal of Perinatology*, *26*(7), 529–536.

Hillier, S. L., Krohn, M. A., Cassen, E., Easterling, T. R., Rabe, L. K., & Eschenbach, D. A. (1995). The role of bacterial vaginosis and vaginal bacteria in amniotic fluid infection in women in preterm labor with intact fetal membranes. *Clinical Infectious Diseases*, *20*(2), S276–S278.

Hillier, S. L., Martius, J., Krohn, M., Kiviat, N., Holmes, K. K., & Eschenbach, D. A. (1998). A case-control study of chorioamnionic infection and histologic chorioamnionitis in prematurity. *New England Journal of Medicine*, *319*, 972–978.

Hillier, S. L., Nugent, R. P., Eschenbach, D. A., Krohn, M. A., Gibbs, R. S., Martin, D. H., et al. (1995). Association between bacterial vaginosis and preterm delivery of a low- birth-weight infant. *New England Journal of Medicine*, *333*(26), 1737–1742.

Hitti, J., Nugent, R., Boutain, D., Gardella, C., Hillier, S. L., & Eschenbach, D. A. (2007). Racial disparity in risk of preterm birth associated with lower genital tract infection. *Paediatric and Perinatal Epidemiology*, *21*(4), 330–337.

Hobel, C. J., Goldstein, A., & Barrett, E. S. (2008). Psychosocial stress and pregnancy outcome. *Clinical Obstetrics and Gynecology*, *51*(2), 333–348.

Hogan, V. K., Culhane, J. F., Hitti, J., Rauh, V. A., McCollum, K. F., & Agnew, K. J. (2007). Relative performance of three methods for diagnosing bacterial vaginosis during pregnancy. *Maternal and Child Health Journal*, *11*(6), 532–539.

Holzman, C., Leventhal, J. M., Qiu, H., Jones, N., Wang, J., & The BV Study Group. (2001). Factors linked to bacterial vaginosis in non-pregnant women. *American Journal of Public Health*, *91*(10), 1664–1670.

How, H. Y., & Sibai, B. M. (2009). Progesterone for the prevention of preterm birth: Indications, when to initiate, efficacy and safety. *Journal of Therapeutics and Clinical Risk Management*, *5*(1), 55–64.

Iams, J. D., Goldenberg, R. L., Meis, P. J., Mercer, B. M., Moawad, A., Das, A., et al. (1996). The length of the cervix and the risk of spontaneous premature delivery. *New England Journal of Medicine*, *334*(9), 567–572.

Institute of Medicine Committee on Understanding Premature Birth and Assuring Healthy Outcomes. (2007). Premature birth: Causes, consequences, and prevention. In R. E. Behrman, A. S. Butler (Eds.), Washington, DC: National Academies Press.

Joesoef, M. R., Hillier, S. L., Wiknojosastro, G., Sumampouw, H., Linnan, M., Norojono, W., et al. (1995). Intravaginal clindamycin treatment for bacterial vaginosis: Effects on preterm delivery and low birth weight. *American Journal of Obstetrics and Gynecology*, *173*(5), 1527–1531.

Johnson, J. W., Austin, K. L., Jones, G. S., Davis, G. H., & King, T. M. (1975). Efficacy of 17 alpha hydroxyprogesterone caproate in the prevention of premature labor. *New England Journal of Medicine*, *293*(14), 675–680.

Keane, F. E. A., Maw, R., Protchard, C., & Ison, C. A. (2005). Methods employed by genitourinary medicine clinics in the United Kingdom to diagnose bacterial vaginosis. *Sexually Transmitted Infections*, *81*(2), 155–157.

Keirse, M. J. (1990). Progesterone administration in pregnancy may prevent preterm delivery. *BJOG*, *97*(2), 149–154.

Keirse, M. J. N. C. (2004). Progesterone and preterm: Seventy years of "déjà vu" or "still to be seen"? *Birth*, *31*(3), 230–235.

Kekki, M., Kurki, T., Pelkonen, J., Kurkinen-Raty, M., Cacciatore, B., & Paavonen, J. (2001). Vaginal clindamycin in preventing preterm birth and peripartal infections in asymptomatic women with bacterial vaginosis: A randomized, controlled trial. *Obstetrics and Gynecology*, *97*(5), 643–648.

Kiss, H., Petricevic, L., & Hussein, P. (2004). Prospective randomized controlled trial of an infection screening programme to reduce rate of preterm delivery. *British Medical Journal*, *329*(7462), 371. doi:10.1136/bmj.3869.519653.EB.

Koumans, E. H., Markowitz, L. E., & Hogan, V. K. (2002). Indications for therapy and treatment recommendations for bacterial vaginosis in nonpregnant and pregnant women: A synthesis of data. *Clinical Infectious Diseases*, *35*, S152–S172.

Kuban, K. C., & Leviton, A. (1994). Cerebral palsy. *New England Journal of Medicine*, *330*(3), 188–195.

Kurkinen-Räty, M., Vuopala, S., Koskela, M., et al. (2000). A randomised controlled trial of vaginal clindamycin for early pregnancy bacterial vaginosis. *BJOG: An International Journal of Obstetrics & Gynaecology*, *107*(11), 1427–1432.

Lamont, R., Duncan, S., Mandal, D., & Bassett, P. (2003). Intravaginal clindamycin to reduce preterm birth in women with abnormal vaginal tract flora. *Obstetrics and Gynecology*, *101*(3), 516–522.

Lamont, R. F., & Jayasooriya, G. S. (2009). Progestational agents for the prevention of preterm birth. *Journal of Perinatal Medicine*, *37*(1), 12–14.

Law, C. M., & Shiell, A. W. (1996). Is blood pressure inversely related to birth weight? The strength of evidence from a systematic review of the literature. *Journal of Hypertension*, *14*, 935–941.

Leon, D. A., Lithell, H., Vagero, D., McKeigue, P., & Koupilova, I. (1997). Biologic and social influences on mortality in a cohort of 15000 Swedes followed form birth to old age. *Journal of Epidemiology and Community Health, 51*, 594 (Abstract).

MacDorman, M., & Mathews, T. (2008). *Recent trends in infant mortality in the United States*. NCHS Data Brief. No. 9.

Mackenzie, R., Walker, M., Armson, A., & Hannah, M. E. (2006). Progesterone for the prevention of preterm birth among women at increased risk: A systematic review and metaanalysis of randomized controlled trials. *American Journal of Obstetrics and Gynecology*, *194*(5), 1234–1242.

Martin, J., Hamilton, B., Sutton, P., Ventura, S., Menacker, F., Kirmeyer, S., et al. (2009). Births: Final Data for 2006. *National Vital Statistics Reports*, *57*(7), 1–102.

McDonald, H. M., Brocklehurst, P., & Gordon, A. (2007). Antibiotics for treating bacterial vaginosis in pregnancy. *Cochrane Database of Systematic Reviews*, (1), CD000262.

McDonald, H. M., O'Loughlin, J. A., Vigneswaran, R., Jolley, P. T., Harvey, J. A., Bof, A., et al. (1997). Impact of metronidazole therapy on preterm birth in women with bacterial vaginosis flora: A randomized placebo controlled trial. *BJOG*, *104*(12), 1391–1397.

McGregor, J. A., French, J. I., Jones, W., Milligan, K., McKinney, P. J., Patterson, E., et al. (1994). Bacterial vaginosis is associated with prematurity and vaginal fluid mucinase and sialidase: Results of a controlled trial of topical clindamycin cream. *American Journal of Obstetrics and Gynecology*, *170*(4), 1048–1059.

Meis, P. J., Klebanoff, M., Thom, E., Dombowski, M. P., Sibai, B., Moawad, A. H., et al. (2003). Prevention of recurrent preterm delivery by 17-p alpha hydroxyprogesterone caproate. *New England Journal of Medicine*, *348*(24), 2379–2385.

Morales, W. J., Schorr, S., & Albritton, J. (1994). Effect of metronidazole in patients with preterm birth in preceding pregnancy and bacterial vaginosis: A placebo-controlled, double-blind study. *American Journal of Obstetrics and Gynecology*, *171*, 345–349.

Ness, A., Dias, T., Damus, K., Burd, I., & Berghella, V. (2006). Impact of the recent randomized trials on the use of progesterone to prevent preterm birth: A 2005 follow-up survey. *American Journal of Obstetrics and Gynecology*, *195*(4), 1174–1179.

Ness, R. B., Kip, K. E., Soper, D. E., Stamm, C. A., Rice, P., & Richter, H. E. (2006). Variability of bacterial vaginosis over 6- to 12-month intervals. *Sexually Transmitted Diseases*, *33*(6), 381–385.

Nicolaides, K., Celik, E., & Fonseca, E. B. (2007). Progesterone and preterm birth. *New England Journal of Medicine*, *357*(22), 2306–2307; author reply 2307.

Norman, J. E., Mackenzie, F., Owen, P., Mactier, H., Hanretty, K., Cooper, S., et al. (2009). Progesterone for the prevention of preterm birth in twin pregnancy (STOPPIT): A randomized, double-blind, placebo-controlled study and meta-analysis. *Lancet*, *373*(9680), 2034–2040.

Nugent, R. P., Krohn, M. A., & Hillier, S. L. (1991). Reliability of diagnosing bacterial vaginosis is improved by a standardized method of Gram stain interpretation. *Journal of Clinical Microbiology*, *29*, 297–301.

Nygren, P., Fu, R., Freeman, M., Bougatsos, C., Klebanoff, M., & Guise, J. M. (2008a). *Screening and treatment for bacterial vaginosis in pregnancy: Systematic review to update the 2001 U.S. Preventive Services Task Force recommendation statement*. Prepared for the Agency for Healthcare Research and Quality by the Oregon Evidence-based Practice Center under Contract No. 290-02-0024. Evidence Synthesis No. 57. AHRQ Publication No. 08-05106. Rockville, MD: Agency for Healthcare Research and Quality.

Nygren, P., Fu, R., Freeman, M., Bougatsos, C., Klebanoff, M., & Guise, J. M. (2008b). Evidence on the benefits and harms of screening and treating pregnant women who are asymptomatic for bacterial vaginosis: An update review for the U.S. Preventive Services Task Force. *Annals of Internal Medicine*, *148*(3), 220–233.

O'Brien, J. M., Adair, C. D., Lewis, D. F., Hall, D. R., Defranco, E. A., Fusey, S., et al. (2007). Progesterone vaginal gel for the reduction of recurrent preterm birth: Primary results from a randomized, double-blind, placebo-controlled trial. *Ultrasound in Obstetrics and Gynecology*, *30*(5), 687–696.

O'Brien, J. M., Defranco, E. A., Adair, C. D., Lewis, D. F., Hall, D. R., How, H., Bsharat, M., Creasy, G.W; Progesterone Vaginal Gel Study Group. (2009) Effect of progesterone on cervical shortening in women at risk for preterm birth: secondary analysis from a multinational, randomized, double-blind, placebo-controlled trial. *Ultrasound in Obstetrics and Gynecology*. 2009 Dec; *34*(6), 653–9.

Odendaal, H., Popov, I., Schoeman, J., Smith, M., & Grove, D. (2002). Preterm labour – is bacterial vaginosis involved? *South African Medical Journal*, *92*(3), 231–234.

Okun, N., Gronau, K. A., & Hannah, M. E. (2005). Antibiotics for bacterial vaginosis or Trichomonas vaginalis in pregnancy: A systematic review. *Obstetrics and Gynecology*, *105*(4), 857–868.

Papiernik-Berkhauer, E. (1970). Etude en double aveugle d'un medicament prevenant la survenue prematuree de l'accouchement chez les femmes a risque eleve d'accouchement premature. *Edition Schering*, *IV*(3), 65–68.

Pararas, M. V., Skevaki, C. L., & Kafetzis, D. A. (2006). Preterm birth due to maternal infection: Causative pathogens and modes of prevention. *European Journal of Clinical Microbiology & Infectious Diseases*, *25*(9), 562–569.

Penston, J. (2005). Large-scale randomized trials – A misguided approach to clinical research. *Medical Hypotheses*, *64*(3), 651–657.

Persson, E., Bergstrom, M., Larsson, P. G., Moberg, P., Platz-Christensen, J. J., Schedvins, K., & Wølner-Hanssen, P. (1996). Infections after hysterectomy: A prospective nationwide Swedish Study. *Acta Obstetrica et Gynecologiaca Scandinavica*, *75*, 757–761.

Petrini, J. R., Callaghan, W. M., Klebanoff, M., Green, N. S., Lackritz, E. M., Howse, J. L., et al. (2005). Estimated effects of 17alpha hydroxyprogesterone caproate on preterm birth in the United States. *Obstetrics and Gynecology*, *105*(2), 267–272.

Pretorius, C., Jagatt, A., & Lamont, R. F. (2007). The relationship between periodontal disease, bacterial vaginosis, and preterm birth. *Journal of Perinatal Medicine*, *35*(2), 93–99.

Public Health Agency of Canada. (1999). *Measuring up. A Health surveillance update on Canadian children and youth*. Canada. Retrieved December 20, 2005, form http://www.phac-aspc.gc.ca/publicat/meas-haut/mu_d_e.html.

Rebarber, A., Istwan, N. B., Russo-Steiglitz, K., Cleary-Goldman, J., Rhea, D. J., Stanziano, G. J., & et al. (2007). Increased incidence of gestational diabetes in women receiving prophylactic 17ª-hydoxyprogesterone caproate for prevention of recurrent preterm delivery. *Diabetes Care*, *30*(9), 2277–2280.

Riduan, J. M., Hillier, S. L., Utomo, B., Wiknjosastro, G., Linnan, M., & Kandun, N. (1993). Bacterial vaginosis and prematurity in Indonesia: Association in early and late pregnancy. *American Journal of Obstetrics and Gynecology*, *169*(1), 175–178.

Rouse, D. J., Caritis, S. N., Peaceman, A. M., Sciscione, A., Thom, E. A., Spong, C. Y., et al. (2007). A trial of 17 alpha-hydroxyprogesterone caproate to prevent prematurity in twins. *New England Journal of Medicine*, *357*(5), 454–461.

Ruiz, R. J., Fullerton, J., & Dudley, D. J. (2003). The interrelationship of maternal stress, endocrine factors and inflammation on gestational length. *Obstetrics Gynecology Survey*, *58*(6), 415–428.

Sacco, G., Carmagnola, D., Abati, S., Luglio, P. F., Ottolenghi, L., Villa, A., et al. (2008). Periodontal disease and preterm birth relationship: A review of the literature. *Minerva Stomatologica*, *57*(5), 233–246, 246–250.

Sanchez-Ramos, L., Kaunitz, A. M., & Delke, I. (2005). Progestational agents to prevent preterm birth: A meta analysis of randomized controlled trials. *Obstetrics and Gynecology*, *105*(2), 273–279.

Sawaya, G. F., Guirguis-Blake, J., LeFevre, M., Harris, R., & Petitti, D. (2007). Update on the methods of the US Preventive Services Task Force: Estimating certainty and magnitude of net benefits. *Annals of Internal Medicine*, *147*, 871–875.

Sewankambo, N., Gray, R. H., Wawer, M. J., Paxton, L., McNaim, D., Wabwire-Angen, F., et al. (1997). HIV-1 infection associated with abnormal vaginal flora morphology and bacterial vaginosis. *Lancet*, *350*(9077), 546–550.

Soper, D. E., Bump, R. C., & Hurt, W. G. (1990). Bacterial vaginosis and trichomoniasis vaginitis are risk factors for cuff cellulitis after abdominal hysterectomy. *American Journal of Obstetrics and Gynecology*, *163*, 1016–1021.

Taha, T. E., Hoover, D. R., Dallabetta, G. A., Kumwenda, N. I., Mtimavalye, L. A., Yang, L. P., et al. (1998). Bacterial vaginosis and disturbances in vaginal flora: Associations with increased acquisition of HIV. *AIDS*, *12*, 1699–1706.

Thornton, J. G. (2007). Progesterone and preterm birth – still no definite answers. *New England Journal of Medicine*, *357*(5), 499–501.

Tita, A. T., & Rouse, D. J. (2009). Progesterone for preterm birth prevention: An evolving intervention. *American Journal of Obstetrics and Gynecology*, *200*(3), 219–224.

Ugwumadu, A., Manyonda, I., Reid, F., & Hay, P. (2003). Effect of early oral clindamycin on late miscarriage and preterm delivery in asymptomatic women with abnormal vaginal flora and bacterial vaginosis: A randomized controlled trial. *Lancet*, *361*, 983–988.

U.S. Department of Health and Human Services. (2000). *Healthy people 2010: Understanding and improving health* (2nd ed.). Washington, DC: U.S. Government Printing Office. Retrieved August 10, 2008, from http://www.healthypeople.gov/document/html/objectives/16-11.htm.

U.S. Preventive Services Task Force. (2001). Screening for bacterial vaginosis in pregnancy: Recommendations and rationale. *American Journal of Preventative Medicine*, *20*(3S), 59–61.

U.S. Preventive Services Task Force. (2008). Screening for bacterial vaginosis in pregnancy to prevent preterm delivery: U.S. Preventive Services Task Force recommendation statement. *Annals of Internal Medicine*, *148*(3), 214–219.

Vermeulen, G., & Bruinse, H. W. (1999). Prophylactic administration of clindamycin 2% vaginal cream to reduce incidence of spontaneous birth in women with increased risk: A placebo controlled randomized double blind study. *BJOG*, *106*, 652–657.

Watts, D. H., Krohn, M. A., Hillier, S. L., & Eschenbach, D. A. (1992). The association of occult amniotic fluid infection with gestational age and neonatal outcome among women in preterm labor. *Obstetrics and Gynecology*, *79*(3), 351–357.

Wiesenfeld, H. C., & Macio, I. (1999). The infrequent use of office based diagnostic tests for vaginitis. *American Journal of Obstetrics and Gynecology*, *181*(1), 39–41.

Yemini, M., Borenstein, R., Dreazen, E., Apelman, Z., Mogilner, B. M., Kessler, I., et al. (1985). Prevention of premature labor by 17-phydroxyprogesterone caproate. *American Journal of Obstetrics and Gynecology*,*151*, 574–577.

Zhang, J., Hatch, M., Zhang, D., Shulman, J., Harville, E., & Thomas, A. G. (2004). Frequency of douching and risk of bacterial vaginosis in African-American women. *Obstetrics and Gynecology*, *104*(4), 756–760.

Zhang, J., & Savitz, D. A. (1992). Preterm birth subtypes among blacks and whites. *Epidemiology*, *3*(5), 428–433.

Zhang, J., Thomas, A. G., & Leybovich, E. (1997). Vaginal douching and adverse health effects: A meta-analysis. *American Journal of Public Health*, *87*(7), 1207–1211.

Chapter 10
Prenatal Case Management of Pregnant Women: What Is the Evidence for Its Contribution to a Reduction of Disparities in Perinatal Outcomes?[†]

L. Michele Issel, Jaime C. Slaughter, and Sarah G. Forrestal

Women who are at high risk for adverse pregnancy outcomes often have inadequate personal, psychological, or financial resources to overcome their multiple social and/or medical risk factors. These women therefore need assistance, support, guidance and oversight from a health or social service professional who meets them "where they are" both geographically and personally. One type of program that provides such assistance and support is case management. Definitions of what constitutes case management vary (Lee, Mackenzie, Dudley-Brown, & Chin, 1998), but generally the definition includes interventions, other than medical or clinical care, that facilitate access to and utilization of health and social services by clients from vulnerable populations (Hall, Carswell, Walsh, Huber, & Jampoler, 2002). For this review, we define case management as a health-focused service provided to pregnant women who are at medical or social risk for adverse pregnancy outcomes with the intent of assuring a healthy birth and improved neonatal outcomes. More specifically, prenatal case management (PCM) is designed to increase appropriate utilization of health and social services by pregnant women through simultaneous attention to their multiple medical and social needs. Notably, PCM is provided through a variety of programmatic structures, rather than representing a single discrete intervention, and is generally provided in the home of the pregnant woman by nurses or other health and social service providers who are based in community clinics or social service agencies. When provided in the home, PCM is typically called home visiting.

Based on the evidence from the Olds program of research (Olds & Kitzman, 1990; Olds & Korfmacher, 1997), which used registered nurses (RNs) to provide PCM through home visiting, the American Nurses Association (1997) published a position paper in support of the use of RNs as providers of PCM. Subsequently, the Maternal and Child Health Bureau of the Health Resources and Services Administration (HRSA) acknowledged the value of PCM, when provided as home visiting, for HRSA Title V programs (AMCHP, 1999), and state-wide nurse home visiting programs were created by innovative states, such as the one created by the Colorado State Board of Health (2000). Despite these professional endorsements of PCM, the magnitude of the effect of PCM on birth outcomes is unclear, particularly for women of color or ethnic minorities. In addition, the effectiveness of PCM when provided by non-nurses is unclear.

The purpose of this chapter is to review studies of PCM provided through various programmatic structures and with various providers, to assess whether PCM is effective in improving birth outcomes, and ultimately in reducing racial/ethnic disparities in these outcomes. As grounding for

[†] The superscript numbers preceding references are used in the text and in the tables to designate that reference.

L.M. Issel (✉)
University of Illinois, School of Public Health, Community Health Sciences,
1603 W. Taylor, Chicago, IL 60612, USA
e-mail: issel@uic.edu

A. Handler et al. (eds.), *Reducing Racial/Ethnic Disparities in Reproductive and Perinatal Outcomes: The Evidence from Population-Based Interventions*, DOI 10.1007/978-1-4419-1499-6_10,

our review, we first present a brief overview of the limitations inherent in the PCM literature. This is followed by a description of our methodology. The chapter concludes with a summary of the evidence and a set of recommendations.

Caveats When Reviewing the Literature on PCM Effectiveness

The Non-Specificity of "Case Management" as a Term

The term "case management" varies in meaning depending on the context and is easily confused with similar terms, such as care coordination and care management. Care coordination and care management generally denote medical management of health and illness processes, or refer to medical supervision provided within the clinical environment. Although important elements of a comprehensive health care system, neither care coordination nor care management is "case management." Similarly, case management does not include the provision of prenatal care, which is a medical service. In some circumstances, the term case management is used to refer to the means by which an insurance company manages the expenses incurred by its enrollees. This fiscal management of services is not case management as discussed in this chapter.

Case management, as a non-medical but health-focused program, is provided in various settings, including outpatient clinics and the client's home. The term "home visiting" is often used rather than case management for historical reasons. Our use of the term prenatal case management or PCM encompasses home visiting. In the definition of PCM used in this chapter, case management of pregnant women at risk or high risk is initiated during the pregnancy, and is generally continued for women and their infants, especially if either the woman or the infant is at risk. In such circumstances, prenatal case management becomes post-partum or general case management with a focus on parenting and infant development.

Case Management as a Variable Program

The complexities of conducting a review of prenatal case management stem from variations inherent in what constitutes case management. Case management is most accurately described as a program, rather than a service. Mueser and colleagues (1998), in a review of models of community-based case management for persons with severe mental illnesses, identified seven features of a case management program: staff-patient ratio, outreach to patients, shared caseloads, consumer input, emphasis on client skills training, frequency of client contacts, and locus of contacts. Interestingly, their list of programmatic characteristics does not include the specific interventions provided by the case manager. Ramey and Ramey (1993) in a review of home visiting programs suggested that home visitation (the older term for PCM) for pregnant women and new parents has the following key program characteristics: program philosophy, strategy, timing of visits, intensity of services, coordination of activities within the program, and sensitivity to social and family context. Since their report in 1993, attention has mostly focused on the structural characteristic of the PCM staff mix, specifically, the use of RNs versus para-professionals for the delivery of PCM (Korfmacher, O'Brien, Hyatt, & Olds, 1999).

There is only one standardized PCM program, the Nurse Family Partnership (n.d.), but this program model has not been universally adopted. Therefore, there is little or no consistency across agencies within and between states in how PCM is provided to women. This broader lack of standardization has potential consequences for interpreting the effectiveness of PCM. One study of a well established and standardized mental health case management model, the ACT model, revealed that fidelity to guidelines ranged from 69 to 89% (Dewa et al., 2003). This highlights a need for

caution in interpreting research on PCM effectiveness, especially given the lack of a standardized program being applied across the wide range of health care organizations providing PCM.

The lack of uniformity in what are included as interventions within PCM may stem from a corresponding lack of theoretical development that is specific to PCM. At best, multiple social interaction theories, such as role modeling, self-efficacy, social network theories, learning theories, and ecological frameworks can be inferred from the types of interventions provided. The literature addressing a theoretical basis for PCM is scant (Issel, 2000; Olds, 2002), providing minimal theoretical guidance for research or program development. The lack of an accepted theoretical basis for PCM makes it difficult to determine which potential maternal and neonatal outcomes would be most likely affected by interventions optimally provided within PCM programs. As such, this has resulted in inconsistency in what are studied as outcomes of PCM.

Variety of Outcomes Attributed to PCM

As the structure as well as interventions vary across individual PCM programs and because studies of PCM have had various purposes, a wide array of maternal outcomes have been considered as effects of PCM in the published literature (Gomby, Culross, & Behrman, 1999). PCM has been associated with improvements in anemia (Hardy, King, & Repke, 1987), increased social competence (Koniak-Griffin, Anderson, Verzemnieks, & Brecht, 2000), increased economic self-sufficiency (Kitzman et al., 1997), and having more problems resolved (Gonzalez-Cavlo, Jackson, Hanford, Woodman, & Remington, 1997). Researchers have also found PCM to be associated with decreased cesarean section rates (Hodnett & Fredericks, 2003) and improved life circumstances of the woman (Olds, Henderson, Tatelman, & Chamberlin, 1988), although other studies have found no such effects (Fraser, Armstrong, Morris, & Dadds, 2000). Some studies of case management of pregnant women have also examined child-focused outcomes (Koniak-Griffin et al., 2000; Roberts, Kramer, & Suisse, 1996), such as improved parenting (Kendrick et al., 2000; St. Pierre & Layzer, 1999), and reduction in childhood injuries (Roberts et al., 1996). Two frequently stated purposes of PCM are to reduce low birth weight (LBW) rate, the number of infants born weighing less than 2,500g per 1,000 births, and to reduce prematurity (births occurring at less than 37 weeks of completed gestation). From the array of reported outcomes from PCM, our literature review on the potential effect of PCM for reducing racial/ethnic disparities in birth outcomes focuses on the two most frequently reported outcomes: utilization of prenatal care and neonatal health as captured in both birth weight and gestational age. We describe the specific inclusion and exclusion criteria used in our review and analysis below.

Methods

Data Gathering

Search Strategy

The search for evidence regarding the effectiveness of PCM began with a review of articles on PCM gathered by the lead author over 15 years. These articles were used as the starting point for both ancestry and descendancy searches (Cooper, 1998) conducted in 2006 and 2007 using CINAHL, Medline, PychInfo, and Social Work Abstract databases. Articles published from 1985 to 2005 were sought. Similarly, reference lists of published meta-analyses were used to identify additional studies. The following keywords were used in various combinations for the literature searches: prenatal, pregnant, high risk, case management, home visiting, outcomes, birth, maternal, neonatal, community, and Medicaid.

In addition, websites of known PCM related programs (e.g., the Nurse Family Partnership) were reviewed for additional or new materials. Public health nursing and maternal and child health listservs were used to solicit unpublished reports and research findings. Three unpublished reports were provided from members of these listservs.

Inclusion and Exclusion Criteria

The first criterion for a study to be included in the review was that the program met the following definition of PCM: non-medical program or service that focused on facilitating utilization of health or social services, and was provided in the home or clinic to pregnant women who were at high medical or social risk for adverse birth outcomes. Thus, studies of medical prenatal care were excluded.

To be included in the review, a qualitative or quantitative study had to meet all of the following inclusion criteria: (a) published during or since 1985; (b) published in English; (c) reported on an intervention provided to pregnant women beginning before or during the third trimester; and, (d) reported on at least one of the following: a maternal health outcome in pregnancy, a maternal intrapartum or postpartum outcome, or an outcome related to the neonate. Any study design was acceptable if the study met all four inclusion criteria. Meta-analyses that met our definition of PCM were also reviewed.

Studies were excluded if they: (a) did not provide information on any of the interventions used within PCM (e.g., Zimmer-Gembeck & Helfan, 1995; Loda, Speizer, Martin, Skatrud, & Bennett, 1997), or did not include PCM as a distinct component of a comprehensive program (e.g., Goldfarb et al., 1991; Korenbrot, Gill, & Zoe-Patterson, 1995); (b) did not include at least one maternal health outcome important in pregnancy, or a maternal intrapartum or postpartum health outcome, or a neonatal outcome (e.g., Poland, Giblin, Waller, Bayer, 1991; Northeast Florida Healthy Start Coalition, n.d.); or, (c) reported only behavioral health outcomes, such as smoking cessation (e.g., Dolan-Mullen et al., 2006). A few reports were excluded because insufficient data were presented in the article to be meaningful in program comparisons (e.g., Stankaitis, Brill, & Walker, 2005). The exclusion criteria regarding lack of maternal or perinatal outcomes was used in combination with inclusion criteria requiring a maternal or neonatal outcome in order to avoid reviewing the large body of literature on effects of case management on parenting behaviors and infant development. The studies of case management initiated after the infant's birth or late in pregnancy tended to focus on parenting and child health outcomes, rather than perinatal outcomes, and thus were excluded. Exclusion criteria regarding lack of maternal and perinatal outcomes also eliminated from consideration studies that reported only on the process of PCM delivery or on costs, without providing any data on maternal or neonatal outcomes.

Data Abstraction and Management

The search strategies yielded 44 published articles meeting the inclusion and exclusion criteria and 8 systematic literature reviews. Studies were reviewed for eligibility as the data abstraction proceeded. This assured us that no study was excluded without adequate consideration, and resulted in the exclusion of studies that had initially appeared eligible based on title or abstract. During the data abstraction process, care was taken not to over-represent a single study that is associated with multiple published reports. For example, there were several reports (Olds, Henderson, Tatelman, & Chamberlin, 1986; Olds et al., 1988, 1997) of one well-known study conducted in Elmira, NY. We chose to abstract data from across the multiple published reports, thus yielding the most complete description possible of that one study. As a result, the 44 reports yielded 39 primary studies. Data abstracted from the 39 primary studies were entered into an Excel spread sheet for data management.

Study Sample Coding

For each study, data were abstracted on the race/ethnic composition of the study participants, with particular attention to whether any comparisons were conducted across racial/ethnic groups on the maternal and neonatal outcomes. Data were abstracted on race/ethnicity as reported for both the control/comparison and the experimental/intervention groups. Of the 39 studies, 18% ($n=7$) did not report the race/ethnicity of the study sample (Table 10.1). In 38% of the studies ($n=15$), at least 50% of the participants were African-American, and 31% ($n=12$) of the studies included some Latinas.

Within-Program Intervention Coding

Each study report was carefully read for any description of the actions performed by the case managers that were intended to have a beneficial effect on the client. We also noted structural characteristics of the PCM program (Table 10.2), specifically, whether a standardized or structured, semi-standardized, or an unstandardized protocol was followed by the case managers, and whether the PCM program was integrated into existing prenatal care services. Descriptions of PCM personnel and qualifications of case managers were used to capture the staff mix of PCM programs (Table 10.2). The RN-only model of PCM was used in 31% ($n=12$) of the studies, whereas 21% ($n = 8$) of the studies reviewed used only paraprofessionals or lay outreach workers. We also abstracted data on the gestational age at which most of the study participants were enrolled in the PCM program (Table 10.3). In nearly two thirds of the studies ($n=23$, 59%), we were unable to determine the gestational age at which the pregnant women were enrolled into PCM. Among studies providing information on gestational age at enrollment, none of the studies reviewed enrolled women in their first trimester.

Table 10.1 Summary of the ethnicity of participants in the PCM studies by study number ($n = 39$ primary studies)

Percent ethnicity	Number of studies	Study number
No ethnicity reported	7	3, 4, 5, 6, 20, 23, 40
White only reported	1	9
African-American		
≥75%	10	2, 10, 11, 14, 17, 18, 22, 25, 30, 31
50–74%	5	1, 8, 19, 26, 38
25–49%	3	12, 16, 39
<25%	7	15, 21, 24, 28, 35, 36, 37
Latina		
≥75%	2	27, 34
50–74%	2	28, 29
25–49%	4	26, 32, 35, 37
<25%	4	24, 33, 38, 39
Other ethnicities reported		
Asian		
25–49%	1	29
<25%	1	12
Native American/aboriginal		
<25%	1	7

Table 10.2 Summary of structural characteristics of PCM by study number ($n = 39$ primary studies)

Structural characteristics	Number of studies ($n = 39$)	Study number
Not standardized/structured/not integrated into PNC	4	17, 35, 37, 39
Semi-standardized/structured/not integrated into PNC	3	4, 19, 20
Standardized/structured/not integrated into PNC	18	2, 3, 5, 9, 10, 14, 15, 18, 21, 22, 25, 27, 28, 32, 34, 36, 38, 40
Integrated into prenatal care, standardization unknown	5	7, 16, 26, 29, 30
Integrated into prenatal care and standardized/structured	8	1, 6, 8, 11, 23, 24, 31, 33
Not Reported or Unclear	1	12
PCM staff mix and profession	**Number of studies ($n = 39$)**	**Study number**
RN only	12	3, 6, 18, 22, 25, 28, 29, 30, 31, 32, 35, 36
Paraprofessional, lay/outreach worker	8	2, 4, 7, 8, 11, 17, 19, 20
Interdisciplinary team[a]	8	1, 10, 21, 23, 24, 26, 27, 33
RN or SW	7	9, 15, 34, 37, 38, 39, 40
Not reported or unclear	4	5, 12, 14, 16

[a] Any combination of RN, SW, MD, and or lay/outreach worker

Table 10.3 Summary of Timing during pregnancy when PCM was initiated (when noted) by study number ($n = 39$ primary studies)

Intervention onset during pregnancy	Number of studies	Study number
1st trimester; through 13 weeks	0	
2nd trimester; 14–28 weeks	12	7, 8, 10, 16, 19, 20, 25, 27, 28, 30, 36, 39
3rd trimester; 29–40 weeks	3	4, 31, 33
Any time during pregnancy	1	36
Not reported or unclear	23	1, 2, 3, 5, 6, 9, 11, 12, 14, 15, 17, 18, 21, 22, 23, 24, 26, 29, 32, 34, 37, 38, 40

Outcome Coding

Initially, data on all outcomes reported in the primary studies were abstracted. This resulted in a wide range of outcomes, such as maternal perinatal morbidity (i.e., cesarean section, labor complications), maternal lifestyle behaviors, as well as neonatal morbidity. An inspection of the data on the various outcomes revealed that many of the outcomes were reported by only one or two studies. However, *prenatal care utilization* and *birth weight* or *gestation (neonatal outcome)*, were reported across the largest percentage of studies and thus were chosen for a more thorough analysis. These outcomes also reflect the impact of PCM for both the mother (prenatal care utilization) and the neonate (birth weight, gestational age), and resulted in the broadest possible inclusion of studies. Of note, although utilization of prenatal care can be viewed as an outcome directly linked to programmatic interventions, birthweight and gestational age are the result of a broader range of influences than receiving maternal PCM interventions. Nonetheless, improvement in birth outcomes is a major objective of PCM, and therefore an important focus for this review.

Statistical data indicative of programmatic effects were abstracted, including simple percentages, odds ratios, relative risk ratios, confidence intervals, and p values. Cooper (1998) outlined a literature synthesis approach, the vote count method, which is based on a simple count of the number of

Table 10.4 Summary of effects on the key outcomes by study number (*n*=39 primary studies)

	Number of studies	Study number
Prenatal care use (*n* = 16)		
Increase significant	6	1, 2, 10, 15, 17, 35
Increase not significant	5	11, 19[a], 27, 38, 40
Increase, unknown significance	2	21, 26
Same or decrease, not significant	2	28, 31
Same or decrease, significant	1	19[b]
Unknown, not significant	1	25
Infant birth outcome-BW/LBW (*n* = 33)		
Improvement significant	11	2, 5, 14, 16, 21, 25, 30, 32, 35, 36, 38
Improvement not significant	5	9, 19[b], 37, 39, 40
Improvement, unknown significance	2	29, 34
Same or worse, not significant	15	1, 3, 4, 6, 8, 11, 12, 17, 18, 19[a], 20, 24, 28, 31, 33
Unknown direction of effect, not significant	1	22
Infant birth outcome-PTB/GA (*n* = 23)		
Improvement significant	6	2, 7, 15[a], 25, 36, 38
Improvement not significant	4	1, 9, 15[b] 32
Improvement, unknown significance	2	28, 29
Same or worse, not significant	11	3, 4, 6, 11, 18, 20, 24, 30, 31, 33, 34
Unknown	1	23
Outcome reported by ethnicity (*n* = 5)		
Improvement significant for women of color	3	16, 25[c], 36[d]
Improvement not significance for women of color	2	15, 19

[a]For white women only
[b]For African-American women only
[c]Improvement only significant in African-American women under 19 years old
[d]Only African-American women studied

studies that show significant results for and against the intervention. The data we abstracted allowed us to conduct the vote count based on the number of studies that reported a significant or non-significant improvement for those in PCM compared to those not in PCM, those that noted an improvement but without any statistical significance reported, and those reporting no statistical improvement for the women who received PCM (Table 10.4).

Study Quality Coding

The quality and rigor of each primary study was assessed (Table 10.5) using the Quality Checklist form based on one developed by Downs and Black (1998) (as suggested by the editors' protocol for this book). For each element in the quality assessment form, the primary study was scored, using 1 for present or 0 for absent. Scores on the quality of reporting ranged from 3 to 11 of a possible 12, with 8.0 (SD = 1.9) as the mean reporting score. The external validity scores ranged from 1 to 4 of a possible 4, with 2.8 (SD = 1.0) as the mean score. The internal validity bias score ranged from 1 to 6 of a possible 7, with 3.6 (SD = 1.3) as the mean score, and the internal validity confounding score ranged from 1 to 5 of a possible 6, with 3.1 (SD = 1.2) as the mean score. Only 4 (10%) of the 39 studies mentioned a calculation of power or sample size (with a score ranging from 0 to 2). Thus, the total quality score, which is sum of the five subscale scores, ranged from 7 to 28 of a possible 31 with a mean of 17.6 (SD = 4.7).

Table 10.5 Summary of the study quality of studies of prenatal case management ($n = 39$)

Study number: author(s), year	Reporting	External validity	Internal validity-bias	Internal validity-confounding	Power	Total quality score (≤14 = poor, 15–19 = fair, ≥20 = good)
Mean (SD) scores:	8.0 (1.9)	2.8 (1.0)	3.6 (1.3)	3.1 (1.2)	0.2 (0.5)	17.6 (4.7)
1: Hardy et al., 1987	10	3	5	4	0	22
2: Heins et al., 1987	6	3	3	3	0	15
3: Olds et al., 1986, 1988, 1997	11	4	6	5	2	28
4: Dawson et al., 1989	8	1	3	3	0	15
5: Korenbrot et al., 1989	9	3	4	3	0	19
6: Oakley et al., 1990	7	2	3	3	2	17
7: Bryce et al., 1991	11	4	4	5	1	25
8: Graham et al., 1992	9	3	4	5	0	21
9: Villar et al., 1992	9	3	4	4	0	20
10: Bradley and Martin, 1994	8	1	3	3	0	15
11: Julnes et al., 1994	3	3	3	3	0	12
12: Zotti and Zahner, 1995	7	3	3	1	0	14
14[a]: Norbeck, Dejoseph, and Smith, 1996	10	4	3	5	0	22
15: Piper et al., 1996	8	3	5	4	0	20
16: Reichman and Florio, 1996	9	2	4	2	0	17
17: Rogers et al., 1996	9	3	5	4	0	21
18: Kitzman et al., 1997; 2000	10	4	3	5	1	23
19: Tessaro et al., 1997	9	4	4	2	0	19
20: Spencer, Thomas, and Morris, 1989	8	4	4	4	0	20
21: Baldwin et al., 1998	8	3	4	3	0	18
22: Gonzalez-Calvo et al., 1998	6	2	2	0	0	10
23: Lear et al., 1998	3	1	1	2	0	7
24: Lowry and Beikirch, 1998	7	4	4	3	0	18
25: Moore et al., 1998	9	4	6	5	0	24
26: Newschaffer et al., 1998	7	3	4	2	0	16
27: Thompson et al., 1998	10	3	4	2	0	19
28: Koniak-Griffin et al, 2000	10	3	5	4	0	22
29: Prozialeck and Pesole, 2000	4	2	1	2	0	9
30: Brooten et al., 2001	9	3	2	3	0	17
31: Margolis et al., 2001	10	3	6	4	0	23
32: Little et al., 2002	8	2	2	3	0	15
33: Jackson et al., 2003	6	1	3	1	0	11
34: Nguyen, Carson, Parris, and Place, 2003	8	3	2	2	0	15
35: Keeton et al., 2004	6	1	4	2	0	13
36: Carabin et al., 2005	9	3	4	2	0	18
37: Ricketts et al., 2005	10	3	4	4	0	21
38: Sangalang et al., 2006	8	4	4	3	0	19
39: Silva et al., 2006	7	1	3	2	0	13
40: Cramer et al., 2007	7	2	1	2	0	12

[a]Study 13 was excluded just prior to publication

Coding Reliability

Each primary study was independently reviewed and coded by each of the chapter authors. In the event of a discrepancy, the primary study was reviewed and discussed until an agreement was reached on the final coding and data abstracted. This iterative coding process resulted in

refinements and more accurate capturing of the information available in the publications and reports.

Review of the Evidence

Meta-Analyses

We located and reviewed eight literature syntheses or meta-analyses of PCM for pregnant women. We also initially considered literature reviews or meta-analyses of comprehensive prenatal care which encompassed PCM or home visiting; however, none of these reviews reported findings specific to PCM or home visiting. Thus, reviews such as those by Stevens-Simons and Orleans (1999) and Landis and Willems Van Dijk (2006) were excluded. The remaining eight reviews are presented separately from the primary studies so as to not bias or skew our findings, given that any primary study in these meta-analyses that fit our inclusion criteria was also included in our literature review (Table 10.6).

In a review of seven randomized trials, Olds and Kitzman (1990) concluded that home visiting for pregnant women was not consistently effective in improving birth outcomes, that effectiveness depended upon the risk characteristics of the sample and program characteristics, with the use of RNs being more effective, and that there was no clear pattern of a relationship between the focus of the program and the modest outcomes achieved. In 2002 Olds published a summary of 25 years of data from quasi-experimental longitudinal research. In this review, he showed that PCM home visiting using only RNs as case managers had better outcomes for pregnant and parenting women than programs using lay or para-professionals. One other meta-analysis found the provision of home visiting to pregnant women to be related to fewer cesarean births (Hodnett & Fredericks, 2003). None of the literature reviews or syntheses investigated effects by race or ethnicity.

PCM Program Characteristics

To describe the general nature of the PCM programs studied, we abstracted information on structural characteristics of the PCM programs (Table 10.2), PCM staff mix (Table 10.2), and the timing of initiation of PCM during the pregnancy (Table 10.3). The two structural characteristics commonly reported were the use of a protocol to ensure a standardized and structured intervention across recipients, and whether the PCM was an integrated part of prenatal care or a separate, distinct program. Eighteen (46%) of the PCM programs studied appeared to use a standardized intervention protocol, and were also stand-alone programs. Slightly fewer ($n = 13$, 33%) appeared to be integrated into prenatal care, with 62% of those ($n = 8$) using a standardized protocol.

The staff mix data revealed that less than one third ($n = 12$) examined the evidence-based Olds model in which RNs with a BSN only are used for PCM. Although the timing of PCM initiation was the only information reported which would allow for an approximation of PCM "dosage," 23 studies (59%) did not report which week or trimester of the pregnancy the woman began receiving PCM (Table 10.3). In 12 of 16 studies that provided information on intervention onset, women enrolled in PCM during the second trimester. Three of the studies noted enrollment in PCM in the third trimester and one reported that enrollment could have been any time during the pregnancy (Table 10.3).

Table 10.6 Summary of literature syntheses and meta-analyses of studies of PCM and home visiting (n=8)

Study: author(s), year	Source	Period covered	Review purpose	Primary study (n)	Conclusions
Combs-Orme, Reis, and Ward, 1985	Public Health Reports	1960–1984	Effectiveness of home visiting by public health RNs	8	Home visiting by public health RN affected knowledge and attitudes
Olds and Kitzman, 1990	Pediatrics	NR	Narrative review of RCT for prenatal and infancy home visiting to look for underlying cause model and at-risk characteristics	RCTs only	Prenatal programs were more effective if used RNs
Lumley, 1991	Inter J Tech Assess Health Care	NR	Review of 4 RCTs to examine effects of social support home visiting on PTB, BW, GA	NR	3 of 4 RCTS showed 25–50% reduction in PTB and 450g increase in mean birthweight
Ciliska et al., 1996	Can J Public Health	1986–1993	Home visiting effectiveness, narrative review, included prenatal, postnatal and preschool programs	77	Home visiting had beneficial effect on range of outcomes; no negative effects from home visiting
Gomby et al., 1999	The Future of Children	1999	Summarize articles in the issue of The Future of the Children and critique evaluation methods used	6	Improved interconceptional spacing for the mother
Ciliska et al., 2001	Primary Care Research and Development	1978–1998	Assess the effectiveness of public health nursing interventions delivered through home visiting	20	No impact on LBW, GA or IM
Olds, 2002	Preventive Science	25 years	Present conceptual model and findings by outcome domains	3 longitudinal studies	Fewer kidney infections, fewer PTB among smokers, increased birth weight among teens, decreased PIH, decreased vaginal infections. Recommends using RNs with a BSN
Hodnett and Fredericks, 2003	Cochrane Database of Sys Revs	1990–2003	Support during pregnancy for women at increased risk of low birthweight babies	18	No significant effect on LBW Significant positive effect on cesarean birth

NR not reported

We were not able to identify the level of risk among the study participants in any of the studies. Ostensibly, only women with multiple risk factors or at high-risk for poor birth and related outcomes receive PCM. However, it appeared that very few programs and their associated studies assessed or adjusted for risk, either in recruitment or during analysis.

Outcome #1: Maternal Use of Prenatal Care

Of the 39 studies reviewed, 16 (41%) reported on prenatal care use among women who received PCM (Table 10.7). These studies were mostly published in the mid- to late-1990s. The primary studies are noted numerically in the text in brackets, corresponding to their listings in Tables 10.7 and 10.8, and to the numbers indicated in the reference list. In addition, we use the notation "E" to represent women receiving the intervention and "C" to represent those not receiving the intervention. Across the studies, the number of women in the PCM treatment group ranged from 42,683 [35] in a statewide cohort study to 49 [11] in a randomized clinical trial. Use of prenatal care was measured as number of prenatal care visits [2, 25], prenatal care enrollment [10], or adequacy of prenatal care utilization typically using the Kotelchuck or Kessner indices [1, 11, 15, 17, 19, 21, 26, 27, 28, 31, 35, 38, 40].

Four [1, 15, 17, 35] of the eight retrospective cohort studies showed a significant positive program effect (Tables 10.4, 10.7). The only matched case-control study [2] found positive program effects, as did both [21, 26] ecological studies (although for the latter, the significance was not known). Across almost all of the studies reviewed, the PCM group had more or better prenatal care use than the comparison group, indicating that PCM has no adverse effects; the only exception was in the sole prospective cohort study [19]. Seven of the 16 studies reported the percent of each group (PCM vs. no PCM) that received adequate prenatal care. The difference between the PCM versus no PCM groups ranged from less than 1 to 31%. Among the 16 studies reporting on prenatal care utilization, 10 had samples which were greater than 50% African-American. Among those ten studies, four [1, 2, 17, 26] found a significant difference in prenatal care utilization favoring the PCM program, and one [19] found the opposite effect**.**

Outcome #2: Infant Birthweight and Gestation

Of the 39 studies reviewed, 36 (92%) reported on neonatal outcomes for women who received PCM (Table 10.8). Most of these studies were published in the mid-1990s. Across the studies, the number of women in the PCM treatment group ranged from 42,683 [35] in a statewide study to 27 [29] in a study of subsequent pregnancies. Birth outcomes were reported as average birth weight in grams, average gestational age in weeks, percent low birth weight (LBW), or percent preterm birth (PTB).

Of 14 randomized clinical trials reporting on infant birth outcomes, five [7, 14, 25, 30, 32] showed a significant positive program effect on one or more neonatal outcomes in certain subgroups or overall (Tables 10.4, 10.8). Four [15, 16, 35, 38] of ten retrospective cohort studies also found significant program effects. Three studies using alternative designs also found significant positive program effects [2, 5, 36].

Across the studies reviewed, the PCM group generally had better neonatal outcomes; however, in a few studies the PCM group had less favorable outcomes than the control or comparison groups. For example, Julnes, Konefal, Pindur, and Kim (1994) [11] reported the percent of study participants with a gestational age less than 38 weeks as 12.2% in the PCM group but only 4.3% in the

Table 10.7 Summary of studies of prenatal case management reporting on use of prenatal care ($n=16$)

Study #: author, year	Study design and study type	Description of intervention	Study sample size and population ethnicity	Key findings related to intervention effectiveness	Caveats/biases	Supports the intervention? For which populations?
Outcome #1: Use of prenatal care						
1: Hardy et al., 1987	Retrospective cohort Published article	Integrated, structured PCM by RN, SW, and health educator	E=744, C=744 Predominantly black	PNC adequacy p=0.006, E=9.2%, C=8.7%	Adolescents only	Yes, for adolescents
2: Heins et al., 1987	Matched case-control Published article	Structured Monthly HV during pregnancy and postpartum by lay "Resource Mother"	E=575, C=565, 89% AA	E had significantly more PNC visits		Yes
10: Bradley and Martin, 1994	Retrospective descriptive study Published article	Structured Multidisciplinary team of RN, SW, 4 outreach workers Care-coordination and HV at least once/month during pregnancy and at least once in first 3 months postpartum	381 enrolled, 310 through postpartum 93% AA	PNC enrollment, $p \leq 0.001$ At admission=85.4% Prior to delivery=99.1%	Compared on admit to program versus at delivery	Yes
11: Julnes et al., 1994	RCT Published article	Integrated, structured E=Lay HV C1=Clinic-based multidisciplinary team C2=no prenatal care	E=49, C1=46, C2=29 AA E=92%, C1=70% C2=79%	More that 6 PNC visits NS E=87.8%, C1=73.9% C2=0.0%		No
15: Piper et al., 1996	Retrospective cohort Published article	Structured RN or SW	E=3,859; C=62,192 AA E=20.9%, C=38.6%	Inadequate PNC OR=0.71 (95% CI 0.61, 0.82) E=5.8%, C=10.8%		Yes. AA in E had better PNC use (OR for inadequate prenatal care=0.46, 95% CI 0.35, 0.60)
17: Rogers et al., 1996	Retrospective cohort Published article	Structured, but no protocol Trained lay health workers Monthly visits during pregnancy and 1 year postpartum	E=1,901 C1=4,613; C2=712 AA E=77% C1=54%, C2=70%	Early PNC RR=1.48 (95% CI 1.32, 1.66) Adequacy of PNC RR=1.58 (95% CI 1.40, 1.78)	Within racial groups no difference in LBW, but better than expected, given risk.	Yes

19: Tessaro et al., 1997	Prospective cohort Published article	Semi-structured Social support, education and outreach from trained lay health workers 28 weeks to 1 year postpartum	E=1,319; C=9,255 AA E=61.8%, C=59.4%	PNC adequacy (Kessner) AA, $p<0.05$ E=60.7%, C=63.8% White NS E=77.4%, C=75.1%	AA participants had lower rates of adequate PNC than AA non-participants. White participants had non-significantly higher rates of PNC than white non-participants	No
21: Baldwin et al., 1998	Ecological Published article	Integrated interdisciplinary team in WA state E received up to 20 HV or office visits; mean number of prenatal HV 3.7 C group were Medicaid women in CO	E in WA state=7,555, but only 378 had PCM only E in CO state=5,065; C in WA state=6,537; C in CO state=4,053 AA E in WA=10.2%, E in CO=13.0, C in WA=9.1%, C in CO=15.0%	% of expected PNC visits completed, $p<0.001$ WA state E=87.1%, C=83.8% E=86.8%, C=84.0% CO state	Reported outcomes are for PCM plus maternity support services	Yes. Higher effects in groups with higher risk
25: Moore et al., 1998	RCT Published article	Structured RN 3 TC per week to 37th week	E=718, C=715AA E=77.6%, C=77.8% White E=22.4%, C=22.2%	NS difference in total number of PNC visits		No
26: Newschaffer et al., 1998	Ecological Published article	Integrated Multidisciplinary team Average 32.5 visits from enrollment to 12 months postpartum	E=240, C=113 AA E=57%, C=55.8% Hispanic E=34%, C=29.2%	Adequate PNC E=71%, C=40% Continuity of PNC E=44%, C=25%		Yes
27: Thompson et al., 1998	Retrospective cohort Published article	Structured PHN with outreach worker At least 3 visits during pregnancy	E=100, C=100 Hispanic E=100%, C=100%	Adequacy of PNC NS Number of PNC visits $p<0.05$	Small sample size due to rural setting of the program. Persistent barriers to PNC decreased program effects	No

(continued)

Table 10.7 (continued)

Study #: author, year	Study design and study type	Description of intervention	Study sample size and population ethnicity	Key findings related to intervention effectiveness	Caveats/biases	Supports the intervention? For which populations?
Outcome #1: Use of prenatal care						
28: Koniak-Griffin et al., 2000	RCT Published article	Structured PHN E=17 home visits C=1 or 2 prenatal home visits	E=62, C=59 White E=19%, C=21%	Mean number of PNC visits NS E=9.7, C=9.3	Pregnant adolescents only	No
31: Margolis et al., 2001	Retrospective cohort Published article	Integrated, structured PHN and ECE 2–4 visits per month	E=105, C=103 AA E=90.3%, C=4.9% White E=14.3%, C=80%	PNC adequacy adjusted, NS E=28.3%, C=27.6%	Interventions also at community and provider levels	No
35: Keeton et al., 2004	Retrospective cohort Published article	Unstructured Case manager with bachelor's level degree Services throughout pregnancy and 1–3 years pospartum	E=42,683; C=31,982 AA E=24%, C=52% Hispanic E=31.3%, C=22.7%	Less than adequate PNC E=37.7%, C=42.2%, $p<0.001$	Premature delivery bias since women could enter program anytime prior to birth	Yes
38: Sangalang et al., 2006	Retrospective cohort Published article	Structured SW or other health service professional 3–4 times/month	E=1,260; C=1,260 AA E=70.2%, C=48.4% Hispanic E=1.1%, C=4.4%	Less than adequate PNC NS E=47.8%, C=49.1%		No
40: Cramer et al., 2007	Retrospective cohort Published article	Structured SW or RN E=weekly contacts	E=79 C=746	PNC adequacy E=63.3% C=70.1% NS 1st trimester PNC E=67.1% C=85.5% (p=0.0011)		No

Abbreviations: *E* experimental group, *C* control group, *AA* African-American, *HV* home visiting, *PNC* prenatal care, *NS* not significant *RN* registered nurse, *SW* social worker, *TC* telephone calls, *PHN* public health nurse, *ECE* early childhood educator, *OI* office intervention, *NA* native American

Table 10.8 Summary of studies of prenatal case management reporting on infant birthweight or gestational age ($n=36$)

Study #: author, year	Study type	Description of intervention	Populations studied and sample size	Key findings related to intervention effectiveness	Caveats/biases	Supports the intervention? For which populations
Outcome #2: Infant birthweight/premature birth						
1: Hardy et al., 1987	Retrospective cohort Published article	Integrated and structured PCM by RN, SW, and health educator	E=930, C=1,080 Predominantly AA	<36 weeks NS E=11.9%, C=14.2% <2,600g NS, E=15.1%, C=15.3%	Adolescents only	NS difference for GA and LBW
2: Heins et al., 1987	Matched case-control Published article	Structured Monthly HV during pregnancy and postpartum by lay "Resource Mother"	E=575, C=565 89% AA	E had significantly fewer SGA and LBW		Yes
3: Olds et al. 1986; 1988; 1997 [Elmira]	RCT Published article	Structured RN E1-Usual care E2-Free transportation only E3-RN prenatal HV once every 2 weeks, average 9 visits total E4-RN prenatal HV once every 2 weeks decreased frequency of visits for 2 years postpartum	E1=90, E2=94, E3=100, E4=116 C=E1+E2 E=E3+E4 [ethnicity not reported]	BW (grams) NS E=3,285; C=3,262 LBW NS E=5.8%, C=2.6% GA (weeks) NS E=39.6, C=39.7 PTB NS E=6.9%, C=7.3%		Overall no difference in BW or GA, but significant difference for % LBW and % PTB for adolescents
4: Dawson et al., 1989	RCT Published article	Semi-structured Lay/peer HV E1=Routine PNC plus weekly HV E2=Routine PNC, weekly HV, invitation to parenting classes every 2 weeks C1=Routine PNC plus occasional HV C2=Mothers attending the clinic in the months following randomization	E1=42, E2=50 C1=53, C2=27 E=E1+E2 C=C1+C2 [ethnicity not reported]	GA (weeks) NS E=39.2, C=39.7 BW (grams) NS E=3,104; C=3,242	HV group had significantly higher stress levels	No

(continued)

Study #: author, year	Study type	Description of intervention	Populations studied and sample size	Key findings related to intervention effectiveness	Caveats/biases	Supports the intervention? For which populations
5: Korenbrot et al., 1989	Static group comparison Published article	Structured PCM by counselors	*n*=411 [ethnicity not reported] 100% adolescents	LBW rates in school $p<0.05$ Post=8.1% Pre=12.0% mean BW $p<0.0001$	School based, adolescents only	School based PCM is effective for teens
6: Oakley et al., 1990	RCT Published article	Integrated and structured Minimum of 3 HV by nurse midwife	E=255, C=254 Afro-Caribbean/Asian E=11%, C=12%	BW (grams) E=2,944, C=2,907 GA (weeks) E=38.6, C=38.6 [significance not calculated]	E had more spontaneous labor and spontaneous delivery and fewer epidurals. C had more antenatal hospital admits	No
7: Bryce et al., 1991	RCT Published article	Integrated Midwives HV and TC at least every 4–6 weeks	E=983, C=987 White E=90.9%, C=89.3%	% ≥37 weeks NS E=80.8%, C=80.3%		No effect of HV and TC, except for a significantly lower risk of PTB associated with HV and TC in upper class women
8: Graham et al., 1992	RCT Published article	Integrated and structured Paraprofessional 1h HV every 2 weeks, 4 total	E=63, C=58 Predominantly AA	LWB NS E (<4 HV)=12.9%, E (4 HV)=7.7%, C=7.5%		No
9: Villar et al., 1992	RCT Published article	Structured SW, RN 1–2h HV, 4 total	E=1,115; C=1,120 White E=60.3%, C=59.8%	LBW NS OR=0.93, E=8.7%, C=9.4% PTB NS OR=0.88, E=11.1%, C=12.5%	Data also from Belizan et al. (1995) and Langer et al. (1993). Conducted in Mexico, Cuba, Argentina, Brazil	No

11: Julnes et al., 1994	RCT Published article	Integrated and structured E=lay HV C1=clinic-based multidisciplinary team C2=no prenatal care	E=49, C1=46, C2=29 AA E=92%, C=70% C2=80%	BW>2,500g NS E=89.8%, C=93.5% GA<38 weeks NS E=12.2%, C=4.3%		No
12: Zotti and Zahner, 1995	Retrospective cohort Published article	[not reported]	E=97, C=301 White E=55%, C=53% Asian E=17%, C=24% AA E=27%, C=22%	LBW NS chi sq	Very weak intervention	No
14: Norbeck et al., 1996	RCT Published article	Structured RN HV every 2 weeks (4 total) plus TC between sessions	E=56, C=58 100% AA	LBWp=0.04 E=9.1%, C=22.4%	Small sample size; C had low social support	Yes
15: Piper et al., 1996	Retrospective cohort Published article	Structured PCM by RN or SW	E=3,859; C=62,192 AA E=20.9%, C=38.6%	PTB OR=1.3 (1.02,1.25) E=12.6%, C=12.8%		Yes
16: Reichman and Florio, 1996	Retrospective cohort Published article	Integrated	AA E=10,908; C=8,617 White E=13,128; C=8,102AA E=45.4%, C=51.5% White E=54.6%, C=48.5%	AA BW (grams) p<0.01, E=3,114; C=3,040 LBW p<0.01, E=11.4%, C=15.8% White BWp<0.01 E=3,294, C=3,271 LBWp<0.01 E=6.9%, C=8.0%		Participation in HealthStart (E) had a significant effect on BW and LBW for AA and whites
17: Rogers et al., 1996	Retrospective cohort Published article	Structured without protocol Trained lay health workers Monthly HV during pregnancy and 1 year postpartum	E=1901 C1=4,613; C2=712 AA E=77%, C1=54%, C2=70%	LBW NS		Within racial/ethnic groups, no program difference for LBW, but better than expected, given risk

(continued)

Table 10.8 (continued)

Study #: author, year	Study type	Description of intervention	Populations studied and sample size	Key findings related to intervention effectiveness	Caveats/biases	Supports the intervention? For which populations
18: Kitzman et al., 1997; 2000 [Memphis]	RCT Published article	Structured E1=Free transportation only E2=Same as E1 plus screening/ referral services E3=Same as E2 plus RN HV (average 7 during pregnancy and 2 postpartum) E4=Same as E3 plus HV for 2 years postpartum (average 26 visits postpartum)	E1+E2 (Control) =771, E3+E4 (Treatment) =458 Primarily AA	BW NS GA NS		Treatment visited group had fewer miscarriages, fewer gestational infection, lower blood pressure, fewer subsequent pregnancies
19: Tessaro et al., 1997	Prospective cohort Published article	Semi-structured Social support, education and outreach from trained lay health workers 28 weeks to 1 year postpartum	E=1,319; C=9,255 AA E=61.8%, C=59.4%	LBW NS		No
20: Spencer et al., 1989	RCT Published article	Unstructured Lay worker 1–2 visits per week	E=655, C=633 [ethnicity not reported]	BW NS SGA NS GA NS	Women were not screened to see who would be more likely to accept family worker	Addition of family worker had no effect on birth outcomes
21: Baldwin et al., 1998	Ecological Published article	Structured interdisciplinary team in WA state. PCM delivered within a more comprehensive maternity support system E received up to 20 HV or office visits; mean number of prenatal HV 3.7 C group were Medicaid receiving women in CO	WA (1992) state=7,555 but only 378 had PCM only CO (1992) state=5,065 WA (1989) state=6,537 CO (1989) state=4,053 AA WA state (1992)=10.2% CO state (1992)=13.0% WA state (1989)=9.1% CO state (1989)=15.0%	LBW decreased in WA state, NS LBW increased in CO state, NS	Outcomes reported for maternity support services overall, not PCM specifically	Significant program effects in groups with higher risk

22: Gonzalez-Calvo et al., 1997	Descriptive Published article	Structured PRNs HV frequency determined by risk status	AA=210 100% AA	[No usable stats reported]	[No stats on E vs. C]	Found that women with home visiting had more problems solved
23: Lear et al., 1998	Non-equivalent control group Published article	Integrated and structured RN PCM	$E_{low\ risk}$=12 $E_{high\ risk}$=142 C=600 [ethnicity not reported]	[No statistics reported] Women stratified by risk	[No stats on E vs. C]	Appears as though less PTB in low risk group
24: Lowry and Beikirch, 1998	Matched retrospective cohort Published article	Integrated and structured RN and multidisciplinary team, seen at clinic, at least 1 visit	E=175, C=175 AA=23% Hispanic=10%	PTB E=15.4%, C=10.0% GA (weeks) p=0.01, E=37.7, C=38.9 BW (grams) NS E=3,236, C=3,290		White infants heavier than black by 257g. Significant difference for NICU admit (p<0.01) E=13% C=7%
25: Moore et al., 1998	RCT Published article	Structured RN 3 TC per week to 37th week	AA=1,113 E=557, C=556 White=320 E=161, C=159	LBW E=10.9%, C=14.0%, GA<37wks RR=0.87, NS E=9.7%, C=11.0% AA LBW RR=0.75(0.55, 1.00) E=11.3%, C=15.3% GA RR=0.73 (0.52, 1.02) E=9.4%, C=12.8%		Significant reduction in LBW and GA in only AA for women >19 years. No similar effect found in white women
28: Koniak-Griffin et al., 2000	RCT Published article	Structured PHN E=HV during pregnancy and 1 year postpartum (total 17) C=1 or 2 prenatal PHN HV	E=62, C=59 White E=19% C=21%	BW (grams) NS E=3,200, C=3,206 <36 weeks E=3.2%, C=8.5%	Pregnant adolescents only	No significant difference in vaginal delivery or infant hospital stay. E had less hospitalization up to 6 wks (p<0.01) and positive effect on life course

(continued)

Study #: author, year	Study type	Description of intervention	Populations studied and sample size	Key findings related to intervention effectiveness	Caveats/biases	Supports the intervention? For which populations
29: Prozialeck and Pesole, 2000	One group pretest-posttest Published article	Integrated and semi-structured PHN PCM	*n*=27 22% White 52% Hispanic 26% Asian Pre=first pregnancy Post=second pregnancy	BW (grams) Pre = 2,928, Post = 2,523 GA (weeks) Pre = 34.4, Post=39.1	Small *n*. No control group used. Calculated CI and found significance on BW, weight gain, and GA between 1st and 2nd pregnancy	Significant increase in BW, weight gain, and GA in subsequent pregnancies
30: Brooten et al., 2001	RCT Published article	Standardized/structured E=every other medical visit with APN in home C=office based medical care	E = 85, C = 88 AA E = 80%, C = 82%	BW (grams) among PTB E=2,263; C=1,960 *t*-test, $p < 0.05$ GA (weeks) of term infants NS E=38.7, C=38.7	Not usual PCM, similar to enhanced PNC; Among premature and twins, intervention group significantly better	Yes
31: Margolis et al., 2001	Retrospective cohort Published article	Integrated and structured PHN HV 2–4 visits per month	E=105, C=103 AA E=90.3%, C=4.9% White E=14.3%, C=80%	BW (grams) adjusted NS E=3,230, C=3,307 GA (weeks) adjusted NS E=39.3, C=39.4	Interventions also at community and provider levels	No
32: Little et al., 2002	RCT Published article	Structured RN PCM TC	E=61, C=50 Hispanic ethnicity E=27.9%, C=30.0% AA E=42.4%, C=39.6% American Indian E=6.8%, C=12.5%	BW*F*-test 9.8,*p*=0.002; GA*F*-test 2.3,*p*=0.132	PCM saves $500/patient	Telephone PCM had effects on BW

33: Jackson et al., 2003	Prospective cohort Published article	Integrated and structured Doctor and CNM care plus PCM	E=2,156; C=1,577 E=8.4%, C=20.3% Hispanic	<37 weeks NS <2,500g NS E=6.4%, C=6.5%	No separate PCM	No
34: Nguyen et al., 2003	RCT Published article	Structured SW, RN E=weekly HV (1st mo), bi-weekly HV until delivery C=3 HV	E=104, C=121 100% Hispanic	[significance not reported] GA (weeks) E=38.9, C=38.4 BW (grams) E=3,294; C=3,130 <37 weeks E=4.3%, C=8.24	Hispanics only in the study	No
35: Keeton et al., 2004	Retrospective cohort Published article	Unstructured Case manager with bachelor's level degree Services throughout pregnancy and 1–3 years postpartum	E=42,683; C=31,982 AA E=24%, C=52% Hispanic E=31.3%, C=22.7%	VLBW.86 OR (95% CI 0.75, 0.99) E=1.4%, C=2.5% LBW .83 OR (95% CI 0.79, 0.89) E=8.3%, C=12.9%		Yes
36: Carabin et al., 2005	Prospective cohort Published article	Structured RN Weekly visit for 1st mo, bi-weekly to delivery, weekly from post partum visit to 6 wks	E=8,598; C=55,737 AA E=13.6%, C=8.5%	PTB NS LBW Significant PTB Adjusted OR=0.89 (95% CI 0.79, 1.00) LBW Adjusted OR=0.86 (95% CI 0.75, 0.98)	Majority of PCM participants came from WIC referrals	Yes, for LBW
37: Ricketts et al., 2005	Static group comparison Published article	Semi-structured client centered model RN or SW 10 visits including 2 HV	E=2,377; C=68,420 (all births in Colorado) White E=43%, C=61%	[no usable stats reported] Among PCM participants, those with resolved risks had 7.0% LBW, compared to 13.2% for those with unresolved risks ($p<0.001$). E and C were not compared		Group receiving full intervention had more risks resolved. If solved, birth outcomes were better

(continued)

Table 10.8 (continued)

Study #: author, year	Study type	Description of intervention	Populations studied and sample size	Key findings related to intervention effectiveness	Caveats/biases	Supports the intervention? For which populations
38: Sangalang et al., 2006	Retrospective cohort Published article	Structured SW or other health service professional PCM 3–4 times per month	E=1,260; C=1,260 AA E=70.2%, C=48.4% Hispanic E=1.1%, C=4.4%	BW (Normal BW vs. LBW) RR=1.67 (95% CI 1.29, 2.16) LBW$p<0.001$ E=9.4%, C=13.95% GA (full-term vs. PTB) RR=1.69 (95% CI 1.37, 2.09) PTB $p<0.001$ E=15.0%, C=22.6%		Program was protective for GA and BW
39: Silva et al., 2006	Retrospective cohort Published article	Unstructured RN, SW PCM	*N*=6,440 26.7% AA 23.2% Hispanic, 48.2% White	[no usable stats reported] LBW rates were lower for PCM participants than the county		Increased number of visits and/or time with CM did not decrease % LBW births
40: Cramer et al., 2007	Retrospective cohort Published article	Structured SW, RNs E=weekly contacts (HV and PCM)	E_{2002}=79, E_{2003}=157 C_{2002}=7,962; C_{2003}=7,987 [ethnicity not reported]	LBW rate per 100,000 live births E_{2002}=151.9 vs. E_{2003}=95.5 (NS) C_{2002}=76.6 vs. C_{2003}=76.6 (NS)		There was an important but not statistically significant decrease in LBW

Abbreviations: *E* experimental group, *C* control group, *AA* African-American, *HV* home visiting, *TC* telephone calls, *PCM* prenatal case management, *PNC* prenatal care, *NS* not significant, *BW* birthweight, *GA* gestational age, *SGA* small for gestational age, *LBW* low birth weight, *PTB* preterm birth, *VLBW* very low birthweight, *RN* registered nurse, *PHN* public health nurse, *CM* case manager, *CNM* certified nurse midwife

control group. Similarly, Lowry and Beikirch (1998) [24] reported that PTB in the PCM group was 15.4% but only 10.0% in the control group. Even Olds et al., (1988, 1997) [3] reported a higher LBW rate in the PCM group (5.8%) compared to the control group (2.6%) in two of his randomized clinical trials. The most frequent explanation was that the PCM group had a higher level of risk. While this explanation seems plausible, these same studies had screened participants and, in general, few significant differences were reported between the PCM and comparison groups on demographic and health risk characteristics.

Of the two studies with infant outcomes that specifically reported neonatal effects of PCM by race/ethnicity, two [16, 25] found more favorable outcomes among minority participants compared to whites. These studies are described below.

The study by Moore et al. (1998) [25] was a randomized clinical trial with retrospective medical record abstraction conducted without knowledge of group assignment. The researchers found that African-American women who received PCM had a reduced rate of LBW (11.3 vs. 15.3%), with a corresponding relative risk of.75 (95% CI 0.55, 1.00), and fewer neonates of gestational age less than 37 weeks (9.4 vs. 12.8%; RR=0.73; 95% CI 0.52, 1.02) compared to women who did not receive PCM. These improvements for African-American women only became significant when the sample was stratified by age (i.e., 18 years old or less and 19 years or older). The study by Reichman and Florio (1996) [16] was a retrospective cohort design using vital records. They found a significant improvement in birth weight for both African-American and white infants born to women who had received PCM.

Summary of the Evidence

Study Quality

Overall, the quality across the 39 primary studies was moderate (Table 10.5), when considering issues of validity, reliability, and power. The 18 randomized clinical trials or prospective cohort studies had fair to high quality, with quality scores ranging from 11 to 28 (mean=19.1; SD=4.4; median=19.5). There were four quasi-experimental studies [5, 23, 29, 37], with quality scores of 19, 7, 9, and 21 respectively. The retrospective studies of various designs used Medicaid claims or medical records and vital records data, and varied widely in quality. The influence of study design on the overall study quality must be taken into consideration when weighing the evidence with respect to PCM's effect on maternal or neonatal outcomes. There were no descriptive qualitative studies included in the review, as the qualitative studies reviewed did not report on either prenatal care use or on birth outcomes.

PCM and Disparities

The major difficulty in assessing the effectiveness of PCM in reducing disparities in birth outcomes is that so few studies reported outcomes by the race/ethnicity of study participants. Nonetheless, the sparse evidence is modestly encouraging. Of the ten studies with at least 75% African-American participants (Table 10.1), four studies showed positive statistically significant outcomes for women receiving PCM [2, 14, 25, 30]. There were too few studies with greater than 75% Latina or Asian participants to draw any conclusions with regard to the effectiveness of PCM for these minority groups. Of the 19 studies with at least 50% of the treatment or experimental group from a minority group (African-American or Latina), three (16%) reported significant improvement in prenatal care

utilization [1, 2, 17], and 4 (21%) reported improvements in LBW [2, 14, 30, 38]. Given that minorities were overrepresented in several study samples, there is some indirect evidence of the positive effect of PCM on birth outcomes among women of color. This suggests that PCM may indeed have a positive effect on decreasing disparities in birth outcomes if uptake is sufficient. It is worth noting that study participants received PCM because they were at risk for adverse birth outcomes. In other words, the modest effects of PCM were achieved for women for whom healthy births would have been the most difficult to achieve.

The data from the studies reviewed do not allow us to conclude, or even speculate, as to which vulnerable groups would benefit most from PCM. We can look, however, at the comments of the authors; several authors discussed their findings in terms of women having reduced risk factors that contribute to adverse birth outcomes as a result of PCM. One interpretation of this is that PCM effects may be more long term. Our review did not attempt to systematically identify or synthesize findings for subsequent pregnancies because only one study in our review [38] reported outcomes for subsequent pregnancies. Sangalang, Barth, and Painter (2006) [38] did not find a significant difference in interconceptional spacing, except among adolescents; their findings were not reported by race/ethnicity.

Commentary on the Outcome Variables Studied

One of the more fascinating findings from this literature review is the breadth of outcomes studied in association with PCM. We chose to focus only on utilization of prenatal care and birthweight/gestational age because those outcomes were reported in sufficient number to draw tentative conclusions, and because the focus of this book is reproductive and perinatal outcomes. The various other outcomes reported and relevant to the reproductive and perinatal period may hold promise as more sensitive indicators of the effects of PCM, but a larger number of studies are needed using the same outcome indicators before we can say whether PCM has an effect on other maternal health indicators, such as hematocrit, or behaviors, such as smoking cessation. We did note that the clinical significance in terms of meaningful health improvements, in addition to the statistical significance of improvements, was not discussed in any of studies reviewed.

A significant relationship was found between receiving PCM and improved neonatal health status in a few, high quality studies [7, 14, 15, 25] (using the quality scores shown in Table 10.5). Aside from these studies, data are lacking for strong support of PCM to improve neonatal health status as measured by birth weight and gestational age. However, these outcome measures are not uniquely sensitive or specific to PCM. Birthweight and gestational age are both influenced by a vast array of factors, from maternal nutrition to competency of medical care to socio-economic conditions. In addition, it is quite likely that the relative contribution of PCM to a change in these neonatal outcomes would be small. Despite these caveats, 14 (35%) studies (Table 10.4) were able to detect a statistically significant positive difference in neonatal outcomes among women who had received PCM.

Recommendations

Practice Recommendations

In general, although not all of the studies were high quality, PCM had a positive, but not always significant effect on birth weight and LBW rates in of the majority of the studies that reported that

outcome, and on PTB and gestational age in almost half the studies reporting that outcome. These findings suggest that PCM does have the potential to improve the health of neonates.

No study, in which gestational age at entry was provided, included participants who began PCM in the first trimester. In the majority of studies in which PCM was initiated in the second trimester there were positive outcomes, although preterm delivery bias is a potential threat to validity. Such bias notwithstanding, it seems that receiving some PCM is more beneficial than not receiving any PCM. With regard to the content of PCM, it is difficult to make evidence-based recommendations for tailoring specific interventions to be included in a PCM program because of the inconsistent and sparse descriptions of interventions used in PCM programs. As such, we recommend that going forward, PCM programs adopt standardized intervention protocols. It may also be necessary to have state- and federal-level policies that reinforce the standardization of PCM programs in terms of frequency of contact, interventions used, and the qualifications of case managers providing PCM.

There was no discernible pattern in effectiveness that distinguished between PCM programs that were integrated into existing prenatal care services versus those that were stand-alone programs, when comparing Tables 10.2 and 10.4. Similarly, there was no apparent pattern of effectiveness with regard to whether the PCM program was standardized or not standardized. Thus, we are reluctant to make recommendations for the appropriate intervention protocol or for the optimal organizational structure for PCM. With regard to RN versus non-RN case managers, the evidence is similarly equivocal. Perhaps a more relevant issue is the quality of the inter-personal relationship and client trust, in addition to the professional training of case managers.

Research Recommendations

Studies of PCM have tended to be of poor to fair quality, with a few notable exceptions. Additionally, the studies reviewed failed to address four areas critical to a more complete evaluation of PCM effectiveness: (1) match of intervention to client needs; (2) intervention standardization and dosage; (3) outcome specificity and sensitivity; and, (4) qualifications of the case manager. Each of these is discussed in turn.

Most of the studies reviewed mentioned that women were screened for eligibility for PCM and described some efforts within the program to tailor the interventions to the needs of the women. The screening used to assign women to PCM generally included social, economic, medical, psychological, and environmental factors. In short, the screening approach was generic and there was no apparent attempt in the studies (or programs) to assign women to different levels of PCM intensity. None of the studies we reviewed used specific interventions, such as referrals or counseling, as one of the variables to predict effectiveness. Only one study stratified the women by degree of risk or by any specific psycho-social risk factor [23]. The lack of attention to measuring interventions, and to matching interventions with specific psycho-social risk, leaves us with little data to identify the groups of women for whom PCM is most beneficial. In other words, we need further research to understand whether PCM has differential effects for different vulnerable groups.

Intervention standardization and dosage remain an issue. Only 26 of the 39 studies had a standardized/structured intervention protocol, 8 of which were integrated into prenatal care (Table 10.2). There seemed to be a heavy reliance on the judgment of clinicians or the trained lay paraprofessional to know what to do for the women. Thus, we need additional research to know whether implementation of PCM protocols specifying various levels of intervention intensity might have differential effects. A somewhat surprising finding of this literature review was the large number of PCM programs that were integrated into existing prenatal care systems. However, none of the studies compared an integrated with a non-integrated approach. Thus, we do not know whether PCM as a stand-alone program would have a greater or lesser effect than PCM provided as an integrated element

of prenatal care. These structural issues, as well as the effect of payment structures and incentives on the PCM program, remain fertile areas for future research.

Across the studies, there did not seem to be a systematic nor theoretical link between the interventions that comprise PCM and the maternal or neonatal outcomes reported. Rather, the studies seemed to rely on the availability of existing maternal and neonatal health data when reporting outcomes. Graham and Campbell (1992) referred to this as the measurement trap, in which available data are used rather than measures that might be more appropriate. A few studies did report on more discrete maternal health outcome indicators, such as hematocrit (Hardy et al., 1987), or on life-course outcome indicators, such as interconceptional spacing (Sangalang et al., 2006). It appeared that the findings using those outcome variables were favorable for PCM. What remains unknown is which maternal and neonatal outcome indicators are most specific and sensitive to potential effects which might be attributable to PCM, and for which racial/ethnic groups.

The last area that deserves further research is the qualifications of case managers and the composition of the PCM team. Given that the professional RN BSN model for PCM has been adopted by so few programs, and that the nation is facing a nursing shortage, research is needed to better understand the circumstances under which combinations of health and service professionals and paraprofessionals will yield the most positive health outcomes for the women receiving PCM. One aspect of human resources that was not addressed in any study is the cultural competence and interpersonal skills of the case managers. Attention to cultural competence seems particularly relevant given that PCM includes home visiting, a highly personal and interactive intervention.

Theory Building Recommendations

For all of the areas needing further research, a strong, explicit theoretical framework is needed. The few studies that did report or mention a theory on which the PCM program was based tended to use an ecological framework. While this is appropriate as a conceptual framework, it does not assist in understanding how the actual interventions lead to maternal health changes important to improving both infant and birth outcomes. Issel (2000), based on a qualitative study of women who received PCM, proposed a conceptual model that links psychosocial variables addressed by case managers to possible birth improvements. This model deserves refinement and subsequent testing. In particular, variables related to culture and context need to be incorporated into theories that might predict which aspects of PCM would be most effective for which maternal risks, as well as for which ethnic and cultural groups. In this way, theory building efforts might lead to the design and development of more effective PCM programs.

Policy Recommendations

Most importantly, the intent of PCM must be reframed in the minds of clinicians and policy makers from averting adverse birth outcomes to improving the maternal life course. The women who receive PCM are at high social and economic risk, meaning that they have complicated lives with problems that are mostly recalcitrant to short-term intervention. Thus, reframing the intent of PCM to a life-course focus is more in keeping with the potential effects of counseling, referrals, support, continuity of healthcare, and long-term monitoring, all of which are part of PCM.

Conclusion

Providing pregnant women who are at high-risk for adverse pregnancy outcomes with case management generally will result in improvement in the use of prenatal care and infant birth outcomes. However, the magnitude of these effects is likely to be small. No pattern was identifiable for variables such as study quality, structure of the PCM program, or trimester during which PCM began. Similarly, due to lack of study specification, we are reluctant to state that PCM has a definitive benefit for women from diverse racial/ethnic minority groups, although the evidence is encouraging. In addition, despite the considerable number of studies of PCM, our cumulative knowledge continues to have gaps, particularly about which elements of PCM are most effective with which sub-groups of high risk pregnant women, and, especially, how best to deliver PCM.

Acknowledgements We wish to express great appreciation to Melissa Sherwin and Anna Wiencrot for their help in preparing this manuscript and to Arden Handler for her careful editing.

References

American Nurses Association. (1997). *Home care for mother, infant and family following birth.* Accessed 3 July 2003, from www.nursingworld.org.

Association of Maternal and Child Health Programs (AMCHP) (1999). *FactSheet: Home visiting: An effective strategy for improving the health of mothers and children*. Accessed 3 July 2003 from www.amchp.org.

[21] Baldwin, L., Larson, E. H., Connell, F. A., Nordlund, D., Cain, K. C., Cawthon, M. L., et al. (1998). The effect of expanding Medicaid prenatal services on birth outcomes. *American Journal of Public Health, 88*, 1623–1629.

[9] Belizan, J., Barros, F., Langer, A., Farnot, U., Victora, C., Villar, J., et al., (1995). Impact of health education during pregnancy on behavior and utilization of health resources. *American Journal of Obstetrics & Gynecology, 173*, 894–899.

[10] Bradley, P. J., & Martin, J. (1994). The impact of home visits on enrollment patterns in pregnancy related services among low-income women. *Public Health Nursing, 11*, 392–398.

[30] Brooten, D., Youngblatt, J. M., Brown, L., Finkler, S. A., Neff, D. F., & Madigan, E. (2001). A randomized trial of nurse specialist home care for women with high-risk pregnancies: Outcomes and costs. *American Journal of Managed Care, 7*, 793–803.

[7] Bryce, R. L., Stanley, F. J., & Garner, J. B. (1991). Randomized control trial of antenatal social support to prevent preterm birth. *British Journal of Obstetrics & Gynaecology, 98*, 1001–1008.

[36] Carabin, H., Cowan, L. D., Beebe, L. A., Skaggs, V. J., Thompson, D., & Agbangla, C. (2005). Does participation in a nurse visitation programme reduce the frequency of adverse perinatal outcomes in first-time mothers? *Pediatric and Perinatal Epidemiology, 19*, 194–205.

Ciliska, D., Hayward, S., Thomas, H., Mitchell, A., Dobbins, M., Underwood, J., et al., (1996). A systematic overview of effectiveness of home visiting as a delivery strategy for public health nursing interventions. *Canadian Journal of Public Health, 87*(3), 193–198.

Ciliska, D., Masgtrilli, P, Ploeg, J., Brunton, G., & Underwood, J. (2001). The effectiveness of home visiting as a delivery strategy for public health nursing interventions to clients in the prenatal and postnatal period: A systematic review. *Primary Health Care Research and Demonstration, 2*, 41–54.

Colorado State Board of Health (2000). *Statement of basis and purpose and specific statutory authority: Nurse home visitor program rules*. Accessed 3 July 2003 from www.cdphe.state.co.us/op/regs/preventionservices/101601nursehomevisitorprogram.pdf.

Cooper, H. (1998). *Synthesizing research* (3rd ed.). Thousand Oaks, CA: Sage.

[40] Cramer, M. E., Chen, L., Roberts, S., & Clue, D. (2007). Evaluating the social and economic impact of community based prenatal care.*Public Health Nursing, 24*, 329–336.

[4] Dawson, P., Van Doorninck, W. J., & Robinson, J. (1989). Effects of home-based, informal social support on child health. *Developmental and Behavioral Pediatrics, 10*, 63–67.

Dewa, C. S., Horgan, S., McIntyre, S., Robinson, G., Krupa, T., & Eastabrook, S. (2003). Direct and indirect time inputs and assertive community treatment. *Community Mental Health Journal, 39*, 17–32.

Dolan-Mullen, P., DiClemente, C. C., Velásquez, M. M., Timpson, S. C., Groff, J. Y., Carbonari, J. P., et al. (2006). Enhanced prenatal case management for low income smokers. *Tobacco Control, 9*, iii75–iii77.

Downs, S. H., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomized and non-randomized studies of health care interventions. *Journal of Epidemiology and Community Health, 52*, 377–387.

Fraser, J. A., Armstrong, K. L., Morris, J. P., & Dadds, M. R. (2000). Home visiting intervention for vulnerable families with newborns: Follow-up results of a randomized controlled trial. *Child Abuse Neglect, 24*, 1399–1349.

Goldfarb, N. I., Hillman, A. L., Eisenberg, J. M., Kelley, M. A., Cohen, A. V., & Dellheim, M. (1991). Impact of mandatory Medicaid case management program on prenatal care and birth outcomes. *Medical Care, 29*, 64–71.

Gomby, D. S., Culross, P. L., & Behrman, R. E. (1999). Home visiting: Recent program evaluations – Analysis and recommendations. *The Future of Children, 9*(1), 4–26.

[22] Gonzalez-Calvo, J., Jackson, J., Hanford, C., Woodman, C., & Remington, N. S. (1997). Nursing case management and its role in perinatal risk reduction: Development, implementation, and evaluation of a culturally competent model for African-American women. *Public Health Nursing, 14*, 190–206.

Graham, W., & Campbell, O. (1992). Maternal health and the measurement trap. *Social Science and Medicine, 35*, 967–977.

[8] Graham, A. V., Frank, S. H., Zyzanski, S. J., Kitson, G. C., & Reeb, K. G. (1992). A clinical trial to reduce the rate of low birth weight in an inner-city black population. *Family Medicine, 24*, 439–446.

Hall, J. A., Carswell, C., Walsh, E., Huber, D. L., & Jampoler, J. S. (2002). Iowa case management: Innovative social work. *Social Work, 47*, 132–141.

[1]Hardy, J. B., King, T. M., & Repke, J. T. (1987). The Johns Hopkins adolescent pregnancy program: An evaluation. *Obstetrics and Gynecology, 69*, 300–306.

[2] Heins, H. C., Nance, N. W., & Ferguson, J. E. (1987). Social support in improving perinatal outcome: The Resource mother program. *Obstetrics & Gynecology, 70*, 263–266.

Hodnett, E. D., & Fredericks, S. (2003). Support during pregnancy for women at increased risk of low birthweight babies. *The Cochrane Database of Systematic Reviews*, CD000198.

Issel, L. M. (2000). Women's perceptions of outcomes from prenatal case management. *Birth, 27*, 120–126.

[33] Jackson, D. J., Lang, J. M., Swartz, W. H., Ganitas, T. G., Fullerton, J., Ecker, J., et al. (2003). Outcomes, safety, and resource utilization in a collaborative care birth center program compared with traditional physician-based perinatal care. *American Journal of Public Health, 93*, 999–1006.

[11] Julnes, G., Konefal, M., Pindur, W., & Kim, P. (1994). Community-based perinatal care for disadvantaged adolescents: Evaluation of the resource mothers program. *Journal of Community Health, 19* (1), 41–53.

[35] Keeton, K., Saunders, S., & Koltun, D. (2004). The effect of the family case management program on 1996 birth outcomes in Illinois. *Journal of Women's Health, 13*, 207–215.

Kendrick, D., Elkan, R., Hewitt, M., Dewey, M., Blair, M., Robinson, J., et al. (2000). Does home visiting improve parenting and the quality of the home environment? A systematic review and meta-analysis. *Archives of Disease in Childhood, 82*, 443–451.

[18] Kitzman, H., Olds, D. L., Henderson, C. R., Hanks, C., Cole, R., Tatelbaum, R., et al. (1997). Effect of prenatal and infancy home visitation by nurses on pregnancy outcomes, childhood injuries, and repeated childbearing. *Journal of American Medical Association, 278*, 644–652.

[18] Kitzman, H., Olds, D. L., Sidora, K., Henderson, C. R., Hanks, C., Cole, R., et al. (2000). Enduring effects of nurse home visitation on maternal life course: A 3-year follow-up of a randomized trial. *Journal of the American Medical Association, 283*, 1983–1989.

[28] Koniak-Griffin, D., Anderson, N. L. R., Verzemnieks, I., & Brecht, M. L. (2000). A public health nursing early intervention program for adolescent mothers: Outcomes from pregnancy through 6 weeks postpartum. *Nursing Research, 49*, 130–138.

[5] Korenbrot, C., Gill, A., & Zoe-Patterson, E. (1995). Evaluation of California's statewide implementation of enhanced prenatal services as Medicaid benefits. *Public Health Reports, 110*, 125–133.

Korenbrot, C. C., Showstack, J., Loomis, A., & Brindis, C. (1989). Birth weight outcomes in a teenage pregnancy case management program. *Journal of Adolescent Health Care, 10*, 97–104.

Korfmacher, J., O'Brien, R., Hyatt, S., & Olds, D. (1999). Differences in program implementation between nurses and paraprofessionals providing home visits during pregnancy and infancy: A randomized trial. *American Journal of Public Health, 89*, 1847–1851.

Landis, M., & Willems Van Dijk, J. (2006). *Facts on the effectiveness of prenatal care coordination: To what extent does evidence suggest enhanced prenatal care services contribute to producing healthy birth outcomes?* Unpublished report from Department of Health and Family Services, WI, USA.

[9] Langer, A., Victora, C., Victora, M., Barros, F., Farnot, U., Balizan, J., et al. (1993). The Latin American trial of psychosocial support during pregnancy: A social intervention evaluated through an experimental design. *Social Science and Medicine, 36*, 495–507.

[23] Lear, D., Schall, L. C., Marsh, G. M., Liu, K. S., & Yao, Y. (1998). Identification and case management in an HMO of patients as risk of preterm labor. *American Journal of Managed Care, 4*, 865–871.

Lee, D. T. F., Mackenzie, A. E., Dudley-Brown, S., & Chin, T. M. (1998). Case management: A review of the definitions and practices. *Journal of Advanced Nursing, 27*, 933–939.

[32] Little, M., Saul, G. D., Testa, K., & Gaziano, C. (2002). Improving pregnancy outcome and reducing avoidable clinical resource utilization through telephonic perinatal care coordination. *Lippincott's Case Management, 7*, 103–112.

Loda, F. A., Speizer, I. S., Martin, K. L., Skatrud, J. D., & Bennett, T. A. (1997). Programs and services to prevent pregnancy, childbearing, and poor birth outcomes among adolescents in rural areas of the Southeastern United States. *Journal of Adolescent Health, 21*, 157–166.

[24] Lowry, L. W., & Beikirch, P. (1998). Effect of comprehensive care on pregnancy outcome. *Applied Nursing Research*, 11(2), 55–61.

Lumley, J. (1991). Preventing and managing prematurity. *International Journal of Technology Assessment in Health Care, 7*, 460–477.

[31] Margolis, P. A., Stevens, R., Bordley, W. C., Stuart, J., Harlan, C., Keyes-Elstein, L., et al. (2001). From concept to application: The impact of community-wide intervention to improve the delivery of preventive services to children. *Pediatrics, 108*(3); URL: http://www.pediatrics.org/cgi/content/full/108/3/e42.

[25] Moore, M. L., Meis, P. J., Ernest, J. M., Wells, H. B., Zaccaro, D. J., & Terrell, T. (1998). A randomized trial of nurse intervention to reduce preterm and low birth weight births. *Obstetrics & Gynecology, 91*, 656–661.

Mueser, K. T., Bond, G. R., Drake, R. E., & Resnick, S. G. (1998). Models of community care for severe mental illness: a review of research on case management. *Schizophrenia Bulletin, 24*, 37–74.

[26] Newschaffer, C. J., Cocroft, J., Hauck, W. W., Fanning, T., & Turner, B. J. (1998). Improved birth outcomes associated with enhanced Medicaid prenatal care in drug-using women infected with the human immunodeficiency virus. *Prenatal Care in HIV-Infected Drug Users, 91*, 885–891.

[34] Nguyen, J. D., Carson, M. L., Parris, K. M., & Place, P. (2003). A comparison pilot study of public health field nursing home visitation program interventions for pregnant Hispanic adolescents. *Public Health Nursing, 20*, 412–418.

[14] Norbeck, J. S., Dejoseph, J. F., & Smith, R. T. (1996). A randomized trial of an empirically-derived social support intervention to prevent low birthweight among African American women. *Social Science Medicine, 43*, 947–954.

Northeast Florida Healthy Start Coalition, Inc. (n.d.). *The Magnolia Project: Empowering women's health and wellness*. Accessed 20 July 2006 from www.unf.edu/~jwill/Magnolia%20Final%20Report%201999–2001.pdf.

Nurse Family Partnership. (n.d.). Accessed 9 September 2010 from www.nursefamilypatnership/about/fact-sheets.

[6] Oakley, A., Rajan, L., & Grant, A. (1990). Social support and pregnancy outcome. *British Journal of Obstetrics and Gynaecology, 97*, 155–162.

[3] Olds, D. L., Eckenrode, J., Henderson, C. R., Kitzman, H., Powers, J., Cole, R., et al. (1997). Long-term effects of home visitation on maternal life course and child abuse and neglect. *Journal of the American Medical Association, 278*, 637–643.

[3] Olds, D. L., Henderson, C. R., Tatelman, R., & Chamberlin, R. (1986). Improving the delivery of prenatal care and outcomes of pregnancy: A randomized trial of nurse home visitation. *Pediatrics, 77*, 16–28.

[3] Olds, D. L., Henderson, C. R., Tatelman, R., & Chamberlin, R. (1988). Improving the life-course development of socially disadvantaged mothers: A randomized trial of nurse home visitation. *American Journal of Public Health, 78*, 1436–1445.

Olds, D. L., & Kitzman, H. (1990). Can home visitation improve the health of women and children at environmental risk? *Pediatrics, 86*, 108–116.

Olds, D., & Korfmacher, J. (1997). The evolution of a program of research on prenatal and early childhood home visitation: Special issue introduction. *Journal of Community Psychology, 25*, 1–7.

Olds, D. L., Robinson, J., O'Brien, R., Luckey, D. W., Petitt, L. M., Henderson, C. R., et al., (2003). Home visiting by paraprofessionals and nurses: A randomized, controlled trial. *Pediatrics, 110*, 286–496.

[15] Piper, M., Mitchel, E. F., & Ray, W. A. (1996). Evaluation of a program for prenatal care case management. *Family Planning Perspective, 28*, 65–68.

Poland, M. L., Giblin, P. T., Waller, J. B., & Bayer, I. S. (1991). Development of a paraprofessional home visiting program for low-income mothers and infants. *American Journal of Preventive Medicine, 7*, 204–207.

[29] Prozialeck, L. L., & Pesole, L. (2000). Performing a program evaluation in a Family Case Management Program: Determining outcomes for low birthweight deliveries. *Public Health Nursing, 12*(3), 195–201.

Ramey, C. T., & Ramey, S. L. (1993). Home visiting programs and the health and development of young children. *The Future of Children, 3*, 129–139.

[16] Reichman, N. E., & Florio, M. J. (1996). The effects of enriched prenatal care services on Medicaid birth outcomes in New Jersey. *Journal of Health Economics, 15*, 455–476.

[37] Ricketts, S. A., Murray, E. K., & Schwalberg, R. (2005). Reducing low birthweight by resolving risks: Results from Colorado's Prenatal Plus Program. *American Journal of Public Health, 95*, 1952–1957.

Roberts, I., Kramer, M. S., & Suisse, S. (1996). Does home visiting prevent childhood injury? A systematic review of randomized controlled trials. *British Medical Journal, 312*(7022), 29–33.

[17] Rogers, M. M., Peoples-Sheps, M. D., & Suchindran, C. (1996). Impact of a social support program on teenage prenatal care use and pregnancy outcomes. *Journal of Adolescent Health, 19*, 132–140.

[38] Sangalang, B. B., Barth, R. P., & Painter, J. S. (2006). First-birth outcomes and timing of second births: A statewide case management program for adolescent mothers. *Health and Social Work, 31*(1), 55–63.

[39] Silva, R., Thomas, M., Caetano, R., & Araganki, C. (2006). Preventing low birth weight in Illinois: Outcomes of the Family Case Management Program. *Maternal and Child Health Journal, 10*(6), 481–488.

[20] Spencer, B., Thomas, H., & Morris, J. (1989). A randomized control trial of the provision of a social support service during pregnancy: the South Manchester Family Worker Project. *British Journal of Obstetrics and Gynaecology, 96*, 281–288.

Stankaitis, J. A., Brill, H. R., & Walker, D. M. (2005). Reduction in neonatal intensive care unit admission rates in Medicaid Managed Care program. *The American Journal of Managed Care, 11*, 166–172.

Stevens-Simons, C., & Orleans, M. (1999). Low-birthweight prevention programs: The enigma of failure. *Birth Issues in Perinatal Care, 26*, 184–191.

St. Pierre, R., & Layzer, J. I. (1999). Using home visits for multiple purposes: The comprehensive Child Development Program. *The Future of Children, 9*, 134–151.

[19] Tessaro, I., Campbell, M., O'Meara, C., Herrick, H., Buescher, P., Meyer, R., et al. (1997). State health department and university evaluation of North Carolina's Maternal Outreach Worker Program. *Research Linkages Between Academia and Public Health Practice, 13*(Suppl 2), 38–44.

[27] Thompson, M., Curry, M. A., & Burton, D. (1998). The effect of nursing case management on the utilization of prenatal care by Mexican-Americans in rural Oregon. *Public Health Nursing, 15*, 82–90.

[9] Villar, J., Farnot, U., Barros, F., Victoria, C., Langer, A., & Balizan, J. (1992). A randomized trial of psychosocial support during high-risk pregnancies. *New England Journal of Medicine, 327*, 1266–1271.

Zimmer-Gembeck, M. J., & Helfan, M. (1995). Low birthweight in a public prenatal care program: Behavioral and psychosocial risk factors and psychosocial intervention. *Social Science Medicine, 43*, 187–197.

[12] Zotti, M. E., & Zahner, S. J. (1995). Evaluation of public health nursing home visits to pregnant women on WIC. *Public Health Nursing, 12*, 294–304.

Chapter 11
Behavioral Treatment Methods for Pregnant Smokers: The Evidence Base for Prenatal Care Programs and Professional Practice

Richard Windsor

Overview

This chapter presents a critical review of salient issues related to the selection, planning, delivery, and evaluation of evidence-based assessment and health education-counseling methods for pregnant smokers. The chapter aims are to present: (a) a synopsis of the epidemiologic evidence documenting the adverse effects caused by active and passive tobacco exposure; (b) smoking prevalence rates and trends during pregnancy for the U.S., including a discussion about the poor validity of measurement of smoking status by patient self-reports; (c) a comprehensive review of the evidence to document the levels of "effectiveness" of treatment methods; (d) a synopsis of the evidence about the use of Nicotine Replacement Therapy (NRT) by pregnant smokers; (e) a description of an evidence-based Smoking Cessation and Reduction In Pregnancy Treatment (SCRIPT) Program that can be routinely provided as a component of prenatal care; and, (f) a synthesis of the evidence base supporting adoption of a comprehensive, system-wide tobacco treatment program designed for adoption and dissemination by large, well-defined populations of pregnant smokers: SCRIPT PLUS. The final section of this chapter provides a discussion of future research needs and program challenges. Recommendations for the dissemination-adoption of evidence-based methods for smoking cessation during pregnancy are made for consideration by professional practice and program development and evaluation leadership.

Tobacco Smoke Ingredients

Tobacco smoke contains more than 4,000 chemicals (Hoffman, Hoffman, & Wynder, 1998). The mother, fetus, infant, and others in close proximity are actively and passively exposed daily, in varying degrees, to all of these toxic chemicals and carcinogens. The gaseous combustion phase of tobacco smoke production contains numerous chemical toxins such as urethanes, ammonia, arsenic, formaldehyde, hydrogen cyanide, and a large number of specific carcinogenic and tumorigenic agents. The particulate phase of tobacco smoke produces an additional variety of other toxins. The adverse effects on maternal, fetal, and infant health produced by daily exposure to each single toxic agent, and especially the combination and synergistic interaction of multiple chemical compounds, before,

R. Windsor (✉)
Department of Prevention and Community Health, School of Public Health and Health Services, George Washington U. Medical Center, 2175 K St. NW # 701 202-416-0086, Washington, DC 20037, USA
e-mail: sphraw@gwumc.edu

A. Handler et al. (eds.), *Reducing Racial/Ethnic Disparities in Reproductive and Perinatal Outcomes: The Evidence from Population-Based Interventions*, DOI 10.1007/978-1-4419-1499-6_11,

during, and after a pregnancy is unequivocal (Benowitz, 1991, 1998; Benowitz & Dempsey, 2004). Active smoking during pregnancy, because the causal evidence is conclusive, has been defined as the most serious and preventable cause of fetal and infant morbidity and mortality by the U.S. Healthy People 1990, 2000, and 2010 Objectives for the Nation (U.S. Department of Health, Education & Welfare, 1979; U.S. Department of Health & Human Services, 1991, 2000).

While the mechanisms through which the constituents of tobacco smoke cause specific adverse outcomes are not completely known, it is well established that nicotine and/or carbon monoxide (CO) alter fetal development and produce specific teratogenic effects on the nervous system and neural development (Benowitz, 1998; Karen, Eliopoulis, & Klein, 1998; Oncken, Hardardottir, & Smelzer, 1998; Slotkin, 1998b). Cyanide from combustion and inhalation contributes to impaired fetal growth and increased morbidity and mortality. Acute and chronic fetal hypoxia is most likely produced when CO from inhaled smoke binds to fetal hemoglobin or when placental blood flow is reduced by nicotine exposure. Fetal hypoxia has been implicated in growth restriction and neurobehavioral deficits in infants of smokers and has been linked to SIDS. It has been reported that daily fetal exposure to CO, nicotine, and other tobacco smoke chemicals and carcinogens is worse than cocaine-related fetal exposures (Slotkin, 1998a, 1998b).

Tobacco Exposure and Maternal, Fetal, and Infant Risks

In 1957, Simpson published the first study which evaluated and confirmed the dose–response relationship between cigarette smoking and infant low birth weight (LBW). She reported that the LBW rate of infants of mothers who smoked (11.0%) was approximately double the rate of non-smokers (5.6%) and that pregnant smokers had babies about 200 g lighter than non-smokers. Since this initial study of 7,500 patients in California, a large number of methodologically more rigorous, prospective epidemiologic studies, involving millions of subjects, and current meta-analyses of this literature, have provided conclusive evidence for an independent, adverse causal effect of daily smoking on multiple perinatal outcomes. Infants of non-smokers are consistently 200+ grams heavier on average when compared to infants of smokers, adjusting for covariates of birth weight (odds ratio ≥2.0; Dewan, Brabin, Wood, Dramond, & Cooper, 2003; Hruba & Kachlik, 2000; Hughes & Brennan, 1996; Meyer, Jonas & Tonascia, 1976; Misra & Nguyen, 1999). Although self-reports of the number of cigarettes per day (CPD) is an inaccurate measure of daily exposure, multiple studies that included a biochemical test to confirm patient self-reports of smoking status have reported a dose–response relationship between CPD and infant birth weight. Estimates of a 12 g (England et al., 2001; Mathai, Skinner, Lawton, & Weindling, 1990), 10 g (Li, Windsor, Perkins, Goldenberg, & Lowe, 1993), and a 27 g (Bernstein et al., 2005) reduction in birth weight for each additional cigarette smoked per day have been documented.

The deleterious effects of tobacco exposure during pregnancy include female and male infertility, placental abruption, placenta previa, fetal growth restriction/LBW, premature delivery, SIDS, spontaneous abortion, and decreased lung function (Augood, Duckitt, & Templeton, 1998; Castles, Adams, Melvin, Kelsch, & Boulton, 1999; Hasselmeyer, Meyer, Catz, & Longo, 1979; Kleinman, Pierre, Madans, Land, & Schramm, 1988; Simpson, 1957; Walsh, 1994). If smoking during pregnancy were eliminated, it has been estimated that there would be 30% fewer low birth weight infants, 10% fewer preterm infants, and 5% fewer infant deaths. Adjusting for other salient predictors of negative perinatal outcomes, if a cohort of 1,000 smoking pregnant women were to quit in their first or early second trimester, this cohort would have a very high probability of having rates of adverse perinatal outcomes equal to a non-smoking cohort of 1,000 women with equivalent initial risk levels.

The harmful effects on fetal and infant growth and development from maternal and infant exposure to environmental tobacco smoke (ETS), especially heavy, daily ETS exposure during the third trimester, have also been confirmed. "Reviews of the literature have documented a decrease of approximately 25–30 g in birth weight and a weighted pooled relative risk estimate of 1.2 associated with ETS exposures [National Cancer Institute (NCI), 1999; Windham, Eaton, & Hopkins, 1999]". Kharrazi et al. (2004) conducted one of the most rigorous studies among pregnant women to precisely measure the risks associated with ultra-low levels of ETS exposure. Analyzing effects within five categories ranging from <0.026 to >0.235 ng/mL, they estimated that ETS levels ≥0.05 ng/mL accounted for 12% of all adverse perinatal health effects among 3,150 Medicaid and non-Medicaid funded pregnancies from 15 counties in California.

ET exposure, in-utero and especially in the first year of infancy, also causes infant respiratory diseases and potentially long-term irreversible decrements in lung function. Meta-analyses of the evidence on the risks caused by infant in-utero and post-partum ETS exposure by the U.S. Working Group on Passive Smoking (Spitzer et al., 1990), the NCI (1999), and other reviews (Kharrazi et al., 2004) concluded that: (a) there is strong evidence that ETS is a primary cause of severe respiratory illnesses, pediatric asthma, and ear inflammation; (b) there is convincing evidence that ETS causes multiple chronic and acute respiratory illnesses; and, (c) small reductions in physiological measures of respiratory function have been documented among both children and adults exposed to residential ETS.

Smoking Rates and Trends During Pregnancy

Each year for the last several years, over 4.1 million women gave birth in the United States (Martin et al., 2006). Table 11.1 presents rates of smoking by race/ethnicity among pregnant women in the U.S from 1990 to 2004, based on self-reported information derived from birth certificate data from 46 states and Washington, DC. As indicated, American Indians and non-Hispanic whites have the highest rates of smoking during pregnancy. African-Americans have rates that are lower than these groups but higher than Hispanics and Asian and Pacific Islanders. As also noted in Table 11.1, prevalence rates during pregnancy decreased substantially from 18.4% (approximately 740,000 smokers) to 10.2% (approximately 420,000 smokers) for all races and ethnic groups between 1990 and 2004.

Table 11.1 Percent smokers during pregnancy by race/ethnicity: 1990–2004[a, b]

Year	Hispanic (H)	Non-H White	Non-H Black	American Indian	Asian/Pacific Islander	Total
1990	6.7	21.0	15.9	22.4	5.5	18.4
1992	5.8	19.7	13.8	22.6	5.6	16.9
1994	4.6	17.7	11.5	21.0	3.8	14.6
1996	4.3	16.9	10.3	21.3	3.3	13.6
1998	4.0	16.2	9.6	20.2	3.1	12.9
2000	3.5	15.6	9.2	20.0	2.8	12.2
2002	3.0	15.0	8.8	19.7	2.5	11.4
2004	2.6	13.8	8.4	18.2	2.2	10.2

[a] Data from "Table 1. Percent of women who smoked during pregnancy by race/ethnicity and age of mother and % change between 1990 and 1999," in Mathews (2001)

[b] Data from "Table 12. Smoking during pregnancy by race, Hispanic origin, age, and education of mother by selected years, 1989–2004," in National Center for Health Statistics (2006)

During this time, the rate of decrease and the actual prevalence varied across racial and ethnic groups. In 1990, the range was from 5.5% for Asian and Pacific Islanders to 22.4% for American Indians. In 2004, the range was 2.2% for Asian and Pacific Islanders to 18.2% for American Indians. Between 1990 and 2004 the decrease in smoking prevalence during pregnancy for non-Hispanic black women was 7.5% (from 15.9 to 8.4), 7.2% for non-Hispanic white women (from 21 to 13.8), and 4.1% for Hispanic women (from 6.7 to 2.6). While cessation clearly decreases the risk of adverse perinatal outcomes for an individual woman, the decrease in smoking prevalence for successive cohorts of pregnant black and white women did not translate into any change in the ratio of adverse outcomes, particularly low birth weight, between the groups.

In contrast to the very large decreases in prevalence rates reported by the NCHS-CDC, the annual national household survey data collected by the Substance Abuse and Mental Health Services Administration (SAMHSA, 1990–2004) documented a prevalence rate of smoking during pregnancy of 20.0% on average in 1990–1992 (800,000 smokers) and of 17.3% on average for 2002–2004 (720,000 smokers). The SAMHSA data in Table 11.2 were derived from face-to-face interviews conducted each month with representative samples of residents, males and females, non-pregnant and pregnant (>1,500/year) respondents. The National Epidemiologic Survey (NES) of Alcohol and Related Conditions, an independent, face-to-face survey of a representative sample of 43,093 adults, including 1,517 pregnant women, was conducted in 2001–2002 (Goodwin, Keyes, & Simuro, 2007). The NES survey reported a smoking prevalence rate during pregnancy of 21.7%, corroborating the SAMSHA rates in Table 11.2.

Since data from the SAMSHA and NES interviews were based on personal interviews of representative samples of respondents, they are more likely to be accurate estimates of the self-reported national smoking rate during pregnancy (Aquilino, 1994). The combined non-biochemically confirmed self-reported average SAMHSA and NES survey prevalence rate estimate of 19.5% provides an estimate of about 800,000 women in the U.S. who smoked during pregnancy in 2003/2004. If a biochemical test had been performed to confirm self-reports in the most recent SAMSHA surveys, assuming a conservative 5% non-disclosure rate, the validated prevalence rate among pregnant women in the U.S., was probably ≥20% in 2007/2008.

Smoking during pregnancy is also a world-wide problem, especially in developed countries. Global and country-level prevalence rate estimates are incomplete, and no national self-reported rates have been corroborated by biochemical means. It has been estimated, based on self-reports, that globally, approximately 12% of females (>12 million) smoke during pregnancy (Windsor, 2001). Smoking prevalence rates of women of childbearing age have typically been stable or not significantly decreased in developed countries in the last 5 years. In addition, because adult male smoking rates are consistently much higher than female rates in both developed and developing countries, the number of fetuses in utero and children up to 5 years of age annually exposed to ETS (one smoker in the home) is at least ten million.

Table 11.2 Average smoking rates of women ages 15–44: 1990–2004[a]

Average years	Smoking rates	
Survey	Pregnant (%)	Non-pregnant (%)
1990–91–92 average	20.0	30.0
1994–95–96 average	20.6	31.8
1999–00–01 average	19.4	30.2
2002–03–04 average	17.3	30.0

[a] Substance Abuse and Mental Health Services Administration (1990–2004)

Evidence-Based Assessment of Smoking Status During Pregnancy

Self-reports of non-smoking status during pregnancy that are not confirmed by a valid biochemical test (e.g., saliva, urinary analysis, or exhaled CO) produce inaccurate prevalence rates. None of the U.S. (or international) estimates of prevalence during pregnancy reflect independent confirmation by a biochemical test of patient non-disclosure of their smoking status. As noted in the studies discussed below, the more socially desirable report of being a non-smoker is often made, especially at the first prenatal visit, or later during care by a patient who smokes.

The primary group of pregnant women who provide inaccurate self-reports of smoking status are women who smoke, but state at the onset of care that they have quit since becoming pregnant. The CDC reported a urinary cotinine-confirmed smoking status non-disclosure rate of 28% among 6,800 pregnant Medicaid patients in Colorado, Maryland, and Missouri at the onset of care (Kendrick et al., 1995). Likewise, in a representative sample of 814 pregnant Medicaid patients in Alabama recruited at their first visit over a 24 month period (Boyd, Windsor, Perkins, & Lowe, 1998), a saliva cotinine test confirmed a false negative rate of 26.2%. Among a representative cohort of new pregnant Medicaid patients (N = 431) from eight randomly selected counties in Alabama, a 24% non-disclosure rate was confirmed by saliva cotinine tests (Windsor, Woodby, et al., 2000), and a comparable non-disclosure rate of 25% was confirmed among new pregnant Medicaid patients in Philadelphia (Webb, Boyd, Messina, & Windsor, 2003). Additionally, a tobacco use measurement study of 3,150 Medicaid and non-Medicaid patients using serum cotinine tests at the onset of care confirmed a 12% false negative rate (Kharrazi et al., 2004).

The studies noted above demonstrate consistently higher non-disclosure rates of smoking among pregnant women on Medicaid. These data are noteworthy given that each year approximately 41% (1.6 million) of women giving birth in the United States receive Medicaid-funded delivery services (National Governors Association, 2007). In addition, these studies and other reviews examining the relationship between the number of cigarettes per day and a cotinine value have confirmed the substantial inaccuracy of reported CPD among pregnant smokers. The correlation, typically r = 0.20 –0.50 between CPD reports and saliva cotinine tests, varies by smoking typology, namely women's smoking patterns. CPD self-reports, however, can be grouped for practical purposes into crude ordinal categories of daily exposure: 1–9 CPD (low), 10–19 CPD (moderate), ≥20 CPD (high), and ≥30 CPD (very high).

Systematic reviews of the literature by Lumley, Oliver, Chamberlain, and Oakley (2004), Russell, Crawford, and Woodby (2004), Windsor, Boyd, and Orleans (1998), and Windsor and Orleans (1986) on the accuracy of self-reports of smoking status have also consistently confirmed high rates of non-disclosure in evaluation studies: frequently ≥20% for the Experimental (E) Group and ≥10% for the Control (C) Group. A CDC intervention evaluation study (Kendrick et al., 1995) confirmed, based on urine cotinine tests, non-disclosure rates of 49% among E group patients and 32% among C group patients. These rates, derived from data from a sample of 6,800 pregnant women, are the highest non-disclosure rates in the scientific literature. In considering this evidence, it is important to keep in mind that unless researchers consistently use accurate cut-off points for active smoking, a very small percent of 'non-disclosure' status may represent environmental tobacco smoke [(ETS); Arheart et al., 2008].

Because of the lack of biochemical confirmation of smoking self-reports, the magnitude of invalid self-reported rates, and the inaccuracy of CPD, two conclusions can be drawn from this body of scientific evidence. Reports of smoking rates during pregnancy significantly underestimate the actual rates, and estimates of progress in achieving national, state, and local program tobacco use objectives, not confirmed by a biochemical test, overestimate progress. A second equally salient conclusion can be drawn from this body of evidence. All epidemiologic studies that have not used biochemical documentation of self-reports for smoking status to measure and evaluate the

association between active and passive exposure and adverse maternal, fetal, and infant effects have underestimated the magnitude, incidence, and prevalence of tobacco attributable perinatal risks.

If, between 1990 and 2004, as indicated by the CDC, there was a 44.6% reduction in smoking prevalence among pregnant women, smoking attributable perinatal morbidity incidence rates and trends, particularly LBW rates, should have been significantly lower during this period. Although research has demonstrated a number of significant determinants of low birth weight in addition to smoking status, including an increase in multiple births, it is notable that LBW rates in the U.S. have increased about 20% in the last 20 years during a time that the CDC has consistently reported very large decreases in the smoking prevalence rates during pregnancy (Martin et al., 2006).

Comprehensive Evaluation of Interventions for Pregnant Smokers

As noted in the "Maternal, Fetal, and Infant Risk" section of this chapter, the evidence confirming active and passive tobacco exposure during pregnancy as the causes of adverse effects is unequivocal. While methods to rigorously evaluate all aspects of an evidence-based tobacco treatment program have been described in the literature for over 20 years, almost two-thirds of the evaluation studies (53/74) initially identified for review and inclusion in this chapter did not meet standard methodological criteria employed for this and other systematic reviews. The majority of studies (a) had small, inadequate sample sizes, (b) lacked sufficient statistical power (Fleiss, Levin, & Paik, 2004), (c) did not biochemically confirm smoking status, and/or (d) failed to confirm the delivery of the intervention. Thus, while the evidence confirming adverse effects is strong, the effectiveness of tobacco treatment methods to both significantly increase cessation and reduce perinatal and reproductive risks among all women and specific racial or ethnic groups is incomplete. One of the early and most methodologically sound evaluations examining the benefits of smoking cessation and its effects on infant birth weight is the study by Sexton and Hebel conducted in the early 1980s and published in 1984. It was a prospective randomized clinical trial that documented the "efficacy" of intensive, regular, and multiple face-to-face and telephone counseling sessions by specialized nursing staff for heavy smokers (≥15 CPD) to improve infant health (Sexton & Hebel, 1984)[1]. Their study confirmed an experimental group quit rate of 27% and a control group quit rate of 3%. They also reported an adjusted, significantly higher mean infant birth weight of 100 g in their experimental group (n = 463) compared to the control group (n = 472).

Review of the Evidence for Smoking Cessation Interventions During Pregnancy

The 1979 Report of the Surgeon General (Hasselmeyer et al., 1979) asked, "How can efforts to actively discourage smoking during pregnancy be made more effective?" Thirty years later, considerable empirical insight with excellent internal validity and good external validity is available to answer this question. Windsor and Orleans (1986) and Windsor, Boyd, et al. (1998) conducted a systematic, comprehensive review of the methodological quality of evaluations of interventions for pregnant smokers from 1972 to 1997. Additional reviews have been performed by Ershoff et al. (1999), Fiore et al. (1996, 2000, 2008), Lumley et al. (2004), Mullen, Ramirez, and Groff (1994),

[1] The study by Sexton and Hebel (1984) is the most rigorous epidemiological study related to smoking and perinatal outcomes to date. While it is presented and discussed in the text of the chapter, it is not included in the tables of evidence as its publication date lies outside of the time parameters established for inclusion in this book.

Orleans, Melvin, Marx, Maibach, and Vose (2004), Walsh and Redman (1993), Windsor (2003), Windsor, Li, et al. (1993), and Windsor, Whiteside, et al. (2000). These systematic reviews thoroughly describe the U.S. and global evidence base in this specialized area of tobacco treatment.

The remainder of this chapter provides an up-to-date integrated review and critique of the evidence with respect to intervention effectiveness of smoking cessation for pregnant women. Only published evaluations 1985 through June 2007 that met the following methodological criteria were included: (a) applied a randomized-experimental design (RCT) to assign patients or sites to an experimental or control group; (b) used a biochemical test to document the validity of patient self-reports of smoking status at base line and end of pregnancy; (c) randomized and followed-up sample sizes of ≥100 patients each in the experimental and control groups to attempt to control for selection biases and the issue of statistical power; and (d) defined their intervention methods and presented process evaluation data documenting experimental and control group patients' exposure.

Table 11.3 Systematic and comprehensive evaluation of interventions for pregnant smokers, U.S. population: 1985–2006

Author (year)	Population	Interventions	Sample size	Quit rates	95% CI
Windsor et al. (1985)	Efficacy	CBT	C = 104	1.9%	1.7–30.6
	Medicaid	Self-help guide	E_1 = 103	5.8%	
	55% Black	10 min counsel	E_2 = 102	13.8%	
Ershoff, Mullen, and Quinn (1989)	Effectiveness	CBT	C = 116	17.2%	0.9–2.4
	HMO	8 Manuals	E = 126	26.2%	
	25% Black	Brief counsel			
Windsor, Lowe, et al. (1993)	Efficacy	CBT	C = 414	8.5%	1.1–2.5
	Medicaid	Self-help guide	(C) = 100[a]	3.0%	
	52% Black	15 min counsel	E = 400	14.2%	
Secker-Walker et al. (1994)	Efficacy	CBT	C = 226	11.3%	0.5–1.7
	Mixed-Medicaid-47% + 0% Black	3 Sessions Booklets	E = 188	12.9%	
Hartmann et al. (1996)	Effectiveness	CBT	C = 100	10.0%	1.0–4.9
	OB residents	Self-help guide	E = 107	20.0%	
	24% Black	Brief counsel			
Gielen et al. (1997)	Efficacy	CBT	C = 198	5.6%	0.5–2.5
	Medicaid	Self-help guide	E = 193	6.2%	
	85% Black	15 min Peer X			
Secker-Walker, Solomon, Flynn, Skelly, and Mead (1998)	Efficacy	Repeated MD	C = 141	11%	0.9–4.0
	Mixed-Medicaid	Referral	E = 135	18%	
Ershoff et al. (1999)	Efficacy	CBT	E_1 = 111	22.5%	0.4–1.3
	HMO	Manuals	E_2 = 120	16.7%	
	15% Black	IT/MI telephone	E_3 = 101	20.8%	
Windsor, Woodby, et al. (2000)	Effectiveness	CBT	C = 126	8.8%	2.2–4.1
	Medicaid	Self-help guide	E = 139	17.3%	
	16% Black	10 min counsel	(C) = 96[a]	6.6%	
Donatelle, Prows, Champeau, and Hudson (2000)	Efficacy	Self-help guide	C = 108	9.0%	2.2–10.3
	Medicaid	Brief counsel	E = 112	32.0%	
	11% Black	$$$/Incentives	(C) = 100[a]	3.0%	
Rigotti et al. (2006)	Efficacy	CBT	C = 212	7.5%	0.7–2.7
	HMO	Tele-counseling	E = 209	10.0%	
	12% Black	5 Calls/68 min			

[a](C) = Pre-Trial Historical Comparison Group

Table 11.4 Systematic and comprehensive evaluation of interventions for pregnant smokers, non-U.S. population: 1985–2006

Author (year)	Population	Intervention	Sample size	Quit rates	95% CI
Hjalmarson et al. (1991)	Effectiveness	CBT	C = 209	8.6%	0.9–2.4
	Sweden	Manual	E = 444	12.6%	
	Maternity clinic	Brief counsel			
Rush, Orme, King, Eiser, and Butler (1992)	Effectiveness	Very brief	C = 144	10.4%	0.7–2.8
	Hospital	Counseling	E = 175	14.1%	
	OB clinics	Home visit			
O'Connor et al. (1992)	Efficacy	CBT + guide	C = 115	6.1%	0.9–6.2
	P.H. nurse	20 min counsel + T	E = 109	13.3%	
Walsh et al. (1997)	Efficacy	MD advice/RN	C = 125	0%	1.0–2.6
	Australia	14 min video	E = 127	9%	
	Midwives	Self-help manual			
Lowe et al. (1998)	Effectiveness	Midwife	C = 119	0%	1.0–2.6
	Australia	Counseling	E = 125	9%	
	RN-midwives	Booklet			
Panjari et al. (1999)	Effectiveness	CBT	C = 317	9.8%	0.5–2.0
	Australia	1st visit = 25 min	E = 278	11.9%	
	RN-midwives	2–4 = 5–10 min			
Hajek et al. (2001)	Efficacy	CBI + manuals	C = 440	7.0%	0.7–1.9
	England	CO-feedback	E = 431	6.0%	
	RN-midwives	10–15 min counsel			
Moore et al. (2002)	Effectiveness	CBT	C = 803	20.7%	0.7–1.1
	England	5 Booklets	E = 724	18.8%	
	RN-midwives	Mailed			
Lawrence et al. (2003)	Effectiveness	CBT	E_1 = 217	1.4%	
	England	TTM-manual	E_2 = 243	2.6%	0.5–5.3
	RN-midwives	IT-sessions	E_3 = 251	3.1%	0.5–5.6
Tappin et al. (2005)	Efficacy	CBT–MI	C = 411	4.6%	0.5–2.0
	Scotland	2–5 Home visits	E = 351	4.8%	
	RN-midwives	>30 min/per			

MI motivational interviewing

The comprehensive review of the literature conducted for this chapter identified 21 studies – 11 U.S. studies and 10 non-U.S. studies that met the above criteria. Table 11.3 presents a synopsis of the U.S. studies, and Table 11.4 presents a synopsis of the non-U.S. studies. While the meta-analyses conducted by Fiore et al. (2000, 2008) and Lumley et al. (2004) are not included in these tables, consistent with their reviews of U.S. evaluation studies, our assessment of the literature confirmed that treatment methods, if effective, were equally effective for black and white pregnant smokers.

Almost all of the intervention evaluation studies conducted in the last decade (1996–2006) applied variations of the treatment methods that the first decade of valid evaluation studies (1985–1995) had confirmed as "efficacious." If a theoretical framework was cited, it typically was Social Cognitive Theory or Cognitive Behavioral Theory (Bandura, 1986). However, reviews of the validity and efficacy of interventions based on the Stages of Change–Transtheoretical Model (Prochaska & Diclemente, 1983) have reported that interventions tailored to stage did not predict behavior change in pregnancy (Lumley et al., 2004; Riemsma et al., 2003; Solomon et al., 1996; Woodby, Windsor, Snyder, Kohler, & Diclemente, 1999).

The evaluation studies that met criteria for further review and inclusion in this chapter incorporated similar core intervention components: a printed cessation manual with information about risks and benefits, a description of specific cessation methods, and single or multiple face-to-face

counseling session(s). Telephone counseling intervention and reinforcement was provided in a few studies. Although a range of providers of care and counseling methods were reported, a nurse midwife was the most typical interventionist. Level of treatment effectiveness was not, however, associated with a specific type of provider. Only one RCT (Table 11.3; Gielen et al., 1997), a replication of the Windsor et al. (1985) Smoking Cessation and Reduction in Pregnancy Treatment (SCRIPT) Program Trial I, evaluated the efficacy of evidence-based methods delivered by a trained (nonprofessional) community health worker (CHW). Gielen et al. found that the CHW interventionist was not efficacious with Medicaid patients enrolled in a prenatal care program at a university-based hospital in Baltimore, Maryland.

In the 21 studies presented here, tobacco treatment methods, if effective, typically produced a cessation rate in the experimental group 6–7% higher than that of the control group. It is important to note that a partial explanation for the lack of significant difference in effectiveness between the experimental and control groups in several evaluation studies is that control group interventions were substantially more intensive than usual care. Particularly in the second generation of U.S. evaluations (1996–2006), pregnant control group smokers were typically provided more intensive interventions than the corresponding control groups in first generation (1985–1995) U.S. efficacy studies.

Summary of Evaluation Study Quality

A summary of the quality of the studies in Tables 11.3 and 11.4, using the methodological quality checklist described by Downs and Black (1998), is presented in Table 11.5. Because the studies met multiple rigorous criteria for inclusion, they all had at least "Good" methodological quality.

Table 11.5 Evaluation study quality

Author (year)	Reporting	External validity	Internal validity – bias	Internal validity – confounding	Power	Score	Quality
Windsor et al. (1985)	11	4	6	5	1	27	EX
Ershoff et al. (1989)	11	4	7	6	1	29	EX
Windsor, Lowe, et al. (1993)	12	4	6	5	2	29	EX
Secker-Walker et al. (1994)	12	4	6	5	2	29	EX
Hartmann et al. (1996)	12	4	6	5	1	28	EX
Gielen et al. (1997)	11	2	5	5	1	24	VG
Secker-Walker et al. (1998)	12	4	6	6	2	30	EX
Ershoff et al. (1999)	11	3	6	5	1	26	EC
Windsor, Woodby, et al. (2000)	11	3	6	4	1	25	VG
Donatelle et al. (2000)	10	3	5	5	0	23	VG
Rigotti et al. (2006)	11	2	6	5	1	25	VG
Hjalmarson et al. (1991)	12	4	6	5	1	28	EX
Rush et al. (1992)	11	2	6	5	0	24	VG
O'Connor et al. (1992)	11	3	6	5	0	25	EX
Walsh et al. (1997)	11	2	6	5	0	24	VG
Lowe et al. (1998)	11	4	6	5	0	26	EX
Panjari et al. (1999)	10	1	5	4	2	22	G
Hajek et al. (2001)	10	1	5	4	1	21	G
Moore et al. (2002)	11	4	6	5	1	27	EX
Lawrence et al. (2003)	10	2	5	4	1	22	G
Tappin et al. (2005)	11	3	6	5	2	27	EX
Range	0–12	0–4	0–7	0–6	0–2	0–31	

Ratings: good (G) = 20–22, very good (VG) = 23–25, excellent (EX) = ≥26

Nevertheless, problems in program implementation and low levels of provider participation have been documented as major process evaluation issues. Hajek et al. (2001), for example, reported that only 61% of midwives recruited patients. Another salient issue confirmed in this review and noted by other reviews was the number of evaluations that reported a high percentage (≥20%) of eligible experimental group patients who did not receive the intervention or received only parts of it. Lawrence, Aveyard, Evans, and Cheng (2003), for example, reported that an average of 50% of patients refused to participate in the evaluation, and 7–23% refused to accept different types of interventions.

Without the participation of >80% of eligible patients, including 100% completion of baseline and 90% completion of follow-up assessments, and the delivery of all intervention procedures to 100% of smokers, the probability of behavioral impact or improved maternal, fetal, and infant health is significantly reduced or eliminated. The internal and external validity of the results are also seriously threatened, and results will be partially or totally compromised by these sources of bias (Windsor, Clark, Boyd, & Goodman, 2004).

Because of the salience of the failure-to-treat issue for evaluations of current and future prenatal smoking cessation programs (Newell, 1992), an evidence-based "Process Evaluation Model" is presented in the future challenges section at the end of this chapter. As the cost of a new intervention is also one criterion in the decision to adopt a new tobacco treatment process, and given that few evaluations have presented cost-effectiveness analyses of tobacco treatment methods for pregnant smokers, cost evaluation is also briefly discussed in the future challenges section of this chapter.

The comprehensive and critical literature review in this chapter confirms that rigorous pilot studies, including qualitative inquiry and input from staff and patients along with insights from large system-wide trials directed by experienced providers and senior investigators, are needed prior to intervention implementation. The lack of pilot tests contributed, at least in part, to the inadequate delivery of smoking cessation interventions, thus resulting in low levels of client smoking cessation. These deficiencies must be addressed in future evaluations. Despite these difficulties, over 20 years of evaluations of smoking cessation programs during pregnancy have provided sufficient information to develop evidence-based smoking during pregnancy cessation programs. These approaches are described below.

Evidence-Based Smoking Cessation Methods for Professional Practice

The Smoking Cessation and Reduction in Pregnancy Treatment (SCRIPT) Program is one of the most efficacious, nationally recognized evidence-based programs for pregnant smokers (Fiore et al., 2000, 2008). The SCRIPT is a synthesis of methods derived from a comprehensive review of the relevant world literature (Lumley et al., 2004; Windsor, Boyd, et al., 1998; Windsor & Orleans, 1986). The objectives of SCRIPT are to increase cessation rates and to reduce the prevalence of smoking during pregnancy. The acceptability of SCRIPT to clients, the feasibility of routine delivery by regular prenatal care staff, and the effectiveness, and cost-effectiveness of SCRIPT methods for Medicaid and non-Medicaid supported clients have been demonstrated by independent evaluations in Alabama, North Carolina, Ohio, Canada, Norway, and Sweden (Gebauer, Kwo, Haynes, & Wewers, 1998; Hartmann, Thorp, Pahel-Short, & Koch, 1996; Hjalmarson, Hahn, & Svanberg, 1991; O'Connor et al., 1992). Dissemination of SCRIPT has been promoted by the Agency for Health Care Research and Quality's Clinical Practice Guidelines (Fiore et al.) and the American College of Obstetricians and Gynecologists (2005).

SCRIPT'S 5 A's Model – Ask, Assess, Advise, Assist, Arrange – is used to organize and structure interventions (ACOG, 2005; Fiore et al., 2000, 2008; Lumley et al., 2004; Windsor, 2003;

Table 11.6 SCRIPT evaluation results: estimated population impact of evidence-based smoking cessation methods

Author (year)	Measurement	E Group	C Group	E – C
Windsor, Woodby, et al. (2000)	S-Cotinine	139 (17.3%)	126 (8.8%)	+8.5%
Hartmann et al. (1996)	CO	107 (20.0%)	100 (10.0%)	+10.0%
Windsor, Lowe, et al. (1993)	S-Cotinine	400 (14.3%)	414 (8.5%)	+5.8%
O'Connor et al. (1992)	U-Cotinine	115 (13.3%)	109 (6.0%)	+4.7%
Hjalmarson et al. (1991)	S-Thiocyanate	444 (12.6%)	209 (8.6%)	+4.0%
Windsor et al. (1985)	S-Thiocyanate	102 (13.7%)	104 (1.9%)	+11.8%
(N = 1,492) U.S. studies		Total = 15.5%	–7.8%	+7.7%
(N = 1,373) Non-U.S. studies		Total = 11.5%	–4.3%	+7.2%
(N = 2,865) All studies		Total = 13.4%	–6.3%	+7.1%

Windsor, Woodby, et al., 2000). There are ten SCRIPT procedures which specify the actions that need to be taken within each step of the 5 A's model. Each step is to be delivered by a trained interventionist within the prenatal care setting (e.g., nurse, social worker, or WIC nutritionist) and then reinforced by the medical care provider (physician or nurse midwife).

Table 11.6 presents a summary of the six evaluation studies which tested the SCRIPT components and procedures presented in Fig. 11.1. These six studies include the three RCTs also shown in Table 11.3 (Windsor et al., 1985; Windsor, Lowe, et al., 1993; Windsor, Woodby, et al., 2000) and three independent evaluation studies of SCRIPT methods by Hjalmarson et al. (1991) in Sweden, O'Connor et al. (1992) in Canada, and Hartmann et al. (1996) in North Carolina (Tables 11.3 and 11.4). As indicated in Tables 11.3 and 11.4, biochemically confirmed excess cessation rates attributable to SCRIPT from the six RCTs was on average 7.7% in the U.S. studies and 5.1% in the non-U.S. studies. If the annual estimated U.S. pregnant smoker cohort (>17% and >700,000) were provided an evidence-based tobacco treatment intervention such as SCRIPT, an estimate of the annual number of additional quitters might be 40,000 or more pregnant women (6%).

Although the impact of SCRIPT and other counseling methods has been consistently small, it is important to emphasize that smoking is a very addictive behavior. While pregnancy is one of the most influential conditions and periods of time in a woman's life, with the potential to lead to significant health related behavioral changes including smoking cessation, from an addiction and psychosocial perspective, pregnant smokers present a behavioral challenge. This review confirmed that each year a very large percent (80–90%) of pregnant smokers are not able to quit after they begin prenatal care, even if they receive recommended treatment methods such as the SCRIPT program (Fiore et al., 2000, 2008). As such, among those women who are moderate to heavy smokers (≥10 CPD, CO ≥8 ppm, or urine Cotinine >100 ng/mL, who have failed a behavioral treatment), there may be ≥300,000 pregnant women smokers per year in the U.S. eligible for Nicotine Replacement Therapy (NRT). The following section provides a synthesis of the evidence about NRT use during pregnancy, if a patient has failed behavioral treatment, but wants to try to quit again.

NRT Risks During Pregnancy

The use of NRT is safe and effective for male and female smokers and it is substantially safer to use than daily smoking. NRT in combination with behavioral counseling methods used over an 8 week period typically doubles smoking cessation rates in females (and males) in the general adult population (Fiore et al, 2000, 2008; Fiore, Smith, Jorenby, & Baker, 1994; Henningfield, 1995a, 1995b; Hughes, Goldstein, Hurt, & Shiffman, 1999; Silagy, Lancaster, Stead, Mant, & Fowler, 2004).

The American College of Obstetricians and Gynecologists (2005) has identified the need for physicians to consider NRT's potential for assisting pregnant smokers who want to quit, even after they have failed to quit with behavioral methods, assuming there are no obstetrical contraindications. In 2002, ACOG recommended monitoring cotinine levels if NRT is used by a patient (Chapin and Root, 2004). In *Treating Tobacco Use and Dependence: A Clinical Practice Guideline*, the U.S. Agency for Healthcare Research and Quality (Fiore et al.) recommended that NRT should be considered when, with the use of behavioral methods, a pregnant woman is unable to quit, and when the benefits and likelihood of quitting outweigh the risks of pharmacotherapy and potential continued smoking.

Windsor, Oncken, Henningfield, Hartmann, and Edwards (2000) have reviewed the risks and benefits of behavioral and pharmacological treatment methods for pregnant women. They reported that the appropriate use of NRT during pregnancy, assuming a patient does not smoke while using NRT, would significantly reduce maternal and fetal cotinine levels. NRT use eliminates daily fetal exposure to CO and hundreds of other chemicals and carcinogens that are the primary causes of adverse perinatal events. Because of the very large, reduced nicotine exposure period, 8 weeks of NRT vs. 39 weeks of nicotine from smoking, and the elimination of CO and the other 3,000 toxic constituents of smoke, other reports have indicated that when behavioral methods have failed for heavy smokers (≥10 CPD) who are pregnant, the benefits of NRT would outweigh the risks (Benowitz, 1991, 1998; Benowitz & Dempsey, 2004; Greenland, Satterfield & Lanes, 1998).

Although sample sizes have been small, multiple studies have reported that the daily dose of nicotine from NRT would be less harmful for pregnant smokers and would be lower than the daily dose of nicotine from heavy daily smoking (≥10 CPD; Schroeder et al., 2002). When compared to daily smoking by women, NRT significantly lowers (typically by 50%) maternal and fetal peak and average nicotine/cotinine exposure levels. In a NRT study of pregnant women, Oncken et al. (1997) confirmed that plasma cotinine levels were significantly lower when nicotine gum is used than when smoking ≥10 CPD. They also demonstrated that nicotine concentrations and maternal and fetal hemodynamic effects were lower in pregnant NRT users than in pregnant smokers. Wright et al. (1997) found that the concentrations of salivary cotinine levels were consistently lower in pregnant women who used transdermal nicotine patches than in smoking non-pregnant adults. The highest dose trans-dermal nicotine patch typically delivers nicotine levels approximately 50% of the nicotine levels received by women who smoke 21–30 cigarettes per day (Gupta, Hwang, Causey, Rolf, & Gorsline, 1995; Tonneson et al., 1988).

This synthesis of the NRT literature suggests that there is an alternative for pregnant women who smoke ≥10 CPD and who have a failed behavioral treatment. If women agree not to smoke, the appropriate use for 8 weeks during the second trimester of a slow absorption NRT patch, while awake, should substantially reduce the typical daily cotinine exposure levels by as much as 50%. NRT use over an 8 week treatment period also eliminates maternal and fetal exposure to CO and the thousands of other risk-producing toxins and carcinogens in tobacco smoke.

Evaluations of NRT During Pregnancy

Three evaluations of the efficacy of NRT with pregnant smokers have been published. Wisborg, Henriksen, Jespersen, and Secher (2000) conducted a study in Denmark with 250 patients who smoked ≥10 cigarettes in total after their first trimester and who either received a behavioral program only, or a behavioral program plus nicotine patches. The behavioral treatment program provided to all patients was intensive. Over 2 h of contact time for all patients was provided, which included multiple one-on-one sessions by a trained midwife and MD-OB reinforcement.

The cessation rate of 21% for women in the intervention group (provided the counseling program plus nicotine patches) was not significantly greater than the rate of 19% among women in the control group (provided the cessation counseling intervention plus a placebo patch). The Wisborg et al. study documented a very large participant non-adherence and process evaluation problem in the nicotine patch component of the study. Only 15% of the intervention group participants used their nicotine patches and only 10% of the control group patients used their placebo patches.

Pollak et al. (2007) conducted an evaluation in North Carolina with 181 patients who smoked ≥5 CPD and were between 13 and 25 weeks pregnant. Approximately 70% of the eligible smokers agreed to participate in the study. Two study groups, a cognitive behavioral treatment (CBT) control group and a CBT + NRT (choice) experimental group were established by random assignment using a 1:2 CBT vs. CBT + NRT ratio. All patients received six one-on-one counseling sessions. Five were face-to-face and one was a telephone session. Patients were assessed using saliva cotinine and CO tests at 7 weeks post-randomization, at 38 weeks gestation, and at 3 months post-partum. The cotinine validated cessation rates were significantly higher for the experimental group at 38 weeks compared to the control group (18% vs. 7%; $p = 0.04$), but not at 3 months post-partum (20% vs. 14%; $p = 0.55$). A Data and Safety Monitoring Board (DSMB), which had set a pre-trial policy for serious adverse events (SAE) in patient outcomes, suspended enrollment at the mid-point of recruitment. The DSMB stated, however, that it did not believe that the SAEs were related to patient NRT use.

Oncken et al. (2008) conducted a double-blind, placebo-controlled trial to evaluate the safety and efficacy of 2 mg nicotine gum with 194 patients who smoked at least one cigarette per day. Approximately 78% of the eligible patients at three study hospitals in Connecticut agreed to participate and the majority (54%) were Hispanic. Two groups were established by random assignment: a Nicotine Gum Group (N = 100) and a Placebo Gum Group (N = 94). All participants received two 35 min counseling sessions previously shown to be effective with this population (Dornelas et al., 2006). No significant differences were found in the biochemically confirmed cessation rates between the Gum (13%) and Placebo (9.6%) groups after 6 weeks of treatment. After reviewing the data for 147 patients, a Data and Safety Monitoring Board (DSMB) recommended that enrollment stop, as the documented effect size and target sample (N = 268) did not have sufficient statistical power to detect a significant difference in treatment efficacy.

Both the Pollak et al. (2007) and Oncken et al. (2008) evaluations reported differences in infant birth weight and other salient perinatal outcomes, suggesting a positive, potentially attributable effect of patch and gum use respectively. Although both evaluations are among the most methodologically rigorous in this body of literature, the small sample sizes and small number of patients who adhered to the NRT treatment regimen confirm a need for future evaluations with much larger samples and detailed process evaluation to document the rates of NRT adherence per week and clinical impact. NRT Trials are also needed to determine the use of different doses with different baseline levels of addiction.

Two evaluations of the efficacy of behavioral treatment plus NRT for pregnant smokers are in progress at the time of the writing of this chapter. The El-Mohandes and Windsor SCRIPT + NRT (Patch) Trial at the George Washington University Medical Center, funded by NICHD-NIH, is designed to evaluate the acceptability and efficacy of nicotine patches among African-American pregnant smokers in Washington, DC (2004–2009). The Coleman et al. SCRIPT + NRT (Patch) Trial at the University of Nottingham School of Medicine, funded by the Medical Research Council of the U.K. (2005–2010), is designed to evaluate the acceptability and efficacy of nicotine patch use among pregnant smokers enrolled in the health services in the Nottingham Region of England (Coleman et al., 2004). The Fiore et al. (2000, 2008) reports and the above body of evidence indicate that more data are needed to assess patient acceptability, safety, and efficacy of NRT use during pregnancy.

Synthesis of Evidence: Implementation of SCRIPT for a System of Care

Over the last decade, a large number of reports have identified a major need to close the gap between population health research and professional practice with respect to smoking cessation interventions. The need for dissemination, diffusion, and adoption of "Best Practice" tobacco treatment programs by systems and clinical practices has received overwhelming support in the literature (Bero et al., 1998; Blum & Solberg, 1996; Curry, Orleans, Keller, & Fiore, 2006; Ershoff, 2004; Glasgow, Bull, Gillette, Klesges, & Dzewaltowski, 2002; Glasgow, Lichtenstein, & Marcus, 2003; Lowe, Balanda, & Clare, 1998; Lumley et al., 2004; McAfee, 2007; Orleans, Barker, Kaufman, & Marx, 2000; Orleans et al., 2004; Rogers, 2003; Windsor, Dalmat, Orleans, & Gritz, 1990; Windsor, Li, et al., 1993) as well as from national organizations such as the Centers for Disease Control and Prevention, March of Dimes Birth Defects Foundation, and the American Lung Association. The following section summarizes the related literature and answers the question: What comprehensive evidence-based tobacco use intervention can be recommended to policy, program, and practice leadership for system-wide dissemination, adoption, and evaluation?

Table 11.7 presents the eight essential, structural components of such a comprehensive program as reported in Fiore et al. (2008). Table 11.7 also shows the evidence for each element for smokers (pregnant and non-pregnant) based on Fiore et al. (2000, 2008), Lumley et al. (2004), and the comprehensive review presented in this chapter. The term SCRIPT PLUS (+) denotes a system-wide

Table 11.7 Estimated behavioral impact of treatment structural variables[a]

Treatment structure variable (studies)	Odds ratio	Est. quit rate	SCRIPT
#1. Levels of intensity (43 studies)			
a. No contact	1.0	10.9%	
b. Low intensity (3–10 min)	1.6	16.0%	S
c. High intensity (>10 min)	2.3	22.1%	S+
#2. Contact time (35 studies)			
a. No minutes	1.0	11.0%	
b. 4–30 min	1.9	18.8%	S
c. 31–90 min	3.0	26.5%	S+
#3. One-on-one sessions (45 studies)			
a. 0–1 Session	1.0	12.4%	S
b. 2–3 Sessions	1.4	16.3%	S+
#4. Number of clinicians (37 studies)			
a. No clinician	1.0	10.8%	
b. 1 Clinician	1.8	18.3%	S
c. 2 or more clinicians	2.5	23.0%	S+
#5. Format type (58 studies)			
a. None	1.0	10.8%	
b. Self help	1.2	12.3%	S
c. Telephone counseling	1.2	13.1%	S+
d. Individual	1.7	16.8%	S/S+
#6. Number of formats (54 studies)			
a. None	1.0	10.8%	
b. 1 Format	1.5	15.1%	
c. 2 Formats	1.9	14.4%	S
d. 3 or 4 Formats	2.5	18.5%	S+
#7. Intra-treatment supp. (50 studies)	1.3	23.2%	S+
#8. Extra-treatment supp. (19 studies)	1.5	16.2%	S+

[a] AHRQ meta-analysis, 2000
S = SCRIPT; S+ = SCRIPT PLUS

treatment program with all components. Each successfully implemented SCRIPT+ component should enhance the effectiveness of all other components.

A SCRIPT PLUS intervention should include: (a) an attractive, user-friendly, pregnancy specific self-help cessation manual at the fifth to sixth grade reading level such as Windsor's (1985/2005) "A Pregnant Woman's Guide to Quit Smoking" (Fiore et al., 2000, 2008; Gebauer et al., 1998; Hartmann et al., 1996; Lumley et al., 2004; O'Connor et al., 1992; Walsh, Redman, Brinsmead, & Fryer, 1997; Windsor et al., 1985; Windsor, Lowe, et al., 1993; Windsor, Woodby, et al., 2000); (b) a tailored video (Parker et al., 2007; Walsh et al.; Windsor et al., 1985; Windsor, Lowe, et al.; Windsor, Woodby, & Crawford, 1998; Windsor, Woodby, et al.); (c) a team of trained providers delivering and reinforcing evidence-based methods during regular prenatal visits (Andrews, Tingen, Waller, & Harper, 2001; Fiore et al.; Lumley et al.; Windsor, Woodby, et al.); (d) use of different types and channels of counseling communication by multiple, regular providers (Andrews et al.; Fiore et al.; Lumley et al.); (e) use of telephone counseling sessions and/or an existing state-wide or national quit line (Borland, Segan, Livingston, & Owen, 2001; Parker et al.; Solomon & Flynn, 2005; Solomon et al., 1996; Zhu et al., 1996, 2002); (f) use of care related home visits to provide cessation methods (Olds et al., 2002); (g) establishment of a non-smoking home environment to eliminate ETS along with home/partner/family/social support to promote and reinforce quit attempts and cessation (McBride et al., 1998; Pollak et al., 2001; Woodby et al., 1999); and, (h) application of organizational policies, staff training, and on-going technical assistance to reinforce routine delivery of effective methods by clinicians participating in a prenatal care system (Fiore et al.; Lumley et al.; Windsor, Woodby, et al.).

As demonstrated by the data in Table 11.7, the estimated base quit rate from little or no treatment (1a + 2a + 3a + 4a + 5a + 6a) is 11.1%. The estimated quit rate from brief informational methods by regular providers (SCRIPT program variables: 1b + 2b + 3a + 4b + 5b + 6b) is 15.5%, and the average estimated quit rate from the optimal application by regular providers of all recommended treatment components included in a SCRIPT PLUS program (1c + 2c + 3b + 4c + 5d + 6d + 7 + 8) may produce an average quit rate of 20.3%. Although these composite rates represent a synthesis of the literature, including Medicaid and non-Medicaid smokers, and do not represent quit rates from a specific evaluation of a SCRIPT PLUS program for pregnant smokers, they do define the combination of treatment components needed to produce optimal effectiveness for a system-wide intervention.

Importantly, all eight intervention treatment components exist in almost all states and prenatal care programs. Most components, however, are typically not provided, not implemented, or are not optimally delivered. There has been no evaluation of a system-wide approach to comprehensive tobacco treatment for pregnant smokers. This deficiency is briefly discussed in the future research section below.

Future Research and Program-Practice Challenges

A number of key issues need to be addressed in order to strengthen the evidence and practice base of smoking cessation interventions for pregnant women. Future systems of care and participating practice sites should: (a) routinely apply valid methods for the assessment and measurement of individual smoking status during program participation; (b) define the structure and process of a treatment program and apply process evaluation methods to document performance levels and implementation costs; (c) identify state and national trends, and determine progress in smoking cessation; and (e) implement and evaluate the impact of a system-wide dissemination of a comprehensive approach such as SCRIPT or SCRIPT PLUS on patient behavior.

Adoption of Best Practice

The following are a set of core questions that need to be discussed and answered by the senior policy and practice leadership of an organization considering the dissemination – adoption of intervention programs such as SCRIPT or SCRIPT PLUS for its care network, practices, and patients. While not all inclusive, they represent the salient areas that should be examined in initial discussions about how to define the problem and how to prepare solutions and dissemination-adoption plans:

1. Is our organization and care network committed to providing "evidence-based" methods to our pregnant patients who smoke?
2. What are our current tobacco treatment policies and procedures?
3. What methods are we currently using to assess and to assist our patients who smoke to quit?
4. Are our assessment and intervention methods evidence-based?
5. Are our staff well trained to deliver them and are we routinely conducting staff performance assessments and a process evaluation of the program?
6. Do we know the current, true smoking prevalence, and non-disclosure rates of our patients on entry and during care?
7. What percent of our pregnant patients are quitting from participation in our tobacco treatment program?
8. What are the costs of current tobacco treatment program for pregnant smokers?
9. What are the smoking attributable risks and excess costs for maternal and infant care of participants in our perinatal programs?
10. What are the next steps we need to take to plan, with program managers, providers, and patients, the process of introducing, routinely applying, and evaluating a SCRIPT program for our system of care?

Evidence-Based Measurement of Smoking Status in Professional Practice and Research

A serious measurement problem has been repeatedly documented in the literature by almost every published smoking assessment study of pregnant women over a 20 year period (Mullen, Carbonari, Tabak, & Glanday, 1991). High non-disclosure rates, multiple categories of reproductive and perinatal risks, as well as increasing costs directly caused by active tobacco exposure are well-documented. Thus, it is unclear why a pregnant woman's blood pressure and clinical tests of urine and blood serum are included in routine standards of care, assessing risks for other salient conditions, while prenatal/OB providers and systems of care continue to rely only on frequently inaccurate self-reported measures of smoking status. "Perinatal programs should apply well-recognized standards and performance measures related to smoking cessation [such as the Health-Plan Employer Data Information Set (HEDIS)] in order to advance public health practice and improve the accuracy of smoking status measurement (National Committee for Quality Assurance, 2007)". Clinical providers, program planners, funders, and policy leaders need to engage and seriously address this deficiency in professional practice.

Assessment of tobacco use by pregnant women should include a urinary cotinine dipstick (Oncken et al., 1998; Parker et al., 2007) as part of routine prenatal care to confirm self-reports of

tobacco use at the onset of care and at least once more during pregnancy to document smoking status. Trained staff with this valid information can more accurately "Assess" individual patient exposure and risks, provide information about the hazards of smoking and benefits of cessation, and if the client is willing, "Assist" them in initiating the cessation process.

Valid measurement methods are also essential for larger clinical programs and providers. At a minimum, biochemical assessments are recommended at the beginning of pregnancy, once during the third trimester, and once (≤3 months) during the post-partum period to accurately evaluate progress in achieving individual patient, practice, and system-wide behavioral and clinical objectives. Biochemically confirmed prevalence rates of tobacco exposure for pregnant women should be required in all future tobacco assessment, intervention, and evaluation reports, as well as peer reviewed publications.

Evidence-Based Population Assessment of Smoking Prevalence and Trends

Self-reports of smoking status during pregnancy are often not accurate, especially for Medicaid participants. There is a need for federal DHHS agencies, like SAMSHA and CDC, independently or in collaboration with any of several National Institutes of Health, to implement an evidence-based measurement study of the smoking status of pregnant women in the United States. Valid smoking status assessments using self-reports, blood levels of CO), and urine or saliva cotinine tests should be performed at the first prenatal care or home visit, and at least once during the pregnancy in the third trimester. A sufficient number of Alaska Native, American Indian, Asian, black, Hispanic, and white patients need to be included in the tobacco use assessment survey to document valid and representative rates during pregnancy for each major racial and ethnic group and for the total U.S. population at risk. Such a study would establish a valid baseline for future surveillance and trend analyses at the beginning of this decade.

Ongoing surveillance should be conducted through national surveys and reported at least every 5 years as part of regular reporting by a lead DHHS Agency, such as SAMSHA and CDC, in collaboration with NICHD, NCI, and/or NIDA, to document "true" national smoking rates, patient non-disclosure rates, and levels of active and passive tobacco exposure during pregnancy. Mid-decade and end of decade surveys using evidence-based measurement methods are essential to assess actual progress in meeting the 2010, 2015, and 2020 smoking and pregnancy Health Promotion-Disease Prevention Objectives for the Nation. It would seem feasible for the Office of Applied Studies, SAMSHA, to implement such a survey by over-sampling pregnant women from all major races and ethnic groups, and by including the direct collection of urine or saliva samples for cotinine analysis in its 2010/2011 household surveys.

Performance Measurement and Process and Cost Evaluations

Another significant problem area, noted in the comprehensive review of this chapter and other reviews during the last 20 years, has been the failure of many evaluations to include detailed descriptions of their intervention methods and to document staff delivery of all core program procedures. Detailed descriptions of specific intervention components are needed, including frequency, duration of time/contact, and the type(s) of intervention procedures delivered. Data documenting group practice levels (staff performance measurement) and overall program implementation should be required in all future reports. This information is essential to plan staff training, prepare process and cost evaluation plans, and for replication by future programs. The description of the core structure, process, and content of SCRIPT presented in Fig. 11.1 is an example of this type of information.

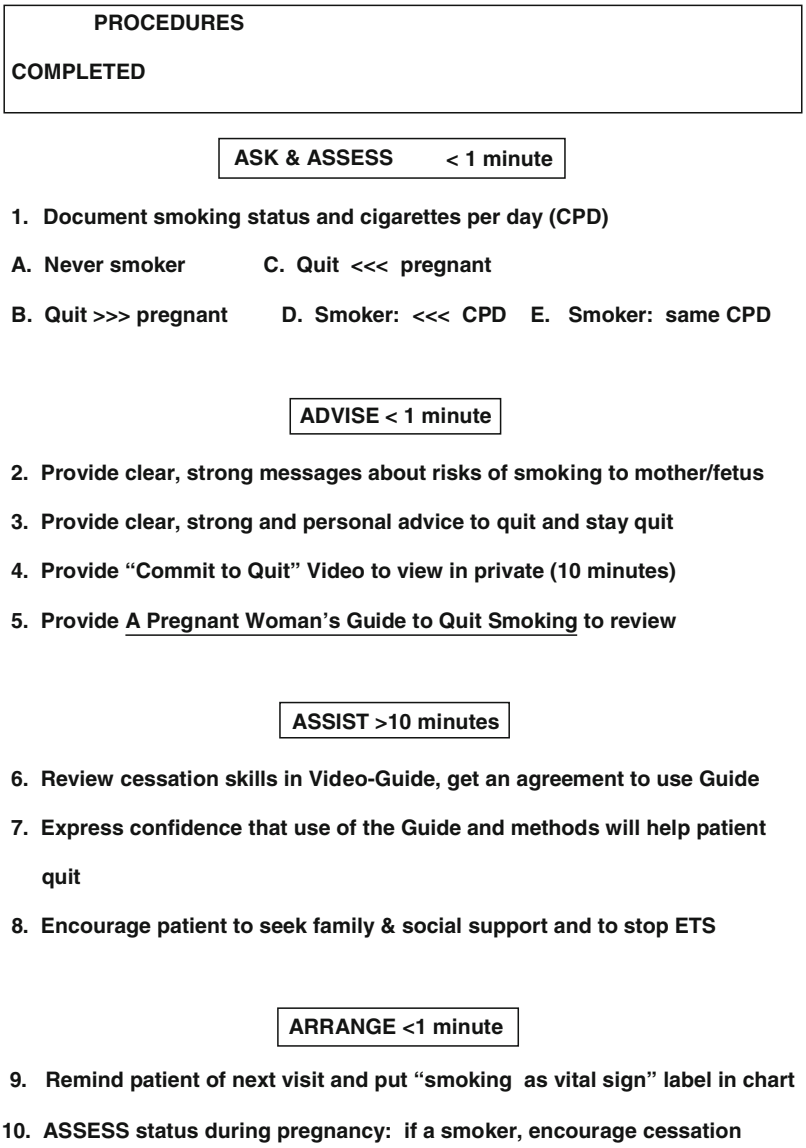


Fig. 11.1 SCRIPT – counseling procedures

A continuous quality improvement model for individual providers, practices, and programs, the SCRIPT Process Evaluation Model (PEM), has been developed for patient counseling programs (Windsor et al., chap. 5, *Process Evaluation*, 2004). A description and applications of the PEM, using data and results from five Robert Wood Johnson Foundation funded intervention programs for pregnant smokers, is presented in *A Process Evaluation Model for Patient Education Programs for Pregnant Smokers* (Windsor, Whiteside, et al., 2000). The SCRIPT-Process Evaluation Model (PEM) requires definitions of the structure, content, methods, and materials of all core patient assessment and counseling procedures such as those defined in Fig. 11.1.

Cost is always an issue when a new treatment program is being introduced into a prenatal care system. A cost effectiveness analysis (CEA) can be performed with valid cessation rates and cost/process

evaluation information. Cost effectiveness and cost-benefit analyses and evaluations of tobacco treatment and SCRIPT methods are described in multiple sources in the literature (Adams, Solanki, & Miller, 1997; Li, Windsor, Lowe, & Goldenberg, 1992; Miller, Villa, Hogue, & Sivapathasundaram, 2001; Parker et al., 2007; Windsor, 2003; Windsor et al., 2004, chap. 7, *Cost Evaluation*; Windsor, Lowe, et al., 1993; Windsor, Warner, Cutter, 1988). It will be important for future programs to routinely incorporate process and cost evaluation methods in their planning and reporting of evaluation results.

Dissemination-Adoption Evaluation and Research

There is consensus that a comprehensive approach to smoking cessation intervention program development and delivery, beyond the individual patient, is needed. When the cumulative body of evidence for pregnant smokers is considered, it is clear that the science, policy and practice base of this specialized area of tobacco cessation treatment has matured to the dissemination and evaluation phase. With established evidence supporting the internal and external validity of tobacco treatment (Fiore et al., 2000, 2008), a trans-NIH announcement of research support was generated: Dissemination and Implementation Research in Health (PAR-02-131). This on-going funding supports rigorous evaluations of programs whose primary aims are to evaluate the institutionalization of evidence-based intervention for disadvantaged populations served by systems of care.

On October 1, 2007, the National Cancer Institute funded an evaluation of such an intervention for pregnant smokers, the West Virginia Right From the Start (RFTS)-SCRIPT Dissemination Program (R. Windsor & J. Clark, Co-Directors, 2007–2011). West Virginia has the highest self-reported state-wide smoking prevalence rate, 45%, for Medicaid patients in the U.S. With a CO confirmed nondisclosure rate >5% among multiple, representative samples of pregnant patients, the true current prevalence rate is approximately 50%. The primary objective of the West Virginia RFTS-SCRIPT Dissemination Program is to evaluate the process and impact of the dissemination and adoption of a comprehensive SCRIPT PLUS Program. SCRIPT PLUS is being delivered as an integrated component of the RFTS Program over a 4 year period by more than 100 nurses and social workers to about 600 pregnant smokers served by the state-wide Perinatal Services Division of the WV Bureau for Public Health.

Although patient willingness to participate and adherence are contributing factors to the overall implementation and success of any evaluation, future intervention evaluation studies planned by scientists and program leadership should critically review and apply the quantitative and qualitative methods, insight, and lessons learned from published systematic reviews (Ershoff, 2004; Fiore et al., 2000, 2008; Lumley et al., 2004; Melvin, Dolan-Mullen, Windsor, Whiteside, & Goldenberg, 2000; Mullen et al., 1994; Windsor, Boyd, et al., 1998; Windsor et al., 1990; Windsor & Orleans, 1986).

Recommendations for Policy-Program-Practice Leadership

Changing policies as well as the structure, process, and intensity of methods in existing treatment programs to support implementation of new, best practice behavioral health interventions is a complex and challenging activity. For any organization, meeting these challenges requires enthusiastic leadership, enduring commitment to institutionalization, competence, patience, modest resources, and sufficient time to plan and prepare staff to try out and adapt the best practice treatments. Direct involvement of agency/system-wide policy makers and program managers is essential. Insight from program staff who routinely counsel pregnant smokers, as well as from representative samples of patients should also be included at the onset of any organizational planning, implementation, and adoption process. Staff training and pilot-testing at all program sites of new patient

assessment and treatment methods are critical to identifying barriers and planning solutions for implementation.

At the onset of discussions regarding adoption of tobacco treatment programs, there needs to be an equal organizational commitment to the application of the well-described planning, process, and impact evaluation methods defined in the literature and this chapter. Ongoing dissemination research and evaluation of innovation within organizations and professional practices implementing evidence-based methods such as SCRIPT is essential. The literature and the 25-plus years of hands-on experience of the author in working with public health policy makers, managers, collaborating scientists, and colleagues in OB/prenatal care practice, along with interactions with patients, confirm that it takes a deliberate and conscious effort to integrate evidence-based interventions within existing programs.

The insight and expertise about how to achieve institutionalization of smoking cessation for pregnant smokers are available to leadership that has recognized smoking as a major cause of maternal, fetal, and infant risk, and has resolved to routinely provide "evidence-based" methods to reduce these risks. Our challenge now is to translate and disseminate science-based interventions into everyday practice. Given the efficacy and cost-effectiveness of evidence-based methods in increasing cessation rates among pregnant women, there is considerable potential for system-wide application of these methods at the national and local levels to improve reproductive health outcomes for individual women and the population of childbearing women overall.

Widespread adoption of evidence-based smoking cessation programs should also contribute to a decrease in adverse perinatal and reproductive outcomes among disadvantaged racial and ethnic groups. It is unclear, however, whether universal adoption and adherence to smoking cessation treatment methods during pregnancy will lead to a reduction in 'disparities' in adverse outcomes between racial/ethnic groups, as women of color, particularly African American women, appear to have lower rates of smoking during pregnancy than white women. Although, as suggested in this review, self-reported smoking rates may very well not be accurate. Nonetheless, the absolute potential for improvement of pregnancy outcomes due to implementation of effective smoking cessation interventions in all groups is unequivocal.

References

Adams, E., Solanki, G., & Miller, L. (1997). Medical-care expenditures attributable to cigarette smoking during pregnancy – United States, 1995. *MMWR: Morbidity & Mortality Weekly Report*, *46*(44), 1048–1050.

American College of Obstetricians and Gynecologists. (2005). Smoking cessation during pregnancy committee opinion number 316. *Obstetrics and Gynecology*, *106*(4), 883–888.

Andrews, J., Tingen, M., Waller, J., & Harper, R. (2001). Provider feedback improves adherence with AHCPR smoking cessation guideline. *Preventive Medicine*, *33*(5), 415–421.

Aquilino, W. (1994). Interview mode effects in surveys of drug and alcohol use: A field experiment. *Public Opinion Quarterly*, *58*(2), 210–240.

Arheart, K., Lee, D., Fleming, L., LeBlanc, W., Dietz, N., McCollister, K., et al. (2008). Accuracy of self-reported smoking and secondhand smoke exposure in the U.S. workforce: The National Health and Nutrition Examination Surveys. *Journal of Occupational and Environmental Medicine*, *50*(12), 1414–1420.

Augood, C., Duckitt, K., & Templeton, A. (1998). Smoking and female infertility: A systematic review and meta-analysis. *Human Reproduction*, *13*(6), 1532–1539.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Upper Saddle River, NJ: Prentice Hall.

Benowitz, N. (1991). Nicotine replacement therapy during pregnancy. *The Journal of the American Medical Association*, *266*(22), 3174–3177.

Benowitz, N. (1998). *Nicotine safety and toxicity*. New York: Oxford University Press.

Benowitz, N., & Dempsey, D. (2004). Pharmacotherapy for smoking cessation during pregnancy. *Nicotine & Tobacco Research*, *6* (Suppl. 2), S189–S202.

Bernstein, I., Mongeon, J., Badger, G., Solomon, L., Heil, S., & Higgins, S. (2005). Maternal smoking and its association with birth weight. *Obstetrics and Gynecology*, *106*(5), 986–992.

Bero, L., Grilli, R., Grimshaw, J., Harvey, E., Oxman, A., & Thomson, M. (1998). Getting research findings into practice: Closing the gap between research and practice: An overview of systematic reviews of interventions to promote the implementation of research findings. *BMJ: British Medical Journal*, *317*(7156), 465.

Blum, A., & Solberg, E. (1996). Real progress vs. public relations in tobacco control. *Seminars in Respiratory and Critical Care Medicine*, *17*(4) 285–288.

Borland, R., Segan, C., Livingston, P., & Owen, N. (2001). The effectiveness of callback counselling for smoking cessation: A randomized trial. *Addiction*, *96*(6), 881–889.

Boyd, N., Windsor, R., Perkins, L., & Lowe, J. (1998). Quality of measurement of smoking status by self-report and saliva cotinine among pregnant women. *Maternal and Child Health Journal*, *2*(2), 77–83.

Castles, A., Adams, E., Melvin, C., Kelsch, C., & Boulton, M. (1999). Effects of smoking during pregnancy five meta-analyses. *American Journal of Preventive Medicine*, *16*(3), 208–215.

Chapin, J., & Root, W. (2004). Improving obstetrician-gynecologist implementation of smoking cessation guidelines for pregnant women: An interim report of the American College of Obstetricians and Gynecologists. *Nicotine & Tobacco Research*, *6*, S253–S257.

Coleman, T., Antoniak, M., Britton, J., Thornton, J., Lewis, S., & Watts, K. (2004). Recruiting pregnant smokers for a placebo-randomized controlled trial of nicotine replacement therapy. *BMC Health Services Research*, *4*(29), 1–7.

Curry, S., Orleans, C., Keller, P., & Fiore, M. (2006). Promoting smoking cessation in the healthcare environment: 10 years later. *American Journal of Preventive Medicine*, *31*(3), 269–272.

Dewan, N., Brabin, B., Wood, L., Dramond, S., & Cooper, C. (2003). The effects of smoking on birthweight-for-gestational-age curves in teenage and adult primigravidae. *Public Health*, *117*(1), 31–35.

Donatelle, R., Prows, S., Champeau, D., & Hudson, D. (2000). Randomised controlled trial using social support and financial incentives for high risk pregnant smokers: Significant other supporter (SOS) program. *British Medical Journal*, *9*(Suppl. 3), 67–69.

Dornelas, E., Magnavita, J., Beazoglou, T., Fischer, E., Oncken, C., Lando, H., et al. (2006). Efficacy and cost-effectiveness of a clinic-based counseling intervention tested in an ethnically diverse sample of pregnant smokers. *Patient Education & Counseling*, *64*(1–3), 342–349.

Downs, S., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *British Medical Journal*, *52*(6), 377–384.

El-Mohandes, A. (2004–2009). *Reduction of tobacco exposure in pregnancy and infancy*. [ongoing clinical trial]. George Washington University Medical Center. National Institutes of Health Project Number 3U10HD036104-10S1.

England, L., Kendrick, J., Wilson, H., Merritt, R., Gargiullo, P., & Zahniser, S. (2001). Effects of smoking reduction during pregnancy on the birth weight of term infants. *American Journal of Epidemiology*, *154*(8), 694–701.

Ershoff, D. (2004). Helping pregnant women quit smoking: Progress and future directions. *Nicotine Tobacco Research*, *6*(Suppl. 2), S95–S277.

Ershoff, D., Mullen, P., & Quinn, V. (1989). A randomized trial of a serialized self-help smoking cessation program for pregnant women in an HMO. *American Journal of Public Health*, *79*(2), 182–187.

Ershoff, D., Quinn, V., Boyd, N., Stern, J., Gregory, M., & Wirtschafter, D. (1999). The Kaiser Permanente prenatal smoking-cessation trial when more isn't better, what is enough? *American Journal of Preventive Medicine*, *17*(3), 161–168.

Fiore, M., Bailey, W., Cohen, S., Dorfman, S., Goldstein, M., Gritz, E., et al. (1996). *Smoking cessation: Clinical practice guideline no. 18* (AHCPR publication no. 96-0692). Rockville, MD: U.S. Department of Health and Human Services. Agency for Health Care Policy and Research.

Fiore, M., Bailey, W., Cohen, S., Dorfman, S., Goldstein, M., Gritz, E., et al. (2000). *Treating tobacco use and dependence: Clinical practice guideline*. Rockville, MD: U.S. Department of Health and Human Services. Public Health Service.

Fiore, M., Jaen, C., Baker, T., Bailey, W., Benowitz, N., Curry, S., et al. (2008). *Treating tobacco use and dependence: 2008 update*. Rockville, MD: USDHHS, PHS, Agency for Healthcare Research and Quality.

Fiore, M., Smith, S., Jorenby, D., & Baker, T. (1994). The effectiveness of the nicotine patch for smoking cessation. *Obstetrical & Gynecological Survey*, *49*(11), 766–768.

Fleiss, J., Levin, B., & Paik, M. (2004). *Statistical methods for rates and proportions* (3rd ed.). Hoboken, NJ: Wiley-Interscience.

Gebauer, C., Kwo, C. Y., Haynes, E. F., & Wewers, M. E. (1998). A nurse-managed smoking cessation intervention during pregnancy. *Journal of Obstetric, Gynecologic, and Neonatal Nursing*, *27*(1), 47–53.

Gielen, A., Windsor, R., Faden, R., O'Campo, P., Repke, J., & Davis, M. (1997). Evaluation of a smoking cessation intervention for pregnant women in an urban prenatal clinic. *Health Education Research*, *12*(2), 247–254.

Glasgow, R., Bull, S., Gillette, C., Klesges, L., & Dzewaltowski, D. (2002). Behavior change intervention research in healthcare settings a review of recent reports with emphasis on external validity. *American Journal of Preventive Medicine*, *23*(1), 62–69.

Glasgow, R., Lichtenstein, E., & Marcus, A. (2003). Why don't we see more translation of health promotion research to practice? Rethinking the efficacy-to-effectiveness transition. *American Journal of Public Health*, *93*(8), 1261–1267.

Goodwin, R., Keyes, K., & Simuro, N. (2007). Mental disorders and nicotine dependence among pregnant women in the United States. *Obstetrics and Gynecology*, *109*(4), 875–883.

Greenland, S., Satterfield, M., & Lanes, S. (1998). A meta-analysis to assess the incidence of adverse effects associated with the transdermal nicotine patch. *Drug Safety*, *18*(4), 297–308.

Gupta, S., Hwang, S., Causey, D., Rolf, C., & Gorsline, J. (1995). Comparison of the nicotine pharmacokinetics of nicoderm (nicotine transdermal system) and half-hourly cigarette smoking. *The Journal of Clinical Pharmacology*, *35*(10), 985–989.

Hajek, P., West, R., Lee, A., Foulds, J., Owen, L., Eiser, J. R., et al. (2001). Randomized controlled trial of a midwife-delivered brief smoking cessation intervention in pregnancy. *Addiction (on-Line)*, *96*(3), 485–494.

Hartmann, K., Thorp, J., Pahel-Short, L., & Koch, M. (1996). A randomized controlled trial of smoking cessation intervention in pregnancy in an academic clinic. *Obstetrics and Gynecology*, *87*(4), 621–626.

Hasselmeyer, E., Meyer, M., Catz, C., & Longo L. D. (1979). Chapter 8: Pregnancy and infant health. In J. M. Pinney (Ed.), *Smoking and health: A report of the surgeon general* (DHEW Publication No. 79-0066, pp. 8.1–8.93) Rockville, MD: Public Health Service. Office on Smoking and Health.

Henningfield, J. (1995a). Introduction to tobacco harm reduction as a complementary strategy to smoking cessation. *Tobacco Control*, *4*, 25–27.

Henningfield, J. (1995b). Nicotine medications for smoking cessation. *New England Journal of Medicine*, *333*(18), 1196–1203.

Hjalmarson, A., Hahn, L., & Svanberg, B. (1991). Stopping smoking in pregnancy: Effect of a self-help manual in controlled trial. *British Journal of Obstetrics & Gynaecology*, *98*(3), 260–264.

Hoffmann, D., Hoffmann, I., & Wynder, E. (1998). *Observations on the changing cigarette, 1950 and 1997*. Paper presented at Canada's expert committee on cigarette toxicity reduction, Toronto, Canada.

Hruba, D., & Kachlik, P. (2000). Influence of maternal active and passive smoking during pregnancy on birthweight in newborns. *Central European Journal of Public Health*, *8*(4), 249–252.

Hughes, E., & Brennan, B. (1996). *The effects of active and passive smoking on human reproduction*. Paper presented at a meeting of Health Canada.

Hughes, J., Goldstein, M., Hurt, R., & Shiffman, S. (1999). Recent advances in the pharmacotherapy of smoking. *Journal of the American Medical Association*, *281*(1), 72–76.

Karen, G., Eliopoulis, C., & Klein, J. (1998). Chapter 10: Measuring fetal exposure to nicotine. In N. L. Benowitz (Ed.), *Nicotine safety and toxicity* (pp. 99–106). New York: Oxford University Press.

Kendrick, J., Zahniser, S., Miller N., Salas, N., Stine, J., Gargiullo, P., et al. (1995). Integrating smoking cessation into routine public prenatal care: The smoking cessation in pregnancy project. *American Journal of Public Health*, *85*(2), 217–222.

Kharrazi, M., DeLorenze, G., Kaufman, F., Eskenazi, B., Bernert Jr., J., Graham, S. et al. (2004). Environmental tobacco smoke and pregnancy outcome. *Epidemiology*, *15*(6), 660–670.

Kleinman, J., Pierre, M., Madans, J., Land, G., & Schramm, W. (1988). The effects of maternal smoking on fetal and infant mortality. *American Journal of Epidemiology*, *127*(2), 274–282.

Lawrence, T., Aveyard, P., Evans, O., & Cheng, K. K. (2003). A cluster randomised controlled trial of smoking cessation in pregnant women comparing interventions based on the transtheoretical (stages of change) model to standard care. *British Medical Journal*, *12*(2), 168–177.

Li, C., Windsor, R., Lowe, J., & Goldenberg, R. (1992). Evaluation of the impact of dissemination of smoking cessation methods on the low birthweight rate and on health care costs: Achieving year 2000 objectives for the nation. *American Journal of Preventive Medicine*, *8*(3), 171–177.

Li, C., Windsor, R., Perkins, L., Goldenberg, R., & Lowe, J. (1993). The impact on infant birth weight and gestational age of cotinine-validated smoking reduction during pregnancy. *Journal of the American Medical Association*, *269*, 1519–1524.

Lowe, J., Balanda, K., & Clare, G. (1998). Evaluation of antenatal smoking cessation programs for pregnant women. *Australian & New Zealand Journal of Public Health*, *22*(1), 55–59.

Lumley, J., Oliver, S., Chamberlain, C., & Oakley, L. (2004). Interventions for promoting smoking cessation during pregnancy. *Cochrane Database of Systematic Reviews*, (2), CD001055.

Martin, J., Hamilton, B., Sutton, P., Ventura, S., Menacker, F., & Kirmeyer, S. (2006). Births: Final data for 2004. *National Vital Statistics Reports*, *55*(1), 1–26.

Mathai, M., Skinner, A., Lawton, K., & Weindling, A. M. (1990). Maternal smoking, urinary cotinine levels and birthweight. *Australian and New Zealand Journal of Obstetrics and Gynaecology*, *30*(1), 33–36.

Mathews, T. (2001). Smoking during pregnancy in the 1990s. *National Vital Statistics Reports*, *49*(7), 1–14.

McAfee, T. (2007). Quit-lines: A tool for research and dissemination of evidence-based cessation practices. *American Journal of Preventive Medicine*, *33*(6S), 357–367.

McBride, C., Curry, S., Grothaus, L., Nelson, J., Lando, H., & Pirie, P. (1998). Partner smoking status and pregnant smoker's perceptions of support for and likelihood of smoking cessation. *Health Psychology*, *17*(1), 63–69.

Melvin, C., Dolan-Mullen, P., Windsor, R., Whiteside, H., & Goldenberg, R. (2000). Recommended cessation counselling for pregnant women who smoke: A review of the evidence. *Tobacco Control*, *9*(Suppl. 3), 80–84.

Meyer, M., Jonas, B., & Tonascia, J. (1976). Perinatal events associated with maternal smoking during pregnancy. *American Journal of Epidemiology*, *103*(5), 464–476.

Miller, D., Villa, K., Hogue, S., & Sivapathasundaram, D. (2001). Birth and first-year costs for mothers and infants attributable to maternal smoking. *Nicotine & Tobacco Research*, *3*(1), 25–35.

Misra, D., & Nguyen, R. (1999). Environmental tobacco smoke and low birth weight: A hazard in the workplace? *Environmental Health Perspectives*, *107*(Suppl 6), 897–904.

Moore, L., Campbell, R., Whelan, A., Mills, N., Lupton, P., Misselbrook, E., et al. (2002). Self help smoking cessation in pregnancy: Cluster randomised controlled trial. *British Medical Journal*, *325*(7377), 1383.

Mullen, P., Carbonari, J., Tabak, E., & Glanday, M. (1991). Improving disclosure of smoking by pregnant women. *American Journal of Obstetrics and Gynecology*, *165*(2), 409–413.

Mullen, P., Ramirez, G., & Groff, J. (1994). A meta-analysis of randomized trials of prenatal smoking cessation interventions. *American Journal of Obstetrics and Gynecology*, *171*(5), 1328–1334.

National Cancer Institute. (1999). *Health effects of exposure to environmental tobacco smoke: The report of the California environmental protection agency*. Smoking and tobacco control monograph no. 10. (NIH pub. no. 99-645). Bethesda, MD: U.S. Department of Health and Human Services.

National Center for Health Statistics. (2006). *Health, United States, 2006: With chartbook on trends in the health of Americans* (DHHS Publication No. 2006-232). Hyattsville, MD: U.S. Department of Health and Human Services.

National Committee for Quality Assurance. (2007). *Health-Plan Employer Data Information Set: HEDIS 2007 summary table of measures and product lines*. Available from http://www.ncqa.org/Portals/0/HEDISQM/Archives/2007/MeasuresList.pdf.

National Governors Association. (2007). *Appendix A. Medicaid births as a percentage of total births*. *Maternal and child health update: States increase eligibility fotates increase eligibility for children's health in 2007*. Washington, DC: National Governors Association. Retrieved from June 10, 2009, from http://www.nga.org/Files/pdf/0811MCHUPDATE.PDF.

Newell, D. (1992). Intention-to-treat analysis: Implications for quantitative and qualitative research. *International Journal of Epidemiology*, *21*(5), 837–841.

O'Connor, A., Davies, B., Dulberg, C., Buhler, P., Nadon, C., McBride, B., et al. (1992). Effectiveness of a pregnancy smoking cessation program. *Journal of Obstetric, Gynecologic, and Neonatal Nursing*, *21*(5), 385–392.

Olds, D., Robinson, J., O'Brien, R., Luckey, D., Pettitt, L., Henderson, C., et al. (2002). Home visiting by paraprofessionals and by nurses: A randomized, controlled trial. *Pediatrics*, *110*(3), 486–496.

Oncken, C., Dornelas, E., Greene, J., Sankey, H., Glasmann, A., Feinn, R., et al. (2008). Nicotine gum for pregnant smokers: A randomized controlled trial. *Obstetrics and Gynecology 112*(4), 859–867.

Oncken, C., Hardardottir, H., & Smeltzer, J. (1998). Human studies of nicotine replacement during pregnancy. In N. L. Benowitz (Ed.), *Nicotine safety and toxicity* (pp. 107–116). New York, NY: Oxford University Press.

Oncken, C., Hatsukami, D., Lupo, V., Lando, H., Gibeau, L., & Hansen, R. (1997). Effects of short-term use of nicotine gum in pregnant smokers. *Obstetrical & Gynecological Survey*, *52*(2), 89.

Orleans, C., Barker, D., Kaufman, N., & Marx, J. (2000). Helping pregnant smokers quit: Meeting the challenge in the next decade. *Tobacco Control*, *9*, III6–III11.

Orleans, T., Melvin, C., Marx, J., Maibach, E., & Vose, K. (2004). National action plan to reduce smoking during pregnancy: The national partnership to help pregnant smokers quit. *Nicotine & Tobacco Research*, *6*(Suppl. 2), 269–277.

Panjari, M., Bell, R., Bishop, S., Astbury, J., Rice, G., & Doery, J. (1999). A randomized controlled trial of a smoking cessation intervention during pregnancy. *Australian and New Zealand Journal of Obstetrics and Gynaecology*, *39*(3), 312–317.

Parker, D., Windsor, R., Roberts, M., Hecht, J., Hardy, N., Strolla, L., et al. (2007). Feasibility, cost, and cost-effectiveness of a telephone-based motivational intervention for underserved pregnant smokers. *Nicotine & Tobacco Research*, *9*(10), 1043–1051.

Pollak, K., McBride, C., Baucom, D., Curry, S., Lando, H., Pirie, P., et al. (2001). Women's perceived and partners' reported support for smoking cessation during pregnancy. *Annals of Behavioral Medicine*, *23*(3), 208–214.

Pollak, K., Oncken, C., Lipkus, I., Lyna, P., Swamy, G., Pletsch, P., et al. (2007). Nicotine replacement and behavioral therapy for smoking cessation in pregnancy. *American Journal of Preventive Medicine*, *33*(4), 297–305.

Prochaska, J., & DiClemente, C. (1983). Stages and processes of self-change of smoking: Toward an integrative model of change. *Journal of Consulting & Clinical Psychology*, *51*(3), 390–395.

Riemsma, R., Pattenden, J., Bridle, C., Sowden, A., Mather, L., Watt, I., et al. (2003). Systematic review of the effectiveness of stage based interventions to promote smoking cessation. *British Medical Journal*, *326*(7400), 1175–1176.

Rigotti, N., Park, E., Regan, S., Chang, Y., Perry, K., Loudin, B., et al. (2006). Efficacy of telephone counseling for pregnant smokers. *Obstetrics and Gynecology*, *108*(1), 83–92.

Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). New York: Free Press.

Rush, D., Orme, J., King, J., Eiser, J. R., & Butler, N. R. (1992). A trial of health education aimed to reduce cigarette smoking among pregnant women. *Paediatric and Perinatal Epidemiology*, *6*(2), 285–297.

Russell, T., Crawford, M., & Woodby, L. (2004). Measurements for active cigarette smoke exposure in prevalence and cessation studies: Why simply asking pregnant women isn't enough. *Nicotine & Tobacco Research*, *6*(Suppl. 2), 141–151.

Schroeder, D., Ogburn, P., Hurt, R., Croghan, I., Ramin, K., Offord, K., et al. (2002). Nicotine patch use in pregnant smokers: Smoking abstinence and delivery outcomes. *The Journal of Maternal-Fetal & Neonatal Medicine*, *11*(2), 100–107.

Secker-Walker, R., Solomon, L., Flynn, B., Skelly, J., Lepage, S., Goodwin, G., et al. (1994). Individualized smoking cessation counseling during prenatal and early postnatal care. *American Journal of Obstetrics and Gynecology*, *171*(5), 1347–1355.

Secker-Walker, R., Solomon, L., Flynn, B., Skelly, J., & Mead, P. (1998). Reducing smoking during pregnancy and postpartum: Physician's advice supported by individual counseling. *Preventive Medicine*, *27*(3), 422–430.

Sexton, M., & Hebel, J. (1984). A clinical trial of change in maternal smoking and its effect on birth weight. *Journal of the American Medical Association*, *251*(7), 911–915.

Silagy, C., Lancaster, T., Stead, L., Mant, D., & Fowler, G. (2004). Nicotine replacement therapy for smoking cessation (review). *Cochrane Database of Systematic Reviews*, (3), CD000146.

Simpson, W. J. (1957). A preliminary report on cigarette smoking and the incidence of prematurity. *Obstetrical & Gynecological Survey*, *12*(6), 808–815.

Slotkin, T. (1998a). Fetal nicotine or cocaine exposure: Which one is worse? *Journal of Pharmacology and Experimental Therapeutics*, *285*(3), 931–945.

Slotkin, T. (1998b). The impact of fetal nicotine exposure on nervous system development and its role in sudden infant death syndrome. In N. L. Benowitz (Ed.), *Nicotine safety and toxicity* (pp. 89–97). New York, NY: Oxford University Press.

Solomon, L., & Flynn, B. (2005). Telephone support for pregnant smokers who want to stop smoking. *Health Promotion Practice*, *6*(1), 105–108.

Solomon, L., Secker-Walker, R., Flynn, B., Christ, S., Dana, G., & Dorwaldt, A. (1996). Proactive peer support by telephone to help women quit smoking. *Health Education Research*, *11*, 377–381.

Spitzer, W., Lawrence, V., Dales, R., Hill, G., Archer, M., Clark, P., et al. (1990). Links between passive smoking and disease: A best-evidence synthesis. A report of the working group on passive smoking. *Clinical and Investigative Medicine. Medecine Clinique Et Experimentale*, *13*(1), 17–42; discussion 43–6.

Substance Abuse and Mental Health Services Administration. (1990–2004). *National Survey on Drug Use and Health (NSDUH) data*. Reports of the Office of Applied Studies. SAMHSA, U.S. Department of Health and Human Services.

Tappin, D., Lumsden, M., Gilmour, W., Crawford, F., McIntyre, D., Stone, D., et al. (2005). Randomised controlled trial of home based motivational interviewing by midwives to help pregnant smokers quit or cut down. *British Medical Journal*, *331*(7513), 373–377.

Tonnesen, P., Fryd, V., Hansen, M., Helsted, J., Gunnersen, A. B., Forchammer, H., et al. (1988). Effect of nicotine chewing gum in combination with group counseling on the cessation of smoking. *New England Journal of Medicine*, *318*(1), 15–18.

U.S. Department of Health, Education & Welfare, Public Health Service. (1979). *Healthy People: The Surgeon General's report on health promotion and disease prevention*. Washington, DC: U.S. Government Printing Office.

U.S. Department of Health and Human Services. (1991). *Healthy People 2000: National health promotion and disease prevention objectives*. Washington, DC: U.S. Government Printing Office.

U.S. Department of Health and Human Services. (2000). *Healthy People 2010: Understanding and improving health* (2nd ed.). Washington, DC: U.S. Government Printing Office.

Walsh, R. (1994). Effects of maternal smoking on adverse pregnancy outcomes: Examination of the criteria of causation. *Human Biology; an International Record of Research*, *66*(6), 1059–1092.

Walsh, R., & Redman, S. (1993). Smoking cessation in pregnancy: Do effective programmes exist? *Health Promotion International*, *8*(2), 111–127.

Walsh, R., Redman, S., Brinsmead, M., & Fryer, J. (1997). Predictors of smoking in pregnancy and attitudes and knowledge of risks of pregnant smokers. *Drug and Alcohol Review*, *16*(1), 41–67.

Webb, D., Boyd, N., Messina, D., & Windsor, R. (2003). The discrepancy between self-reported smoking status and urine cotinine levels among women enrolled in prenatal care at four publicly funded clinical sites. *Journal of Public Health Management and Practice*, *9*(4), 322–325.

Windham, G., Eaton, A., & Hopkins, B. (1999). Evidence for an association between environmental tobacco smoke exposure and birthweight: A meta-analysis and new data. *Paediatric and Perinatal Epidemiology*, *13*, 35–57.

Windsor, R. (1985). *A pregnant woman's guide to quit smoking*. Washington, DC: The Society for Public Health Education.

Windsor, R. (2001). Smoking, cessation and pregnancy. In J. Samet & S.-Y. Yoon (Eds.), *Women and the tobacco epidemic: Challenges for the 21st century* (pp. 147–164). Canada: World Health Organization. Proceedings of the WHO Meeting, Kobe, Japan, 1999.

Windsor, R. (2003). Smoking Cessation or Reduction in Pregnancy Treatment (SCRIPT) methods. *American Journal of Medical Sciences326*(4), 216–222.

Windsor, R. (2005). *A pregnant woman's guide to quit smoking*. Washington, DC: The Society for Public Health Education.

Windsor, R. (2007–2011). *The West Virginia smoking cessation or reduction in pregnancy treatment trial*. [ongoing clinical trial]. George Washington University Medical Center. National Institutes of Health Project Number 5R01CA12442903.

Windsor, R., Boyd, N., & Orleans, C. (1998). A meta-evaluation of smoking cessation intervention research among pregnant women: Improving the science and art. *Health Education Research*, *13*(3), 419–438.

Windsor, R., Clark, N., Boyd, N., & Goodman, R. (2004). *Evaluation of health promotion, health education and disease prevention programs* (3rd ed.). New York: McGraw-Hill.

Windsor, R., Cutter, G., Morris, J., Reese, Y., Manzella, B., Bartlett, E., et al. (1985). The effectiveness of smoking cessation methods for smokers in public health maternity clinics: A randomized trial. *American Journal of Public Health*, *75*(12), 1389–1392.

Windsor, R., Dalmat, M., Orleans, C., & Gritz, E. (1990). *The handbook to plan, implement and evaluate smoking cessation programs for pregnant women*. White Plains, New York: March of Dimes Birth Defects Foundation.

Windsor, R., Li, C., Lowe, J., Perkins, L., Ershoff, D., & Glynn, T. (1993). The dissemination of smoking cessation methods for pregnant women: Achieving the year 2000 objectives. *American Journal of Public Health*, *83*(2), 173–178.

Windsor, R., Lowe, J., Perkins, L., Smith-Yoder, D., Artz, L., Crawford, M., et al. (1993). Health education for pregnant smokers: Its behavioral impact and cost benefit. *American Journal of Public Health*, *83*(2), 201–206.

Windsor, R, Oncken, C., Henningfield, J., Hartmann, K., & Edwards, N. (2000). Behavioral and pharmacological treatment methods for pregnant smokers: Issues for clinical practice. *Journal of the American Medical Women's Association*, *55*(5), 304–310.

Windsor, R., & Orleans, C. (1986). Guidelines and methodological standards for smoking cessation intervention research among pregnant women: Improving the science and art. *Health Education Quarterly*, *13*(2), 131–161.

Windsor, R., Warner, K., & Cutter, G. (1988). A cost-effectiveness analysis of self-help smoking cessation methods for pregnant women. *Public Health Reports*, *103*(1), 83.

Windsor, R., Whiteside, H., Solomon, L., Prows, S., Donatelle, R., Cinciripini, P., et al. (2000). A process evaluation model for patient education programs for pregnant smokers. *Tobacco Control*, *9*(Suppl. 3), 29–35.

Windsor R., Woodby, L., & Crawford M. (1998). *Commit to quit during and after pregnancy* [video]. Birmingham, AL: University of Alabama at Birmingham Medical Television. (Note: Edited by Windsor for West Virginia Bureau of Public Health in 2004 + 2006).

Windsor, R., Woodby, L., Miller, T., Hardin, J., Crawford, M., & DiClemente, C. (2000). Effectiveness of agency for health care policy and research clinical practice guideline and patient education methods for pregnant smokers in Medicaid maternity care. *American Journal of Obstetrics and Gynecology*, *182*(1 Part 1), 68–75.

Wisborg, K., Henriksen, T., Jespersen, L., & Secher, N. (2000). Nicotine patches for pregnant smokers: A randomized controlled study. *Obstetrics and Gynecology*, *96*(6), 967–971.

Woodby, L., Windsor, R., Snyder, S., Kohler, C., & Diclemente, C. (1999). Predictors of smoking cessation during pregnancy. *Addiction*, *94*(2), 283–292.

Wright, L., Thorp Jr, J., Kuller, J., Shrewsbury, R., Ananth, C., & Hartmann, K. (1997). Transdermal nicotine replacement in pregnancy: Maternal pharmacokinetics and fetal effects. *American Journal of Obstetrics and Gynecology*, *176*(5), 1090–1094.

Zhu, S., Anderson, C., Tedeschi, G., Rosbrook, B., Johnson, C., Byrd, M., et al. (2002). Evidence of real-world effectiveness of a telephone quitline for smokers. *New England Journal of Medicine*, *347*(14), 1087–1093.

Zhu, S., Stretch, V., Balabanis, M., Rosbrook, B., Sadler, G., & Pierce, J. (1996). Telephone counseling for smoking cessation: Effects of single-session and multiple-session interventions. *Journal of Consulting and Clinical Psychology*, *64*(1), 202–211.

Chapter 12
Substance Use in Pregnancy: The Impact of Screening and Treatment on Improving Perinatal Outcomes and Reducing Racial and Ethnic Disparities

Suzanne Carlberg-Racich and Ellen Mason

Introduction

Women's use of alcohol and psychoactive substances[1] before conception and during pregnancy is common and occurs in all ethnic groups, all socio-economic strata and all geographic regions of the U.S. The biologic and social effects of substance use on perinatal outcomes and on the long-term physical and psychological development of exposed children are not fully understood, despite active research for over three decades. Though substantial evidence exists to link substance use in pregnancy to a wide variety of adverse maternal, fetal and neonatal outcomes, it is often not possible to attribute a specific negative outcome directly to the effects of a particular substance. Many adverse outcomes are due to the effects of multiple drugs, drugs and alcohol used in combination, and these combined with concurrent tobacco use. Since smoking is the focus of another chapter in this volume, this chapter will focus primarily on the use of alcohol as well as illicit and illegal drugs during pregnancy. (Please see Box 12.1 for explanation of how these terms are used here).

The chapter will begin with an overview of what is known about the prevalence of alcohol and illicit substance use by pregnant women and those at risk for pregnancy, and a summary of key associated perinatal outcomes. This information will provide background and perspective for a brief review of perinatal substance use screening, and for a more comprehensive review of the treatment of substance use disorders in pregnant women. The review will focus on evidence that substance use screening and treatment improve perinatal outcomes. To the extent that there is variation among racial/ethnic groups in patterns of substance use, rates and methods of screening, and responses to treatment modalities, conclusions about the effects of screening and interventions are relevant to efforts to reduce racial and ethnic disparities in maternal and infant health outcomes.

[1]Use of the word "substance" in this chapter refers to any solid, liquid, or vapor that can be ingested orally, inhaled, sniffed, snorted, smoked, or injected into the body, for the purpose of mood alteration whether for recreational purposes, for self-medicating physical or emotional symptoms, or other reasons.

S. Carlberg-Racich (✉)
Training Specialist, Midwest AIDS Training and Education Center (MATEC), University of Illinois at Chicago, 1640 W. Roosevelt Road, Suite 511 (M/C 779), Chicago, IL 60608, USA
e-mail: scarlb1@uic.edu

A. Handler et al. (eds.), *Reducing Racial/Ethnic Disparities in Reproductive and Perinatal Outcomes: The Evidence from Population-Based Interventions*, DOI 10.1007/978-1-4419-1499-6_12,

Box 12.1 Vocabulary and Definitions[2]

Substance Dependence and Substance Abuse

Substance Dependence

Substance dependence is defined as a maladaptive pattern of substance use leading to clinically significant impairment or distress, as manifested by three (or more) of the following, occurring any time in the same 12-month period:

Tolerance, as defined by either of the following: (a) A need for markedly increased amounts of the substance to achieve intoxication or the desired effect or (b) Markedly diminished effect with continued use of the same amount of the substance.

Withdrawal, as manifested by either of the following: (a) The characteristic withdrawal syndrome for the substance or (b) The same (or closely related) substance is taken to relieve or avoid withdrawal symptoms.

The substance is often taken in larger amounts or over a longer period than intended.

There is a persistent desire or unsuccessful efforts to cut down or control substance use.

A great deal of time is spent in activities necessary to obtain the substance, use the substance, or recover from its effects.

Important social, occupational, or recreational activities are given up or reduced because of substance use.

The substance use is continued despite knowledge of having a persistent physical or psychological problem that is likely to have been caused or exacerbated by the substance (for example, current cocaine use despite recognition of cocaine-induced depression or continued drinking despite recognition that an ulcer was made worse by alcohol consumption).

Substance Abuse

A. A maladaptive pattern of substance use leading to clinically significant impairment or distress as manifested by one (or more) lf the following, occurring within a 12 month period:

   1. Recurrent substance use resulting a failure to fulfill major role obligations at work, school or at home, including but not limited to neglect of children or household.[3]
   2. Recurrent substance use in situations in which it is physically hazardous (e.g. driving an automobile or operating machinery)
   3. Recurrent substance-related legal problems
   4. Continued substance use despite having persistent or recurrent social or inter-personal problems caused or exacerbated by the effects of the substance (e.g. spousal arguments, physical fights)

B. The symptoms have never met the criteria for Substance Dependence for this class of substance.

[2] Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision 2000, Section on Substance-Related Disorders. American Psychiatric Association, Arlington, VA.

[3] It should be noted that the term *substance abuse* rather than *use* is frequently employed in perinatal research and clinical practice. This is because *any* use of illegal mood altering drugs, *any* illicit use of pharmaceutically produced mood altering drugs (see below) or *any* use of alcohol by pregnant women has historically been considered *abuse* under an extrapolation of Criteria #1 for Substance Abuse, which assumes that substance use in pregnancy constitutes high probability of risk of harm to maternal health or potential harm to fetal growth and well-being.

(continued)

Box 12.1 (continued)

Illegal and Illicit Drugs[4]

The term *illegal drug* is applied to drugs classified by the United States Drug Enforcement Administration (DEA) as "Schedule I", meaning they are substances with high potential for abuse, have *no* currently accepted medical use in treatment in the United States and lack accepted safety for use under medical supervision. Examples include heroin, mescaline, and psilocybin.

The term *illicit drug* is applied to substances classified by the DEA as Schedule II, III, IV and V. These are substances that do have an accepted medical use in treatment in the United States but can lead to psychological or physiologic dependence and abuse, which is reduced as the Schedule progresses (V is less than IV, which is less than III, etc.). The DEA prohibits or restricts use of narcotic drugs, as well as depressant or stimulant substances by assigning felony drug offense charges to non-prescribed, unsupervised or improper use of these substances under specified conditions and in specified amounts.

For the Sake of Brevity and Ease of Reading, All Illegal and Illicit Substances Will Be Referred to as "Illicit Substances"

Under DEA definitions, illicit drugs may also comprise substances that are available for over the counter purchase, without a prescription, such as dextromethorphan containing cough syrups and pseudoephedrine-containing cold and cough remedies. Other substances which may be used for mood alteration, and which may cause physical dependence and harm, primarily inhalant substances such as toluene, aerosols, correction fluid and paint thinner, are not classified as illicit drugs by the DEA, although these substances may be abused and cause dependence.

[4] (U.S. Drug Enforcement Administration: www.usdoj.gov/dea).

Prevalence of Alcohol and Drug Use in Pregnancy: The Scope of the Problem

Reliable prevalence data on drug and alcohol use in pregnancy are difficult to obtain. Most population-based U.S. data on substance use in the general population are derived from national surveys, the major source being the National Survey on Drug Use and Health (NSDUH) conducted annually by the Substance Abuse and Mental Health Services Administration (SAMHSA). The NSDUH collects information on recent tobacco, alcohol and illicit drug use (use in the last 30 days), along with socio-demographic information (e.g., age, gender, ethnicity, current contraceptive use and pregnancy status). Frequency of alcohol use is recorded as number of drinks per week, and survey items on illicit drugs inquire about specific popular drugs of abuse. Another source of information is the Behavioral Risk Factor Surveillance System (BRFSS), which was developed by the Centers for Disease Control and Prevention (CDC) and is conducted monthly in all 50 states, the District of Columbia, Puerto Rico, the U.S. Virgin Islands, and Guam. The Pregnancy Risk Assessment Monitoring System (PRAMS), also conducted by CDC, collects information about alcohol and tobacco use from pregnant and recently delivered women in 23 participating states and the District of Columbia. The most comprehensive and detailed information about illicit drug, alcohol and tobacco use in pregnancy is derived from population-based epidemiologic studies that gather data exclusively from pregnant women. Studies like this are very resource-intensive and hence infrequently conducted. In 1992, the National Institute on Drug Abuse (NIDA) conducted the National Pregnancy and Health Survey (NPHS), the first and only national survey of drug use in pregnancy. The NPHS collected self report data on substance use from a national sample of 2,613 women delivering in 52 urban and rural health care facilities (NIDA, 2006).

In regularly collected national surveys on drug and alcohol use in the general U.S. population, questions on perinatal use vary from year to year and survey to survey. Individual surveys have posed different questions about use during the preconception period, during pregnancy and postpartum. Older surveys sometimes contain detailed data about perinatal substance use, broken down by age groups, ethnic groups, socioeconomic status and education status, while more recent surveys may not provide this level of detail. In an attempt to provide the most complete prevalence data, including data on racial and ethnic groups, the information presented below comes from both older and more recent surveys.

Alcohol Use in Pregnancy

In the 2007 NSDUH survey, rates of *any* and *heavy* alcohol use (heavy use being five or more drinks on the same occasion on each of 5 or more days in the past 30 days) during pregnancy by women aged 15–44 were estimated to be 11.6% and 0.7% respectively (SAMSHA, 2008), 6.6% of pregnant women reported binge drinking (greater than 5 drinks at one occasion or on 1 day) during their first trimester. More recent data from the NSDUH (SAMHSA, 2009) looked at drinking rates in later pregnancy and found that rates of alcohol use dropped as pregnancy progressed, with 7.8% of women drinking in the second trimester and 6.2% of women drinking in the third trimester of pregnancy. Summary data from the NSDUH indicate that overall drinking rates in pregnancy have remained stable throughout the past decade (Denny, Tsai, Floyd, & Green, 2009).

Perinatal alcohol use data from other sources provide information on use in different regions, as well as additional estimates of use. PRAMS data show significant state-by-state variability, with rates of alcohol use during the last 3 months of pregnancy ranging from 3.4% in Nebraska to 9.9% in Colorado (Phares et al., 2004). BRFSS data gathered from 1991 to 1999 found that rates of binge and frequent drinking (the latter defined in this survey as more than 7 drinks per week) during pregnancy were about 2.7% and 3.3% respectively (Sidhu & Floyd, 2002). The 1992 NPHS had a larger sample of pregnant women than any of the above sources and found that 18.8% of women drank alcohol during pregnancy.

Alcohol use in both pregnant and non-pregnant women varies by age and race/ethnicity. The 2004 NSDUH survey (SAMHSA, 2005) found that pregnant Native American/Alaskan American women and non-Hispanic white women were the most likely of any of the ethnic group to binge or drink heavily.

Given the current high rates of unplanned pregnancy among young adult women along with the documented negative impact of early alcohol exposure in pregnancy, it is important to look at the overall rates of drinking among non-pregnant women. The rates are highest for women aged 18–25, with 57.9% of women in this demographic reporting alcohol consumption within the last month. Heavy alcohol use and binge drinking are also relatively high in this age group. In 2006, 26.6% of young adult women reported binge drinking in the past month (SAMHSA, 2007). Since alcohol consumption itself is associated with elevated risk of unintended pregnancy (Project CHOICES Research Group, 2002), these patterns of heavy and binge drinking among young women may partly account for the worrisome rates of fetal alcohol diagnoses found in national health reports (Stratton, Howe, & Battaglia, 1996; CDC, 1995).

National surveys also report patterns of alcohol use among postpartum women. A recent analysis of NSDUH data from 2002 to 2007 found overall drinking rates of 63.0% in women who were not pregnant and had no children living in the household. While alcohol use rates during pregnancy are much lower, drinking rates appear to rise rapidly in postpartum women. Women who had children less than 3 months old in their household had a 31.9% alcohol use rate, while rates were as high as 54.9% among women with children older than 14 months. Postpartum increases in drinking rates are cause for concern both because of the potential for alcohol teratogenicity in the periconceptional period in subsequent pregnancies and the potential for increased risk of postpartum child abuse and neglect, particularly for mothers who resume binge drinking after pregnancy (SAMHSA, 2008).

Illicit Drug Use in Pregnancy

Among pregnant women aged 15–44 years, the NSDUH found that approximately 5.2% reported using illicit drugs in the past month (SAMHSA, 2008). Additionally, the NSDUH found that rates of illicit drug use by pregnant women remained the same between 2004–2005 and 2006–2007. Interestingly, data from the NPHS done 15 years earlier found the same percentage of illicit drug use as is currently reported – 5.5% among pregnant women. In the older study, the rates of use of specific drugs were broken out, with 2.9% reporting use of marijuana or cocaine. The NPHS also looked at polysubstance abuse and found that 20.4% of their sample smoked cigarettes. The survey found a high correlation between cigarette smoking and the use of alcohol or illicit drugs during pregnancy.

Alcohol and Drug Use by Age, Ethnicity, Income and Education

In NSDUH surveys, very young pregnant women report high rates of illicit drug use when compared to their older counterparts. In the 2007 survey, pregnant women aged 18–25 had drug use rates of about 7.2%. Younger pregnant women had even higher drug use rates, and unlike any other pregnant group studied, had higher rates of illicit drug use than their non-pregnant counterparts (22.6 vs. 13.3%; SAMHSA, 2008). Similarly, BRFSS data also indicate relatively high rates of illicit drug use in younger reproductive age women.

Regarding racial/ethnic and income disparities in the use of illicit drugs, recent information is sparse. Data from the 1992 NPHS indicated that more white women than African-American or Latina women reported *ever* using illicit drugs, but the percentage who used drugs at least monthly was highest among African-Americans and Latinas. Similarly, women with higher incomes were more likely to *ever* have used illicit drugs, but monthly users were more likely to be lower income (Howell, Heiser, & Harrington, 1999). Data from the 2006 NSDUH survey which looked at gender and ethnicity indicated that illicit drug use rates among non-pregnant women were highest for white women at 11.6%, compared to 9.4% for African-American women and 7.4% for non-pregnant Hispanic women. Although more recent NSDUH surveys have not collected pregnancy data by ethnicity, in the 2004 NSDUH survey of pregnant women, African-Americans had the highest rate (8.0%), followed by 4.4% for white women and 3.0% for Hispanics (SAMHSA, 2005). In that survey, African-American women had the smallest difference between non-pregnant and pregnant use of illicit drugs (9.4% and 8% respectively), suggesting that African-American women were less likely than other groups to curtail use when they become pregnant.

Surveys of substance use in pregnancy, however carefully executed, have inherent limitations that can bias study findings. Despite use of the most sensitive and innovative interviewing techniques available to maximize the level of honest reporting and assure the confidentiality of respondents, responses may be affected by recall and reporting errors. The impact of maternal denial, guilt, stigma and fear of legal consequences is a decrease in the likelihood that illicit substance use will be accurately reported. Additionally, since most major surveys rely on use of telephones or mail to contact subjects, there is likely under-representation of women who are homeless, have insecure housing or are institutionalized. These are vulnerable populations in whom substance use problems may be higher, or who may suffer greater impact from such use. On the whole, these sources of bias are most likely to result in underestimates of the prevalence of substance misuse in pregnancy.

Given the limitations of self-report survey data, an important epidemiologic approach to estimating prevalence rates of substance use by pregnant women is the anonymous and routine testing of pregnant women or their neonates for drugs and drug metabolites. Testing of urine is most common, although neonatal meconium and hair have also been used. In the 1990s, individual states including

California (Vega, Kolody, Hwang, & Noble, 1993), Utah (Buchi & Varner, 1994), Rhode Island (Hollinshead et al., 1990) and South Carolina (Nalty, 1991) conducted large-scale studies of substance use in pregnancy using anonymous peripartum maternal and neonatal urine samples, sometimes coupled with neonatal meconium analysis to estimate prevalence of maternal drug and/or alcohol use. These state-wide studies found rates of drug or alcohol use that were equal to or greater than contemporaneous national perinatal self report surveys.

Effects of Substance Use on Pregnancy Outcomes, Infant, Child, and Women's Health

Antenatal use of alcohol and psychoactive drugs is significantly associated with a number of adverse pregnancy outcomes. For example, placenta previa, placental abruption, fetal growth restriction and preterm delivery have all been linked to vasoactive drugs such as cocaine or amphetamines and to frequent or binge use of alcohol (Hulse, English, Milne, Holman, & Bower, 1997; Ananth, Berkowitz, Savitz, & Lapinski, 1999; Macones, Sehdev, Parry, Morgan, & Berlin, 1997; Handler, Mason, Rosenberg, & Davis, 1993). Although the lay public and media often connect use of illegal drugs during pregnancy to birth defects, only antenatal alcohol use is predictably associated with an increased risk of congenital anomalies. Alcohol-related birth abnormalities were initially called fetal alcohol syndrome (FAS), because the triad of growth restriction, neurobehavioral abnormalities and facial dysmorphology was the first discrete clinical entity to be associated with prenatal drinking (Jones & Smith, 1973) As the field of alcohol related teratology has grown, FAS has become a subset of alcohol-mediated abnormalities, now referred to by broader terms such as fetal alcohol spectrum disorder (FASD), or fetal alcohol effect (FAE; Sokol, Delaney-Black, & Nordstrom, 2003; Weber, Floyd, Riley, & Snider, 2002). The combined prevalence of FAS and other alcohol related disorders is estimated to be 9.1 per 1,000 births (Sokol et al., 2003).

Opioid use in pregnancy also has detrimental effects on the newborn, including a form of neonatal withdrawal known as Neonatal Abstinence Syndrome (NAS). Approximately 50% of infants exposed to heroin will exhibit signs of withdrawal within 24–48 hours after birth. Most cases of NAS require observation in a neonatal intensive care nursery (NICU; American Academy of Pediatrics, 1998; Saiki, Lee, Hannam, & Greenough, 2010; O'Grady, Hopewell, & White, 2009). Although symptoms can be sustained and severe, the majority of infants do not suffer from mortality or long term morbidity.

Developmental effects of maternal substance use on children are challenging to study, given a multiplicity of confounding factors, such as poverty, inadequate access to health services, increased rates of child neglect and mistreatment, as well as increased risk of foster care placement and other home disruptions (Schempf, 2007). A recent review of evidence on maternal antenatal substance use and child development found overwhelming evidence for negative neuro-developmental consequences of prenatal alcohol exposure (Thompson, Levitt, & Stanwood, 2009). The review also found that prenatal cocaine exposure is associated with behavioral abnormalities which increase the likelihood of affected children requiring special needs programs. Hans (1996) has reported subtle behavioral abnormalities in opiate-exposed children followed longitudinally.

Finally, very important domains of women's health are impacted by substance use. Extensive literature correlates substance use in pregnancy with increased risks of violence-related injury, intimate partner violence, sexual abuse, psychiatric complications and high rates of infections including hepatitis and HIV (Tardiff et al., 1994; Thompson & Kingree, 1998; Berenson, Stiglich, Wilkinson, & Anderson, 1991; Bauer et al., 2002). A review of maternal mortality in California found that parturient women who used drugs or alcohol had double the rate of maternal death compared to non-using counterparts (Wolfe, Davis, Guydish, & Delucchi, 2005)

Challenges to Establishing Efficacy of Screening and Treatment

Though the impacts of alcohol and illicit substance use during pregnancy on women and their offspring can be substantial, they can theoretically be ameliorated if women (and infants) at risk are identified early and provided with appropriate services. However, efforts to study the efficacy of screening and treatment of substance use during pregnancy face a number of challenges. For example, studies of the impact of screening and treatment often focus on intermediate outcomes, such as utilization of prenatal care rather than primary outcomes such as birth weight or neurological deficits. Furthermore, although drug and alcohol exposure during fetal life may affect neurodevelopment and behavior throughout childhood and young adulthood, it is difficult to maintain and follow cohorts of exposed children to monitor these long-term effects and the possible ameliorating influences of screening combined with antepartum or early postpartum interventions. The Maternal Lifestyles Study (MLS) is a notable example of a study that was designed and resourced to achieve such aims through extensive longitudinal follow-up. The MLS is a prospective multicenter study sponsored by NIH that gathers longitudinal data on cocaine and opiate-exposed children to assess the long-term effects of substance exposure as well as interventions on school performance and child behavior. Results from the MLS show that prenatal cocaine exposure is associated with behavioral abnormalities that can be identified by targeted screening of pre-school and school age children (Messinger et al., 2004).

Timing of screening and intervention for substance use during pregnancy, as well as the "dose" of treatment delivered are variables that can be difficult to measure and that vary considerably across studies. Ethical considerations preclude randomized trials that put at-risk women into a condition of "no treatment". Studies may compare women participating in a "full push" intervention to those receiving less intensive intervention, or the outcomes of women and infants who fully participate in a program may be compared to those who drop out or who do not fully adhere to the intervention. Such a design can be problematic since women who are less engaged in the program may be more severely affected by substance use or have fewer resources than their more fully engaged counterparts. Furthermore, in controlled studies, women in the treatment group who continue substance use during the intervention may have positive outcomes related to their receipt of intensive and personal attention, making it a challenge to identify the specific components of the intervention that create positive change in outcomes. Even extreme care in data analysis may not eliminate such biases.

Finally, the fact that illicit drug abuse is a criminal activity constrains the ability of investigators to study the problem. Although a landmark United States Supreme Court decision in 2001 affirmed that pregnant women cannot legally be non-consensually tested for illicit drugs (Annas, 2001), women who use illicit drugs during pregnancy continue to experience and fear legal repercussions if they choose to access health services or participate in research studies. Most substance-using pregnant women (and many health care providers) are unaware that drug and alcohol testing must be consensual, unless it is done in an acute care setting such as an emergency room or a delivery suite. Pregnant women often fear that testing under any circumstances, whether consensual or not, will result in notification of child protective service agencies or even result in criminal charges. These fears are not unreasonable, given that federal statutes mandate reporting of suspected child abuse or neglect by health care providers and researchers. Since child protective statutes have often been interpreted in a way that equates antepartum maternal substance use with "in utero child abuse" and postpartum use with "child neglect", perinatal substance abuse is often equated with de facto child abuse and neglect (Chavkin, 1991). Fear of experiencing negative consequences can prevent women from disclosing their substance use and can serve as a significant barrier to their participation in substance use research/treatment.

With these limitations in mind, we will briefly discuss the current status of screening for the use of illicit drugs as well as alcohol during pregnancy, followed by a systematic review of the evidence for efficacy of treatment approaches in optimizing perinatal outcomes.

Screening for Illicit Drug and Alcohol Use

Screening for alcohol and recreational drug use in pregnant women and other women of childbearing age is an established part of prenatal and periconceptional care. In 1995, the March of Dimes (MOD) developed and disseminated educational materials to train obstetric providers to screen prenatal patients for substance use. Currently, many leading professional societies, such as the American College of Obstetrics and Gynecology (ACOG), the American Academy of Family Physicians (AFP) and the American College of Nurse Midwifery recommend universal screening of all prenatal patients coupled with patient education as a strategy for preventing or minimizing perinatal substance exposure, and for identifying high-risk women who may benefit from interventions aimed at improving perinatal outcomes and long-term maternal health status. In primary care settings, questions about licit and illicit substance use are commonly incorporated into new patient health histories. They are also part of most pre-formatted obstetric care instruments such as the Hollister Prenatal and Newborn Record. The American Academy of Pediatrics (1998) has also developed recommendations and standards for selective screening of newborns for drug and alcohol exposure so that further evaluation and intervention can be offered. Despite these recommendations, substance use screening is not always conducted by providers, even in prenatal care settings. One study utilized qualitative methodology to gather information from providers and patients on substance use screening and prevention practices in the first prenatal visit (Chang et al., 2008). They found that providers were most likely to engage patients in risk assessment and risk reduction counseling for cigarette smoking, and were less likely to do the same for alcohol or illicit drug use. When substance abuse was addressed, the conversations were brief and focused on referral to specialty treatment services.

Despite the fact that screening is recommended practice, the evidence for the benefits of screening women for illicit drug use is not robust. In a recent statement, The United States Preventive Services Task Force (USPSTF, 2008) concluded that although illicit drug use is one of the ten leading preventable risk factors for death and disability among adolescents, adults and pregnant women, there is insufficient evidence to assess the balance of benefits and harms of screening. The USPSTF noted that validated methods for detecting drug use/abuse are lacking, as are studies that would establish the short and long-term efficacy of interventions offered to individuals, particularly pregnant women who are not seeking treatment. Therefore, the USPSTF concluded that existing evidence does not support current screening recommendations and practices, and that screening for drug use has the potential for harm. For example, low-income women of color are more likely to be asked about or tested for drug use than other groups (Kerker, Leventhal, Schlesinger, & Horwitz, 2006). Published results of perinatal screening may therefore reinforce biased public stereotypes about who uses substances in pregnancy. Screening may also serve to disproportionately penalize low-income or minority women. The early 1990s found a trend toward increasing prosecution of poor African-American women for child abuse based on antenatal drug use, with the number tripling in only 2 years (1990–1992; Breitbart, Chavkin, & Wise, 1994). Both the American College of Nurse Midwives and ACOG oppose any legislation which criminalizes the use of substances by pregnant women; however, many providers appear to support an approach in which illicit drug users are mandated to treatment, which can in some cases result in a criminal record (Abel & Krueger, 2002). It is important to recognize that provider attitudes toward substance use in pregnancy can be reflected in patient care, and can affect the quality of assessment and risk reduction counseling they provide to their pregnant patients. One author advises that "above all, the obstetrician must remember that he or she is dealing with a patient with a medically recognized disease, and not merely moral deficiency or criminal intent" (Christensen, 2008).

In contrast to illicit drug use screening, there is more consensus about the value of screening for alcohol use and misuse in pregnant women and children, and non-pregnant women in their reproductive years, particularly when interventions accompany the screening. In 2004 the United States

Preventive Services Task Force (USPSTF, 2004) issued a report recommending screening, coupled with behavioral interventions to reduce alcohol misuse in all adults, including pregnant women. Screening and intervention is recommended in primary care settings, including prenatal clinics. The guideline states that all pregnant women and women contemplating pregnancy should be informed of the harmful effects of alcohol on the fetus. In this clinical guideline, all pregnant women screening positive for any alcohol use should receive a brief motivational interview using a behavioral counseling framework based on the "5 A's" (Assess, Advise, Agree, Assist, Arrange) that enhances patient knowledge, motivation to change and self-help skills. This intervention has been given a grade B recommendation (high certainty that there is at least a moderate net benefit) by the USPSTF. In the case of this recommendation, screening by itself, along with brief counseling is considered a useful intervention by the Task Force although the guideline acknowledges that more research is needed to establish the efficacy of screening and behavioral intervention for pregnant women. The discussion of treatment in this chapter expands upon this issue.

What valid uses are there then for perinatal drug and alcohol screening? Regarding alcohol, the "best" method of screening women of reproductive age for alcohol use is not currently known, nor has any single screening instrument been identified as having the highest yield of true positives in pregnant women. Similarly, the frequency at which women should be screened for alcohol use before or during pregnancy is not established. Given the evidence that small amounts of alcohol taken during pregnancy can have measurable effects on children's growth and development, a number of screening instruments have been developed for use in reproductive age women as well as prenatal populations. These instruments are specifically designed to focus less on heavy drinking and more on "at risk drinking", meaning any use of alcohol that has the potential to cause fetal harm (Sokol, Martier, & Ager, 1989; Chang et al., 1998). Use of specialized perinatal alcohol screening instruments (in contrast to standardized instruments for adults such as the 4-item CAGE or the 10-item AUDIT screening tests) has been shown to increase identification of at-risk drinking prior to or during pregnancy. Among the questionnaires designed for use in pregnancy, the T-ACE alcohol questionnaire has been shown to be superior in identifying alcohol use in pregnancy in comparison with clinician-administered brief face to face screens (Chang et al., 1998). However, these studies do not establish that improved identification of risky pregnant drinkers alone translates to improved perinatal outcomes in the absence of accompanying interventions (Chang et al., 2008). Similarly, routine use of toxicology testing or testing for biomarkers (such as alcohol metabolites or abnormal levels of, proteins, enzymes or fatty acids associated with alcohol use in pregnant women or their offspring,) in the absence of intervention, has not been proven useful in improving maternal and infant health outcomes or reducing racial/ethnic disparities (Bearer et al., 2005).

Whether it is for alcohol or illicit drugs, perinatal screening done in isolation from intervention has two main objectives: when performed anonymously, screening is useful in estimating prevalence of use in specific regions or institutions. In these types of investigations, maternal history is often coupled with toxicology testing (Ostrea, Brady, Gause, Raymundo, & Stevens, 1992). Similarly, perinatal drug screening studies often combine maternal substance use interviews or specific substance use screening instruments with toxicology testing or biomarkers of drug or alcohol use in order to establish the superiority of one interview approach or one testing instrument over others (Russell et al., 1996). In both types of screening studies, a serious drawback is the lack of a "gold standard" for establishing use. Toxicology testing of blood or urine (maternal or neonatal) is often treated as a gold standard that unequivocally establishes substance use; however, commonly used toxicology tests can only provide evidence of use at some point in time, without quantifying the amount used, or the frequency of use. Very recently, improved drug tests have been developed for evaluating neonatal meconium and hair, which are better standardized and more reliable than other tests used now or in the past. These new tests offer the possibility of accurately quantifying drug exposure during the course of gestation (Araojo, McCune, & Feibus, 2008). Future studies of the benefits of routine neonatal screening for drug or alcohol exposure may be more meaningful if they use these superior

biologic tests, and combine them with careful diagnostic evaluation of exposed children and rigorous evaluation of the effect that special services and interventions have for these children.

From Screening to Treatment: Factors Influencing Access and Utilization

As discussed above, screening for alcohol and substance misuse is common in many health care settings, but screening by itself may be insufficient to generate desired improvements in perinatal outcomes, unless followed by treatment. Even when women are properly screened, or identify their own need for treatment, they may be unable to access it. In fact, the research literature documents a shortage of treatment options for pregnant women through the 1980s, including a disturbing study in which two thirds of major hospitals were unaware of any services to which they could refer their pregnant patients for substance use disorder treatment (Howell et al., 1999). Although data from the 1990s suggest improvements in access for pregnant women, barriers to treatment remained, including a lack of providers willing to accept Medicaid as a payment option, and a lack of childcare services. Although acceptance of Medicaid payment for treatment has improved, challenges persist to finding funded treatment options for pregnant women. For example, in a study of pregnant women in five U.S. cities, being a Medicaid recipient and needing child care remained significant barriers to accessing treatment. Thus, poor women with children were less likely to be able to access care than those with more financial resources or who did not already have young children (Breitbart et al., 1994).

Some women who have access to treatment may choose not to enter a program. Haller, Miles, and Dawson (2003) compared women who enrolled in treatment to women who chose not to enroll. The two groups were demographically similar – mostly African-American with a high school education and lack of employment, and a mean age of 27. Those who chose to enroll were more likely to report use of crack cocaine as their "drug of choice" and to experience more severe addiction. In comparison to those who declined treatment, they were also more likely to report the following: family problems, psychiatric problems or emotional distress, and greater involvement with criminal justice or the legal system. The authors proposed that those who chose not to enter may not have felt as much need for intensive treatment. Thus, treatment options which reflect the entire spectrum of use may encourage women who see their use as less problematic to be more receptive to assistance during pregnancy. Some work is currently being done to assist providers in determining the extent of treatment required by an individual. For example, Christensen (2008) proposed a set of questions to assist physicians in determining the level of treatment that a pregnant patient may require, from minimal outpatient support to intensive, medically managed inpatient facilities. Such questions cover a range of issues that may suggest a need for a higher level of treatment; examples are current intoxication or potential for experiencing withdrawal (particularly where withdrawal may compromise the pregnancy, such as in alcohol or opioid addiction), medical conditions or complicating factors, co-occurring disorders, patient readiness for treatment, history of relapse, current living situation, and funding. Such efforts may promote entry to treatment at many levels of use and be assessed for their ability to improve maternal and infant outcomes.

Another deterrent to entering treatment is the predominant emphasis on achieving abstinence, or on entering treatment drug-free. Pressures to achieve abstinence during pregnancy may result in stress and convince women not to seek help for their addiction (Winklebaur et al., 2008). The cycle of struggling to be abstinent and then relapsing again into use, a struggle that is commonplace in addiction, can cause distress to a fetus. Christensen (2008) concurred, comparing addiction to other chronic conditions such as diabetes and hypertension, both of which are well-known for cycles of relapse. However, such chronic conditions are not stigmatized like addiction, which is often viewed as a character weakness rather than a medical problem. These authors' efforts exemplify a recent trend in treatment approaches that address addiction as a chronic illness rather than a moral failing.

Box 12.2 The Wellstone and Domenici Act

The Paul Wellstone and Pete Domenici Mental Health Parity and Addiction Equity Act stipulates that benefits for mental health and substance abuse cannot be any more restrictive than benefits for medical or surgical care, including frequency of treatment, number of visits, days of coverage, etc. However, the Act is far from perfect. For example, it does not require that companies offer mental health or substance abuse treatment benefits to their employees. It also provides exemptions for companies under 50 employees and those that can prove financial hardship associated with the Act (SAMHSA, 2008). Nevertheless, it is a step toward parity in insurance coverage where little currently exits, and has the potential to improve coverage for millions of people, including those enrolled in Medicaid managed-care plans and SCHIP programs for disadvantaged families with children on the state level.

A final deterrent to entering treatment is the lack of adequate coverage for the cost of addiction treatment, even when individuals are insured. Insurance coverage often has limits which may be inadequate to address an individual's addiction, and may interfere with the necessary retention to produce long-term change. However, this may be changing. In January of 2010, the Paul Wellstone and Pete Domenici Mental Health Parity and Addiction Equity Act (see Box 12.2) went into effect; this legislation is designed to increase access to mental health/substance use treatment benefits and reduce discrimination in healthcare, a necessary first step in improving access.

Disparities in Treatment for Substance Use Disorders

Although the literature on substance use in pregnancy does not contain a great deal of research regarding racial and ethnic disparities in maternal and infant outcomes, the general substance abuse literature documents a number of disparities related to treatment, and these are likely relevant for the pregnant substance using population. Minorities have been found to have poorer alcoholism treatment outcomes than Caucasians, which may be related to a host of factors, including: less engagement in or access to treatment, receiving services of inferior quality, reporting lower satisfaction with services, having inadequate retention in programs, and having reduced access to routine medical care. Such disparities remain after controlling for a variety of potential confounders, including differences in income, insurance coverage, and education (Schmidt, Greenfield, & Mulia, 2006). This is compounded by more intensive treatment needs of minorities, as indicated by a documented disparity in negative health and social consequences resulting from drinking, even when socioeconomic status is accounted for. Despite this disparity in need, African Americans and Hispanics do not report participation in programs such as Alcoholics Anonymous at the same rate as Caucasians, which may be due to barriers such as lack of locally available services, lack of knowledge of available services and lack of child care (Schmidt et al., 2006). Similarly, disparities in involvement with the criminal justice system may inhibit participation in treatment for minority populations. In one study, African American women were disproportionately reported for court intervention and mandatory treatment, at a rate ten times higher than Caucasians (Chasnoff, Landress, & Barrett, 1990), which may lead to greater stigma attached to treatment. Such concerns are compounded by pregnancy in that many women fear losing children if their drug use is discovered (Wolfe, Guydish, Santos, Delucchi, & Gleghorn, 2007). Additional barriers that may also influence treatment utilization for pregnant women and increase disparities include lack of transportation, heightened stigma of drug use in pregnancy, lack of provider knowledge about managing

substance use disorders in pregnant women, and provider refusal to accept Medicaid as payment for services.

Disparities related to gender may amplify disparities related to race and ethnicity. Women's treatment needs are inherently complex, as women may enter treatment with a host of gender-related bio-psycho-social issues to be addressed, and these can be compounded by pregnancy (Uziel-Miller & Lyons, 2000). General treatment programs may not be adept at providing services that meet both culture and gender-based needs. For example, Lewis (2004) conducted an in-depth qualitative exploration into the substance use disorder treatment needs of African-American women and found that a lack of African-American counselors served as a significant barrier to recovery in a treatment center. Similarly, the dearth of African-American staff made women feel that the environment was not a "safe space" for their children, often resulting in having their children stay with relatives even when childcare services were available at the program. One way to reduce gender-related barriers to treatment is to employ gender-based approaches, in which specific needs of women are incorporated into the treatment design and additional supportive services are offered. Studies suggest that such approaches result in better completion rates among women than programs with a mixed-gender focus (Niv & Hser, 2007). Women-focused programs may involve a host of services including case management, family therapy, childcare, pregnancy-related care, individual counseling, professional development, parenting classes, and "trauma-informed" programming to address the role of trauma in women's addictions. This type of program is reviewed later in the chapter.

Retention in treatment has been found to increase treatment success, but African-American women often do not remain in treatment a sufficient amount of time to ensure success (Lewis, 2004). Adequate prenatal care is also critically important for pregnant substance using women. Programs that integrate prenatal care and treatment for substance use disorders have been found to result in improved birth weights for infants born to cocaine-dependent mothers (Chazotte, Youchah, & Comerford Freda, 1995). Other studies have demonstrated the value of tailoring treatment programs to meet the specific needs of clients, particularly pregnant and parenting women, for whom educational and legal barriers may inhibit success. For example, providing opportunities to complete a high school degree can result in better treatment retention and improved outcomes (Knight, Logan, & Simpson, 2001). Along with treatment retention, early identification of substance use disorders in pregnancy is critical for achieving either a reduction in use or abstinence (Corse & Smith, 1998).

There is a growing body of evidence suggesting that overall, pregnant women who obtain substance abuse treatment have better birth outcomes (such as fewer days in the NICU, higher gestational age, higher birth weight, and better Apgar scores) when compared to pregnant women who do not receive treatment. There is also evidence that treatment is cost-effective, with a mean estimated savings of $4,644 per mother-infant pair, mostly due to avoiding the potentially extensive NICU costs associated with the non-treatment population (Svikis et al., 1997). There are a wide variety of treatment options, and it is important to identify those that have demonstrated efficacy in improving outcomes for pregnant women and their infants. The following is a systematic review of substance abuse treatment types and their efficacy in the context of pregnancy.

A Systematic Review of Treatment Efficacy in Pregnant, Substance-Using/Addicted Women: Perinatal Outcomes and Health Disparities

There are many challenges in determining the efficacy of treatment for substance use disorders among pregnant women, and assessing whether such treatment reduces racial/ethnic disparities in perinatal outcomes. There is a general shortage of literature in this area, due in part to the fact that

access to treatment for pregnant women has historically been limited. A wide variety of clinical trials and other biomedical studies have excluded women of childbearing potential (Greenfield et al., 2007). This phenomenon has resulted particularly in a dearth of studies examining the impact of pharmacotherapy treatment approaches for pregnant women (Rayburn & Bogenschutz, 2004). Methadone maintenance therapy (MMT) is a case in point. In general there is solid research and abundant literature on the efficacy of MMT, but studies that include pregnant women and that focus on perinatal outcomes are relatively rare and suffer from methodological limitations. There are studies on the relationship between MMT program retention and perinatal outcomes, but sample sizes are small and of limited generalizability due to recruitment from single treatment sites rather than involving larger, multi-site trials (Burns, Mattick, Lim, & Wallce, 2006). Despite the limitations in the data, MMT remains the standard of care for pregnant, opioid-dependent women, and is discussed below.

Attempts to connect treatment interventions with measurable outcomes have many of the same limitations already noted concerning efforts to tie illicit drug use to adverse perinatal outcomes. Factors such as poverty, poor nutrition, lack of access to prenatal care, reproductive tract infections, physical and sexual abuse, unwanted pregnancy, stress, mental health symptoms and lack of social support may have unmeasured influences on the effectiveness of substance use interventions. Studies may be biased due to having samples derived primarily from prenatal care environments, potentially missing those at greatest risk (Schempf, 2007). Small sample sizes in general can hamper the ability to evaluate effectiveness. Finally, previously described issues with underreporting drug and alcohol consumption, particularly during pregnancy, may bias results of intervention effectiveness studies.

Treatment Methodologies

Despite the limited research on the efficacy of treatment in pregnancy, there are some promising treatment options to assist in achieving more favorable perinatal outcomes. The following review summarizes current scientific evidence for the efficacy of substance use disorder treatment interventions for preventing adverse maternal and infant outcomes, and includes the following treatment types: medications, outpatient behavioral treatment, and inpatient treatment (National Institute on Drug Abuse, 2009). A brief explanation of each type of treatment follows.

Pharmacotherapy (Medications)

Substance use disorders may be treated with medications designed either to replace the illicit substance by helping to maintain normal brain function and reduce cravings, or for detoxification purposes (medically assisted withdrawal), and can be used in conjunction with behavioral therapies. Such medications usually act directly on the specific receptors in the brain that are occupied by the substance of abuse for example, agonists are medications that fully activate the receptors and stabilize them thereby reducing cravings for the drug, and include methadone for opioid dependence and nicotine replacement therapies such as the patch or nicotine gum. Partial agonists do not fully activate the receptors and may also act as competitive antagonists which reduce the effects of a full agonist (for example, if a person takes heroin while on a partial agonist therapy, it may reduce the effects of the heroin). Examples of partial agonists include buprenorphine for opioid dependence and varenicline for nicotine dependence. Finally, antagonists have a greater affinity for and bind tightly to the receptors, but they do not activate the receptors, and thus block the effects of the drug. For example, naltrexone is an antagonist used for both opioid and alcohol dependence; it acts as a barrier to experiencing the effects of either the opioid or the alcohol.

Outpatient Behavioral Treatments

Outpatient behavioral treatments are designed to engage and retain patients in treatment, address issues related to their addiction, and assist them in developing life skills. Behavioral treatments can be offered on an outpatient or inpatient (residential) basis, and may be used in conjunction with medication-based therapies. Some examples follow: Cognitive Behavioral Therapy assists patients in identifying factors that facilitate their drug use, and in learning to avoid or respond appropriately without using drugs. Multidimensional Family Therapy is usually focused on adolescent drug/alcohol users and helps to identify maladaptive patterns and influences that trigger use, for both the adolescent and their family. Motivational Interviewing is a communication or counseling technique designed to assist patients in resolving ambiguity about change and increasing readiness for treatment, and may be used in a variety of treatment settings. Finally, Contingency Management (also called Motivational Incentives or "pee for pay") involves using incentives (money, vouchers, etc.) to encourage retention in treatment and abstinence from drugs. Such incentives are used in a variety of programs as positive reinforcement for patients who have good attendance records or urine screening tests that are negative for illicit drugs. Outpatient behavioral treatment may be brief in nature, or may require multiple visits over time.

Residential (Inpatient) Treatment

Inpatient treatment may involve many of the behavioral treatments listed previously, but is designed to assist an individual in developing a drug or alcohol free lifestyle through intensive programming while remaining housed in a residential facility. Patients may be in inpatient programs which utilize Cognitive Behavioral Therapy, or a 12-step support model. Programs often have a great deal of structure and may be time-intensive, as much as 6–12 months for Therapeutic Communities. Such programs are perhaps most appropriate for those with longer, more problematic addiction histories and impaired social function.

Systematic Review Methodology

The following search engines were employed for this chapter: MEDLINE, CINAHL, Popline, WHO Reproductive Health Library, Web of Science, Cochrane Library, OCLC First Search and Academic Search Elite. Papers were included that met the design criteria, namely randomized controlled trials, observational studies, mixed-methods studies, meta-analyses, and expert reports. Search terms were utilized in a variety of combinations with each search engine to be sure not to miss any significant papers. For example, *substance use, substance abuse, chemical dependency and addiction* were combined with *treatment, pregnant, pregnancy and perinatal*. Searching on *racial/ethnic disparities*, did not yield satisfactory search results, so citations were perused manually to find papers that discussed disparities. A broad array of outcomes were considered in the search as well, including: treatment retention, abstinence, infant and maternal health outcomes, etc. As this chapter evolved over more than a year, multiple searches were performed to seek more recent articles that may have appeared since the first search. Articles published prior to 1985 were excluded from the review, as were those that were more descriptive in nature, and did not examine outcomes. In the course of the searches, some major review articles were found that documented outcomes. Though papers reviewed elsewhere are not included in our systematic review, we have provided an overview of the major review articles in order to summarize the literature and identify any gaps not covered by these reviews. Following this overview (summarized in Table 12.1) is our review of articles not reviewed elsewhere.

Table 12.1 Summary of literature syntheses and meta-analyses related to treatment outcomes among pregnant women and their infants

Study author(s)	Source	Period covered	Review purpose	Primary study (*n*)	Effect sizes	Conclusions
Ashley et al. (2003)	The American Journal of Drug and Alcohol Abuse	1980–2000	Review the effectiveness of treatment for women, using articles with demonstrated outcomes	$n = 38$ total, $n = 17$ had pregnant or postpartum women as focus	N/A	• Evidence was demonstrated for the following program elements: child care services, prenatal care, women-only admissions or women-specific services or workshops, mental health services, and comprehensive approaches • Cited limitations to evidence included: a preponderance of studies lacking randomization, small sample sizes, use of self-report data, lack of follow-up, and differences in outcome measures, instruments, etc.
Greenfield et al. (2007)	Drug and alcohol dependence	1975–2005	Review the effectiveness of treatment for women, using articles with demonstrated outcomes	$n = 283$ total, $n = 31$ had pregnant or postpartum women as focus	N/A	• Women may be less likely to enter treatment than men • Women and men have comparable rates of retention and completion in treatment • Programs which are tailored to the needs of women may be more efficacious • Limitations included: small sample sizes and lack of randomization
Handmaker and Wilbourne (2001)	Alcohol Research and Health	Not specified	Review of interventions to address alcohol use as offered in prenatal care clinics, which measured outcomes	$n = 22$	N/A	• Even brief interventions in prenatal care can impact outcomes (retention in treatment, higher birth weights, reduction in use) • Observed limitations included: a lack of studies incorporating a comparison group, small sample sizes, and inability to control for confounders

(continued)

Table 12.1 (continued)

Study author(s)	Source	Period covered	Review purpose	Primary study (*n*)	Effect sizes	Conclusions
Howell et al. (1999)	Journal of Substance Abuse Treatment	1980–1998	Review of all literature related to the effectiveness of treatment among pregnant women	*n* = 16 prevalence studies, *n* = 19 SA and Outcomes (birthweight, gestational age) *n* = 17 treatment effectiveness	N/A	• Findings from limited studies are consistent with the non-pregnant treatment literature • Treatment retention is associated with increased success at reducing use • More intensive service provision results in improved retention • Pregnant women in MMT programs have greater retention than opiate-dependent pregnant women who are not on MMT • Provision of child care is associated with greater retention • Residential programs followed by outpatient are associated with greater retention and outcomes • Availability of research regarding substance abuse treatment effectiveness is limited, due to small sample sizes and lack of randomization
Minozzi et al. (2008)	The Cochrane Library	Through 2007	Review of outcomes related to the effectiveness of maintenance agonist treatments for opioid-dependent pregnant women, as compared to placebo or another intervention, or maintenance plus psychosocial intervention	*n* = 3	N/A	• No difference in outcomes (NAS, birth weight) were consistently noted between methadone and buprenorphine • Observed limitations included: a dearth of studies meeting the inclusion criteria, small sample sizes, and not controlling for cigarette smoking, a potential confounder

Rayburn and Bogenschutz (2004)	American Journal of Obstetrics and Gynecology	1979–2003	To review available literature examining the role of pharmacotherapy in pregnant women	Unknown *n*	N/A	• Pharmacological interventions (with counseling) are indicated for alcohol withdrawal (benzodiazepines) and opioid dependence (methadone, buprenorphine)
Terplan and Lui (2007)	The Cochrane Library	Through 2006	To assess the effectiveness of psychosocial interventions for pregnant drug using women on the following variables: Perinatal outcomes Treatment attendance and retention Maternal and infant abstinence	$n = 9$	N/A	• Contingency Management interventions resulted in better treatment retention, while motivational interviewing appears to have either no effect or a negative effect on treatment retention
Winklebaur et al. (2008)	Addiction	1998–2006	To review recent studies related to treatment of pregnant women who are opioid dependent	Unknown *n*	N/A	• Maintenance therapy is the dominant, recommended approach for opioid dependent women • Treatment of NAS deserves special attention

Summary of Review Articles (Located in Table 12.1)

Minozzi, Amato, Vecci, and Davoli (2008) conducted a systematic Cochrane review of clinical trials comparing opioid substitution medications during pregnancy. Only three trials met the criteria for the review. Two compared methadone with buprenorphine, while the third compared methadone with slow-dose morphine. Sample sizes were quite small, and outcomes (such as retention in treatment) were similar between the groups. In one study, slow-release morphine appeared to be effective in assisting women to abstain from illicit opioids. However, the study in question utilized a simple urine screen that does not distinguish morphine from other illicit opioids, relying instead on examination of patients for evidence of injection to determine whether illicit opioids were used. Patients who used illicit opioids in a non-injected form may have been inadvertently recorded as non-users. In addition, use of visual inspection may have resulted in failure to detect more "hidden" injection sites such as the groin. Finally, the three studies were plagued by other methodological limitations, such as small sample sizes and not controlling for cigarette smoking, which can produce symptoms similar to Neonatal Abstinence Syndrome and decrease birth weight in infants of heavy users. Nevertheless, it appears that buprenorphine and methadone may be similarly efficacious for retaining pregnant women in treatment, while slow-dose morphine may be more advantageous for assisting with abstinence from illicit opioids. Slow-dose morphine remains an exciting possibility for treatment among pregnant women, though further studies are needed to address some of the methodological concerns above.

Terplan and Lui (2007) conducted a Cochran review of psychosocial interventions in outpatient settings for illicit drug users. Randomized controlled trials were included if they included an experimental group that received a psychosocial intervention, and a control group who received a different psychosocial intervention, a pharmacological intervention or no intervention/placebo. Nine studies were included in the review; five included contingency management techniques and four used motivational interviewing. The authors concluded that psychosocial interventions did not show sufficient evidence of improvements in either infant outcomes or drug abstinence. Contingency management showed efficacy in retaining women in treatment, but motivational interviewing did not, and in fact the latter had poorer retention than control groups in some cases. Studies in this review were subject to a variety of limitations, including the exclusion of women with poor obstetrical outcomes, heterogeneity in outcomes examined, and failure to distinguish women who were mandated into treatment from those who volunteered (which is important because interventions may be less effective among coerced individuals).

Another Cochrane review conducted by the same group examined the effectiveness of psychosocial interventions in improving outcomes for pregnant women dependent on alcohol (Lui, Terplan, & Smith, 2008). Though the authors found a number of articles relevant to the topic ($n = 26$), none met the inclusion criteria of comparing psychosocial treatment to other treatment methods or no treatment in order to assess efficacy. The review is therefore not included in Table 12.1, but it is nonetheless worth mentioning if only to highlight the notable gap in the literature. Implications for future research include the obvious need for randomized controlled trials or quasi-experimental designs to address the dearth of as outcomes-based research regarding alcohol treatment and psychosocial interventions in pregnancy.

Ashley, Marsden, and Brady (2003) completed a review of substance abuse treatment among women, both pregnant and non-pregnant. Included were 38 studies with demonstrated outcomes, such as treatment retention, birth weight, gestational age, number of prenatal care visits, abstinence, or changes in any of the following: drug use behaviors, criminal activity, care giving for dependent children, housing status, and self-esteem. Seventeen of these studies were focused on the perinatal/postpartum population. While less systematic than the Cochrane reviews summarized above, this review was able to identify several program components that showed consistency in improving treatment outcomes for women. These included: provision of child care services, prenatal care,

women-only admissions, or services or workshops specific to the needs of women, mental health services, and comprehensive approaches to treatment. All of these components are included in the Continuum of Care established by the Treatment Improvement Protocol (TIP) for Pregnant, Substance-Using Women (SAMHSA, 1995), which is intended to guide providers in meeting the needs of women in this population. The authors also identified various weaknesses in the literature examined, including the lack of randomized design, inadequate sample sizes, the reliance on self-report data, lack of follow-up to track consistency in outcomes, and wide ranges in outcome measures, instruments, etc. across studies, making it more difficult to discern trends in the literature.

Greenfield and colleagues (2007), reviewed a much larger group of studies (n = 283) dealing with gender and substance abuse treatment, of which 31 focused specifically on perinatal or post-partum populations. Among these 31 studies, treatment modalities included both pharmacotherapy treatment and behavioral approaches, and both maternal and infant outcomes were documented. This review concurred with the findings of Ashley and colleagues (2003) that programs tailored to the needs of women may be more efficacious. Consistent with the screening and access discussion above, the authors also discussed the finding that women may be less likely to enter treatment than men, although treatment retention and completion rates appeared comparable. These reviews indicate the possibility that the inclusion of services tailored to the needs of women (such as child care) may remove significant barriers to treatment.

Howell and colleagues (1999) reviewed 19 studies related to substance abuse treatment efficacy among pregnant women, some of which addressed perinatal outcomes such as birthweight and gestational age. They noted that findings were generally consistent with those in the substance abuse treatment literature for non-pregnant persons, namely that success increased as a result of longer retention in treatment and more intensive service delivery As in other reviews, service enrichment such as provision of child care services was associated with greater retention. In addition, opioid-dependent women who received methadone maintenance stayed in treatment longer than those who did not receive methadone. (This is consistent with another review of pharmacological treatments in pregnancy, which concluded that methadone is indicated for opioid dependence during pregnancy; Rayburn & Bogenschutz, 2004). Again, several limitations were noted in the Howell review, including small sample sizes and lack of randomized designs, and a general dearth of studies that demonstrate impact of interventions on outcomes.

According to a systematic review of treatment for pregnant women who use opioids (Winklebaur et al., 2008), the following strategies have demonstrated effectiveness: contingency management, psychosocial support (particularly earlier in pregnancy), and maintenance therapy with methadone (the "gold standard"), and buprenorphine (not as well documented in the literature due to arriving much later on the U.S. market). The authors note an association between use of MMT and more "standardized, intense" prenatal care, which together result in better perinatal outcomes when compared to ongoing use of illicit opioids. They also note a dose-response relationship for MMT, with greater improvement at higher dosing. European studies of buprenorphine show similar results to MMT, with evidence of less severe neonatal abstinence syndrome (NAS). Clearly, MMT and buprenorphine are valuable treatment options for opioid-dependent pregnant women, and can be integral in improving outcomes for both mothers and infants.

Recent evidence suggests that rates of binge or frequent drinking by pregnant women have not declined substantially despite overall reductions in alcohol consumption during pregnancy (Chang, McNamara, Oray, & Wilkins-Haug, 2006). Since many pregnant women do not seek treatment for alcohol problems, interventions in prenatal care settings can be particularly critical. Handmaker and Wilbourne (2001) reviewed evidence for effectiveness of alcohol-related interventions that are incorporated into prenatal care. Twenty-two studies met the inclusion criteria for review, namely that the study took place in connection with prenatal care services, included a clear measurement of drinking, and considered any of a variety of outcomes as measures of effectiveness. Most of the studies involved "brief interventions" and supportive counseling delivered in the prenatal care setting, while others were

more intensive and involved home visits, case management and referral to treatment services. The authors noted that many interventions are associated with reductions in drinking or increased levels of abstinence, but that the shortage of controlled studies with sufficiently large samples and appropriate comparison groups can make it difficult to conclusively attribute the results to the interventions; this is especially problematic since many women spontaneously reduce or abstain from alcohol consumption during pregnancy. On the whole, the authors concluded that the literature reviewed appears "consistent with the broader treatment literature, which shows that brief interventions and motivational interventions have strong track records for reducing alcohol consumption by both problem drinkers and dependent drinkers" (Handmaker & Wilbourne, 2001). There was also some evidence that babies born to women receiving such interventions had higher birth weights.

Systematic Review of Studies Not Included in Review Articles (See Table 12.2)

Brief Interventions for Alcohol Use in Pregnancy

One of the studies reviewed by Handmaker and Wilbourne (2001) that showed positive results for a prenatal alcohol intervention was conducted by Chang and colleagues (2000). This same group of investigators conducted a second randomized trial after publication of the review article. In this study, 304 pregnant women who had a positive result on the T-ACE alcohol screening were enrolled in a trial of a single-session brief intervention that was offered to both the woman and her partner (Chang et al., 2005). Though women in both the treatment and control groups reduced alcohol consumption over the course of pregnancy, women in the treatment group experience greater reductions. The intervention was most effective among women with higher rates of drinking at enrollment; a subgroup analysis suggested that for these heavier drinkers, partner involvement enhanced the effectiveness of the intervention, suggesting a need for further studies to explore the role of a partner in supporting pregnant women as they attempt to change their substance use behavior (Chang et al., 2005). In a separate analysis, these investigators confirmed their earlier findings that goal selection during the brief intervention was highly predictive of subsequent drinking behavior (Chang et al., 2006). Specifically, those women who chose abstinence as their goal were more likely to achieve that goal or to reduce alcohol consumption than were women who chose "cutting down" as a goal, indicating the importance of goal selection in positive behavior change.

Opioid Substitution Therapy

Withdrawal from opioids during pregnancy can result in miscarriage or premature labor (Burns et al., 2006), as well as fetal distress and intrauterine meconium aspiration. Methadone maintenance therapy (MMT) seeks to stabilize the pregnant woman and reduce the probability of encountering withdrawal, due to the 24-h time-released dose. Methadone as a substitute also guards against the ingestion of street-purchased opioids of uncertain dose and potential unknown adulterants in the "cut" of the drug. Evidence-based benefits of methadone plus prenatal care during pregnancy for the mother include reduced maternal mortality and pregnancy complications. Benefits for the baby include lower fetal morbidity and fetal wastage, while enhancing growth and stability (Burns et al., 2006). Such mutual mother-baby benefits make this a clear treatment choice during pregnancy for opioid-dependent women.

Opioid-using women often experience amenorrhea and may therefore assume that their lack of menstrual cycle is due to the drugs rather than a pregnancy (Burns et al., 2006). As a consequence, many women who chronically use opioids either fail to enter treatment when pregnant, or enter late in the pregnancy. Both early entry and retention in treatment are associated with positive outcomes.

Similarly, relapse and attrition are related to poor infant and maternal outcomes (Jones, Haug, Stitzer, & Svikis, 2000, reviewed elsewhere). For opioid-addicted women, retention in methadone maintenance therapy (MMT) is a key factor in improving perinatal outcomes, and several studies show a linear relationship between time spent in treatment and improved outcomes, such as retention in prenatal care and increased birth weights (Burns et al., 2006). Further, engaging in methadone therapy for the purposes of detoxification or short-term transition off of heroin has not been shown to be effective. Burns and colleagues (2006 – covered in another review) examined outcomes related to entry into MMT and retention and found that earlier entry to treatment resulted in reduced prematurity when compared to those who entered treatment late in pregnancy. Also, the study found that those who entered treatment late were more likely to have infants who needed specialized care or admittance into the NICU. However, those with late entry also had higher smoking rates than those in the early entry group, a possible confounder. Similarly, polysubstance use was not addressed in this study and it cannot be determined whether effects were due to time at entry into care or the use of other substances. Despite these limitations, the study is consistent with other studies indicating that MMT and prenatal care support better fetal growth.

A cohort study of 260 infants in France examined mothers who received either methadone or high-dose buprenorphine (Lejeune, Simmat-Durand, Gourarier, Aubisson, & Groupe d'Etudes Grossesse et Addictions, 2006). In the French system, all pregnant women (including undocumented women or those without a form of payment) have health coverage, and the study population included 35 different public perinatal centers. Data were collected via self-report using a very comprehensive instrument, and the study employed a multidisciplinary team whose goal was to assist the patient in building a solid parent-child bond and reduce potential chaos in the environment caused by addiction. The authors noted that women with poor prenatal care had higher rates of premature delivery than those with good care. The authors concluded that there were no measurable differences in infant outcomes when comparing methadone with buprenorphine (such as gestational age, birth weight, Neonatal Abstinence Syndrome or NAS, and Lipsitz score), and could therefore be considered comparable treatments for this population. This study was prospective in nature, eliminating the potential problems with recall inherent in retrospective studies. However, the study was not blinded, and patients were not randomized to either the buprenorphine or methadone treatment groups, making findings somewhat less compelling. Methadone requires enrollment in specialty programs while buprenorphine can be prescribed by participating physicians in the French system. This resulted in a higher proportion of study enrollees taking buprenorphine therapy as opposed to methadone. Also, the study included only live births, perhaps biasing the results by not including pregnancies with poorer outcomes. The study also did not include women who discontinued use of either therapy, following only those women who continued until delivery.

Kakko, Heilig, and Sarman (2008) conducted another comparison of methadone and buprenorphine in which infant outcomes were assessed. They compared population-based prospective data on pregnant women taking buprenorphine treatment to retrospective data on a group of pregnant women on methadone maintenance therapy. The purpose was to assess any differences in infant outcomes related to the use of each medication. The study found that women taking MMT had lower birth weight infants than those taking buprenorphine, although when gestational age was assessed as a covariate, statistical significance of this finding disappeared. The authors also found that NAS was two times more severe for infants whose mothers took MMT. This finding remained significant even after Bonferroni correction, revealing that 21% of infants whose mothers took buprenorphine required treatment for NAS, while 68% of infants required treatment after their mothers took methadone. This translated into statistically significant longer lengths of stay in the hospital for infants born to MMT recipients. The results suggested that buprenorphine may have some advantages in terms of infant outcome. Limitations to the study included a lack of random assignment, resulting in very different populations in each treatment group and perhaps biasing the results. Similarly, the two groups were studied in different time periods, making time a possible confounder in the results.

Table 12.2 Efficacy of substance abuse treatment in pregnant women: summary of reviewed studies

Authors and year of publication	Study design	Study type and journal	Description of intervention (what, how, where)	Population studied (ages, race ethnicity, sample size)
Burns et al. (2006)	Cross Sectional-Record Linkage, New South Wales, Australia Included all records for a 10 year period (1992–2002)	Journal (Addiction)	Compare infant outcomes for women who were retained on MMT throughout their pregnancies, as compared to two other groups: women who began MMT late in pregnancy (<6 months prior to birth), and women who had previous MMT experience (discontinued at least 1 year prior to birth)	*n* = 2,230 (*n* = 1,213 early entry, *n* = 306 late entry, *n* = 711 previous treatment)
Chang et al. (2005)	Brigham and Women's Hospital – Boston, Massachusetts (given to patients initiating prenatal care) 1 of 3 obstetric practices (clinic, faculty, private group)	Journal (Obstetrics and Gynecology)	Assess the effectiveness of a brief intervention (for the pregnant woman and her partner) during pregnancy on abstinence and reduction in consumption until delivery	*n* = 304 Pregnant women and their partners 78% white 8.6% AA 13% other Median age tx group = 32.0 Median age Control group = 30.7
Chang et al. (2006)	Brigham and Women's Hospital – Boston, Massachusetts (given to patients initiating prenatal care) 1 of 3 obstetric practices (clinic, faculty, private group)	Journal (Journal of Studies on Alcohol)	To assess whether drinking-related goals had an impact on consumption	*n* = 115 pairs (enrolled in the above randomized trial) 78% white 8.6% AA 13% other Median age = 32 years

Addressed disparities (yes/no)?	Key Findings related to intervention effectiveness (OR with CI or p values reflecting the intervention-outcome relationship)	Caveats/biases	Findings support the intervention? (Yes/No) for which populations?
No	Women who presented late were more likely to have infants who were <37 weeks gestation, adjusted odds ratio 2.6 (95% CI 1.5–3.3) ($p < 0.001$), and required specialty care in the NICU or SCN, adjusted odds ratio 2.8 (95% CI 1.9–4.1 ($p < 0.001$) Women with previous MMT experience (up to 1 year prior to the birth of their children) had infants who were least likely to experience NAS and least likely to require specialty care after birth.	Linking several large datasets may miss some women or include errors in matching. Datasets did not include information on: income, exposure to violence, polydrug use, methadone dose, or treatment policies, which may influence outcomes	Yes
No	The brief intervention was more effective in women who reported higher rates of drinking at enrollment $p < 0.01$ Partner involvement enhances the effectiveness of the intervention $p < 0.05$	Sample overwhelmingly white, educated and married, with above average incomes (not able to address disparities, but this group has higher risk for drinking during pregnancy) Sample included partners, so findings may be affected by stability and support of a partner which potentially enhances one's ability to change Sample only recruited from those seeking prenatal care May have assembly bias due to those who are motivated being interested in participating Researchers and interventionists not blinded to group assignment	Yes
No	Women who were abstinent at enrollment and chose abstinence as a goal were most likely to remain abstinent during pregnancy (75% of this group) Among women who were not abstinent at enrollment and chose abstinence as a goal, 50% achieved it None of the women who chose "cutting down" achieved abstinence and were less successful in changing drinking behaviors Older subjects were less likely to be abstinent and more likely to choose "cutting down" as their goal ($p = 0.06$)	52% already abstinent – not a "high risk" population Lack of random assignment Self-report data Sample may have had greater partner-related support and home stability, limiting the generalizability of results	Yes Drinking goal selection can be predictive of subsequent behavior, particularly among patients who choose "abstinence" as a goal

(continued)

Table 12.2 (continued)

Authors and year of publication	Study design	Study type and journal	Description of intervention (what, how, where)	Population studied (ages, race ethnicity, sample size)
Goler et al. (2008)	Northern California 21 outpatient Obstetric clinics (managed care) 1999–2003 Retrospective cohort	Journal of Perinatology, 1–7.	Evaluated the effectiveness of Early Start (screening and treatment program integrated into obstetric clinics) – four group model: Group 1 – SAT (screened, assessed and treated) *n* = 2,073 Group 2 – SA (screened and assessed) *n* = 1,203 Group 3 – S (screened) *n* = 156 Group 4 – Controls *n* = 46,553	*n* = 49,985 Group 1 = 31% white, 26% AA, 12% hisp. Group 2 = 36% white, 20% AA, 14% hisp. Group 3 = 23% white, 31% AA 19% hisp. Group 4 = 25% white, 8% AA, 27% hisp.
Kakko et al. (2008)	Population-based observational study, Sweden	Drug and Alcohol Dependence	To assess differences in perinatal outcomes between pregnant patients taking MMT and pregnant patients taking buprenorphine	*n* = 47 (Bup) *n* = 35 (MMT)
Lejeune et al. (2006)	Cohort Study – conducted in 35 public clinics in France	Drug and Alcohol Dependence	To assess differences in perinatal outcomes between pregnant patients taking MMT and pregnant patients taking high-dose buprenorphine	*n* = 35
May et al. (2007)	New Mexico Indian Health Service	Maternal Child Health Journal	Evaluate the effectiveness of enhanced case management in preventing FAS	*n* = 137

Addressed disparities (yes/no)?	Key Findings related to intervention effectiveness (OR with CI or p values reflecting the intervention-outcome relationship)	Caveats/biases	Findings support the intervention? (Yes/No) for which populations?
No	Group 3 had higher rates of the following than Group 1: Neonatal assisted ventilation 2.2 AOR (1.1–4.4) Preterm delivery 2.1 AOR (1.3–3.2) Low birth weight 1.8 AOR (1.1–3.1) Placental abruption 6.8 AOR (3.0–15.5) Intrauterine fetal demise 16.2 AOR (6.0–43.8) Group 2 exhibited intermediate levels in all three areas	No randomization (but propensity score analysis done to eliminate possible confounding with same results) Did not exclude co-morbidities (diabetes, hypertension, etc.)	Yes
No	• MMT recipients had lower birth weight infants than those on bup ($p = 0.008$), but decreased when co-variate analysis indicated this was related to gestational age ($p = 0.07$) • NAS was 2xs higher for those on MMT ($p = 0.0008$), significant after Bonferroni correction (21% of infants on bup needed treatment vs. 68% of those on MMT) • Average length of hospital stay for MMT was longer ($p = 0.0009$)	• Populations were different (bup recipients were younger $p = 0.003$, MMT subjects had longer drug histories), due to lack of random assignment • Observation periods were different (one prospective and one retrospective), making time a potential confounder • Incomplete monitoring of illicit drug use • Both treatment conditions were supported by a comprehensive model, making results less generalizable to programs which do not provide this support	Yes, bup may offer benefits in reducing NAS symptoms and length of hospital stay among infants
No	• No differences were observed between pregnant patients taking MMT and high dose buprenorphine, for perinatal outcomes • Women with poor prenatal care had higher rates of premature delivery than those with good care ($p = 0.229$).	• No random assignment • Included only women in prenatal care • Included only live births, and women who were retained in treatment until delivery, causing potential bias to results	Yes, bup and MMT do not show differences in efficacy with respect to outcomes
NO, but study done with American Indians, with high rates of alcohol consumption	Clients average number of times reported being drunk or high in the past 6 months dropped from 15 to 4.3 at the 6 month follow-up ($p = 0.048$)	Limited sample size as women were divided into four sites who administered intervention with variability No control group No efforts to control confounders	Yes, but on a limited basis

There has also been an examination of the efficacy of methadone maintenance therapy in pregnancy as a function of when a woman enrolls in treatment. In an Australian study, Burns et al. (2006) compared women who were engaged in MMT throughout their pregnancies to a group who began MMT late in pregnancy, and another group who discontinued MMT at least 1 year prior to the birth of their child. Data analysis was retrospective, linking treatment information from three regional systems: the New South Wales (NSW) Addiction System, the NSW Inpatient Statistics Collection and NSW Midwives Data Collection. The study covered the years 1992 through 2002. Women who entered treatment late were more likely to have infants who were <37 weeks gestation and required specialty care in the NICU or SCN. These women were also more likely to be younger, single, belong to indigenous populations, smoke more heavily, and engage in antenatal services late. In addition, those with previous MMT experience (at least 1 year prior to the birth of their children) had infants who were least likely to experience NAS and least likely to require specialty care after birth. Strengths of the study included the ability to link several large regional datasets (which may not be possible in other geographical areas), which resulted in a relatively large sample size, despite the challenges inherent in linking data sets which rely on names and addresses. Limitations of these databases include a lack of information on factors that may influence treatment outcome, including: income, exposure to violence, polysubstance use, methadone dosage, and differences in treatment policies across settings. Nevertheless, it is clear that MMT and buprenorphine have a place in treatment of opioid dependence among pregnant women.

Integrated Programs

Other innovative programs have begun to emerge, which address substance use by pregnant women with an integrated, nonjudgmental approach. The Early Start program was created to address substance use among pregnant women in the Kaiser Permanente clinic network in Northern California (Taillac, Goler, Armstrong, Haley, & Osejo, 2007; see Box 12.3 for description.) To assess the efficacy of the program, a study was conducted with a sample of 49,986 who completed initial screening requirements between 1999 and 2003, and were divided into four groups for comparison: women who were screened, assessed and treated (SAT), women who were screened and assessed (SA), women who were screened only (S) and those with negative screening results who served as the control group. Groups were compared for important differences regarding polysubstance use, smoking status, family history of addiction, and frequency of use and found to be very similar, eliminating the concern of selection bias. Results from the study indicated that women in the screening-only group experienced more negative infant outcomes than those in the SAT group, including a higher risk of intrauterine fetal demise, placental abruption, neonatal assisted ventilation, preterm delivery and low birth weight (Goler, Armstrong, Taillac, & Osejo, 2008). Women in the screening and assessment (SA) group exhibited intermediate levels of all of these outcomes, suggesting that the intervention, rather than simply screening and assessment, is an essential component in changing behavior and improving outcomes. Results remained significant after controlling for ethnicity, amount of prenatal care, and maternal age. Limitations to the study included a failure to control for co-morbidities which may affect infant outcomes such as diabetes, hypertension, psychiatric illness and uterine anomalies. Similarly, the study did not randomize individual women to the four groups; however, a subsequent propensity score analysis resulted in consistent findings.

Enhanced Case Management and Treatment Retention

Some programs use case management (CM) approaches, or CM enhanced with psychosocial services to assist women to enroll in and continue with treatment. Burns et al. (2006) found that case management increases program attendance as well as birth weight. Enhanced case management may

Box 12.3 The Early Start Program

The Early Start Program operates within prenatal care networks, utilizing both survey-based and voluntary urine screening at the first visit to identify women who may need treatment. In addition, all pregnant women are educated regarding risks of substance use. This education occurs multiple times within the pregnancy, and utilizes a variety of approaches, including one-on-one counseling, group education, pamphlets/newsletters, a hotline, and a website. Clinicians in the network receive extensive training regarding drug exposure during pregnancy, with a focus on understanding that addiction is just one of a number of diseases that may affect pregnancy outcomes. This aspect is essential in creating and maintaining a non-judgmental atmosphere toward women who may be using during pregnancy. Clinicians are also trained to make appropriate treatment referrals. A full multi-disciplinary team supports the intervention, including drug treatment professionals, psychiatrists, case managers, social service professionals, and substance abuse treatment specialists. Perhaps the most unique aspect of the Early Start program is having a substance abuse treatment specialist as a part of the prenatal care team who resides in the clinic for ease of referral (Taillac et al., 2007).

For women who are identified through screening as being at risk, ongoing case management and counseling are provided to reduce the probability of negative outcomes. Another unique aspect of the program is the voluntary continuation of this service post-pregnancy, for up to 1 year. This is unique in a service culture that emphasizes pregnancy outcomes and where services are often discontinued after the baby is born. This discontinuation can be premature for women who have come to rely on such support.

provide a more tailored approach to the needs of pregnant women for drug treatment. For example, some studies indicate that having childcare services available at treatment may increase treatment retention (Greenfield et al., 2007).

To study the impact of enhanced case management on drinking behaviors among pregnant American Indians, a cross-sectional, longitudinal study was conducted to assess the potential for case management (in conjunction with a brief intervention based in motivational interviewing techniques) to prevent fetal alcohol spectrum disorders (FASD) (May et al., 2007). The program employed evidence-based recommendations from the literature including: screening all women of childbearing age (without mandated drug testing), focusing on protecting the infant by engaging the mother in prenatal care, and removing barriers to addiction treatment. Other strategies identified in the literature were part of the design as well, such as: a focus on building trust and communication, having coordinated services, communicating with the significant other, removing barriers such as transportation, and providing support to prevention advocates. In each of the four participating communities, activities and services were tailored to specific needs, and four other communities that did not have such services were chosen as comparison groups. Results of the study indicated that women enrolled in case management services experienced a statistically significant reduction in the number of times they reported being drunk or high in the previous 6 months, dropping from 15 to 4.43 at the 6 month follow-up, but statistical significance was not retained at the 12-month follow-up (May et al., 2007).

In the study by May and colleagues, 12.8% of pregnant women were lost to follow-up and 8.1% were still pregnant at the time of the analysis. Among the less favorable birth outcomes were 9.2% who experienced miscarriage or stillbirth, one case of severe fetal alcohol spectrum disorders, one pending confirmation of fetal alcohol syndrome, and one death from sudden infant death syndrome. Treatment retention was poor for 38% of women who withdrew from case management against staff recommendations. The authors cited that the most important finding was that 76% of pregnancy outcomes were

favorable (i.e. experienced normal deliveries), among those pregnancies that could be tracked, supporting case management even if in a limited manner. They also noted the ethical barriers in studying alcohol in pregnancy with an RCT design, noting that cross-sectional designs are most feasible.

Discussion and Recommendations for Future Research

The review of literature on treatment for alcohol and substance use in pregnancy as it relates to infant outcomes and health disparities helped to identify several areas where more research is warranted. For example, this comprehensive search did not identify any studies which examined treatment efficacy as a means for reducing ethnic or racial health disparities. However, it is clear that the preponderance of studies in this area include samples that are predominantly comprised of minority women of disadvantaged economic status who may have less access to services. Thus, those interventions that have identified improvement in outcomes in this population can arguably be applied in other similar settings with low-income minority women with a relative likelihood of producing efficacious results. As evidence-based treatments are expanded, disparities in treatment access may narrow and outcomes may improve. Other disparities exist in treatment-based research, in which racial and ethnic minorities may not be recruited, and women (particularly pregnant women) may be underrepresented, leading to a deficiency in understanding the addiction trajectories and unique treatment needs of minority women (Weiss, Kung, & Pearson, 2003). Clearly, more research is needed in this area.

Randomized controlled trials and experimental approaches are not common in this literature for reasons identified earlier in this chapter; more such studies are certainly warranted to rectify methodological issues in previous research. Those that do exist are often limited by small sample sizes and the inability to control for important confounding variables. Other types of studies can help to fill the gap by using epidemiologic approaches to link large datasets and augment the existing RCT evidence. One such study was included in this review (see Burns et al., 2006), and despite its limitations the results are encouraging. Other limitations to current research include: lack of ability to control for polysubstance use or smoking, sampling issues (small samples, in particular), and lack of randomization and blinding (Winklebaur et al., 2008). Finally, this review was focused in the peer-reviewed literature, and may not represent the full range of research findings. Ashley and colleagues (2003) warn that peer-reviewed articles appear more likely to publish positive results than papers which are published in the grey (or "fugitive") literature.

It is clear from this review and other major reviews discussed above that certain approaches have been shown to reduce negative infant outcomes. The first of these well-researched treatments is methadone maintenance therapy for opioid addicts. Although the literature for pregnant women is plagued by small samples, the results from studies appear to demonstrate benefits of MMT, such as stabilization of opioid dosing to prevent withdrawal and possible miscarriage. This is consistent with current practice and standard of care for opioid-dependent pregnant women. However, more studies are warranted which examine the dosing of MMT for pregnant women and subsequent infant effects, which are not well-represented in the literature. More studies are also needed to determine the full efficacy of buprenorphine as an alternative to MMT, although the studies presented here appear to support the value of this medication for improving outcomes. This research should be prioritized considering the barriers and controls inherent in the MMT treatment system in the United States. For example, take-home dosing of MMT has been shown to increase retention, as has program intensity and stabilization of the mother's dosing schedule (Burns et al., 2006). Unlike MMT, buprenorphine is initially prescribed from a physician's office with take-home dosing rather than in-person daily at a clinic, and thus may increase retention by reducing barriers to receiving the medication. Generally, MMT take-home dosing occurs after a period of time where the patient

has a demonstrated record of program compliance and urine screens which have been negative for illicit opioids. If buprenorphine was determined to be a recommended treatment for pregnant women, this option could be made available in participating obstetric practices and prenatal clinics across the nation, diffusing a potentially efficacious medication to all who may need it, and potentially reducing the stigma associated with addiction treatment. There is a need for controlled studies on the use of buprenorphine in pregnancy (Rayburn & Bogenschutz, 2004), as well as for phase four studies with pregnant women in pharmacological interventions. The double-blind MOTHER trial compares methadone to buprenorphine therapy. Results are not yet released but are eagerly anticipated, and depending on the findings the study may alter current practice approaches to opioid substitution during pregnancy.

It is also clear from the current literature (included in the reviews listed above) that contingency management approaches are an important mechanism for retaining pregnant women in treatment. Because treatment retention is related to improved outcomes for both mother and baby, incentives to remain in treatment are a promising strategy to improve perinatal outcomes. CM can be used alone or in conjunction with other treatment methods. CM approaches have been shown to increase abstinence from cocaine during pregnancy (Elk et al., 1995; Seracini, Nunes, Tross, & Spano, 1997). When using CM, the amount or type of incentive may make a difference; for example, one study that found better attendance for women receiving payments of $5.00–$10.00 in comparison to those who received no payment or a payment of $1.00 per session (Svikis, Lee, Haug, & Stitzer, 1997). However, even contingency management studies are subject to the limitations of very small samples. An RCT in a prenatal care setting that compared a program using contingency management and drug counseling with one behavioral counseling alone had a sample of only 12 (Elk, Magnus, Rhodes, Andres, & Grabowski, 1998); another study, which compared perinatal outcomes for women enrolled in MMT plus counseling to women receiving MMT plus counseling and contingency management, had a sample of 14, 78% of whom were non-minority (Carroll, Chang, Behr, Clinton, & Kosten, 2005) (Table 12.3). Despite the small sample, findings indicated that those

Table 12.3 Summary of scoring: quality checklists for RCTs/observational studies

Author, year	Reporting	External validity	Internal validity (bias)	Internal validity (confounding)	Power	Total quality score <14 = poor 15–19 = fair >20 = good	Suitability of study to assess effectiveness: greatest, moderate, least
Burns et al. (2006)	11	2	6	3	0	22 (good)	Greatest
Chang et al. (2005)	11	2	5	5	2	25 (good)	Greatest
Chang et al. (2006)	9	2	4	2	0	17 (fair)	Moderate
Goler et al. (2008)	10	2	6	3	0	21 (good)	Greatest
Kakko et al. (2008)	12	2	5	0	0	19 (fair)	Moderate
Lejeune et al. (2006)	9	1	5	1	0	16 (fair)	Moderate
May et al. (2007)	9	2	2	1	0	14 (poor)	Least

in the contingency management group had infants with healthier birth weights, suggesting CM's promise as an added component of treatment. An extension of the contingency management approach is a "therapeutic workplace" model, in which treatment participants are provided incentives not just for abstinence, but for workplace performance and attendance at a job. This approach has been found to be effective in a small study of pregnant and post-partum women, who were chronically unemployed/underemployed, and dependent on opiates and/or cocaine (Silverman, Svikis, Robles, Stitzer, & Bigelow, 2001). Such programs may provide opportunities for both economic independence and professional development in a highly disenfranchised population, and are stimulating pathways for further research.

Other areas of research which require more methodologically rigorous study include women-centered treatment approaches, which may be more adept at addressing the plethora of complications faced by women addicted to substances. Females have higher rates of mood disorders, PTSD, anxiety disorders, eating disorders, trauma histories and lack of family/partner support, all of which interfere with accessing treatment (Greenfield, et al., 2007). In addition, 61–75% of women in substance abuse treatment programs reported a history of sexual abuse (Howell et al., 1999), prompting a need to consider trauma-informed and gender-specific programming and services. It may be that for substance-abuse treatment for women (including pregnant women) to be effective, programs need to be tailored to gender-specific counseling regarding sexual abuse. Such programs could then be evaluated to see if they produce greater improvements in maternal and infant outcomes than programs that are not tailored to women's needs.

Females also progress more quickly from their first use of a given drug to regular use, and then to entry into treatment than do males. Severity of addiction symptoms among women may be equal to that of males, despite the average differences in number of years of use or quantity of drug used. Females may be less likely to seek or enter treatment due to a host of barriers, especially pregnancy, compounded by the lack of services available for pregnant women and the fear of child protective service involvement. Some studies also show less retention in treatment for females in comparison to males; however, when women *are* retained in treatment, some studies show that treatment outcomes may be better for females than males (Greenfield et al., 2007). Thus, more study is needed to determine which approaches best assist with retention in treatment, as well as how health, gender and racial disparities may affect retention.

Many women-centered or women-only programs have been designed to address known barriers such as inadequate child care and transportation, in order to enhance retention during the treatment process. There is some evidence that women-only programs are not only more likely to address these barriers, but that they are also more likely to provide services to address a wide range of psychosocial needs of women, including priority spots for pregnant users, prenatal and well-baby care, job and life skills training, advocacy, housing and transportation assistance (Greenfield et al. 2007). There is also a suggestion that these programs may better address disparities by offering services specific to groups in most need, such as Latina women. Because the programs providing more comprehensive services, they also tend to treat women with more intensive needs. However, studies that examine women-centered treatment have had serious methodological limitations, such as the lack of control groups. For example, one study reported on a unique women-centered treatment program in Harlem, NY provided tailored, comprehensive services to substance using women (see Box 12.4 for program description; McMurtrie, Rosenberg, Kerker, Kan, & Graham, 1999). The researchers attempted to evaluate program effectiveness by utilizing comparison groups developed retrospectively, although information about matching such groups was rather limited. They concluded that the PACE program facilitated better perinatal outcomes, including reduced low birth weight and intrauterine growth retardation and greater mean birth weight than all comparison groups. However, results must be interpreted cautiously given the lack of adequately matched comparison groups.

Box 12.4 The PACE Program

Services included: on-site prenatal, postpartum and pediatric care, nutritional assessments, on-site WIC enrollment and referral to social services, group and individual level counseling, parenting education and professional development (GED courses and vocational courses on-site). These services were provided in an atmosphere that recognizes relapse as a normal part of addiction, in direct opposition to most abstinence-based programs. For example, rather than using positive urine screens to deny services, the program counseled women with positive urine screens and encouraged them to stay in the program. This is an important step in addiction treatment, in direct contrast to how positive urine screens are generally treated. Many times, individuals are expelled from treatment when a urine screen tests positive, basically for exhibiting symptoms of the disease at hand. A critical, but unanswered question is whether this type of comprehensive program is more cost effective than the usual cycle of incarceration, foster care and urgent medical care that women with addictions suffer. In fact, rather than build this treatment program on a "jail" concept, whereby women must be isolated from the outside world in order to be treated, the PACE program allowed maximum flexibility, resulting in reduced relapse and better retention. In addition, this program sought to allow women time to adjust into treatment, a full 6 weeks to stabilize one's life and build trust with the multidisciplinary staff (physician, nurse, substance abuse counselors, child care providers, etc.).

Another methodological issue among studies examining women-only treatment programs is the lack of examination of infant outcomes. For example, Hser and Niv (2006) conducted a large-scale study comparing the treatment needs and services provided to pregnant women in women-only versus mixed-gender treatment programs based in California. They found that pregnant women who were enrolled in women-only treatment programs had greater addiction severity than those in mixed gender programs. They were also two times as likely to be homeless, significantly less likely to be employed, had a history of more drug treatment episodes, were more likely to have psychiatric problems, more likely to have been arrested and had greater legal problems. Despite these issues, women-only treatment programs were less likely to offer psychiatric services or medications, legal services, or assistance with employment than mixed-gender programs, a problematic finding in a population in such need of these services. However, women-only treatment programs were more likely to offer childcare, mental health services for children, HIV testing, pregnancy testing and consultation, services needed by the pregnant, addicted population they serve.

Such programs, although resource-heavy and understudied, are an exciting future prospect for researchers in the field, although it is clear that more improvements can be made to increase the woman-centered offerings of programs, particularly given that fewer than half of existing substance abuse treatment programs in the U.S. offer services targeted specifically to their female clients (SAMHSA, 2006). Additional research is warranted to determine whether these additional services result in healthier women and infants, and whether they have the potential to reduce disparities, about which virtually nothing exists in the literature. Women-only or women-centered treatment programs, which reduce barriers to enrollment, remain an exciting opportunity for outcomes-based research.

Given the potential costs associated with more comprehensive and gender-focused programs, it is easy to see why such programs are scarce. However, decisions about investing in programs should be based on their impact on women and infants, their ability to be cost-effective, and their potential to reduce disparities. In addition, "Program rules and policies should be patient-centered, empowering (not punitive), and reflect a barrier-reduction approach to treatment." (Jones et al., 2009). If such

programs have a greater potential to have an impact on parameters beyond the biomedical, they may well ameliorate many of the environmental concerns which precipitate harm to both mothers and infants. Such efforts would require a less punitive focus on mothers using substances like crack cocaine and heroin, to increasing incentives to participate in prenatal care and treatment and reducing the devastation such substances can have on women's lives. Similarly, biomedical concerns related to the use of dangerous but legal substances like alcohol require an expedient approach for early identification and treatment with evidence-based programming. Whether licit or illicit, prevailing attitudes of stigma compounded by racial and ethnic stereotypes toward pregnant women who use substances remain a significant barrier to overcome in both research and practice.

References

Abel E. L., & Krueger M. (2002). Physician attitudes concerning legal coercion of pregnant alcohol and drug abusers. *American Journal of Obstetrics and Gynecology*, *186*, 768–772.

American Academy of Pediatrics (1998). Neonatal drug withdrawal. Committee on drugs. *Pediatrics*,*101*(6), 1079–1088.

Ananth, C. V., Berkowitz, G. S., Savitz, D. A., & Lapinski, R. H. (1999). Placental abruption and adverse perinatal outcomes. *Journal of the American Medical Association*, *282*(17), 1646–1651.

Annas, G. J. (2001). Testing poor pregnant women for cocaine-physicians as police investitgators. *New England Journal of Medicine*, *344*(22), 1729–1732.

Araojo, R., McCune, S., & Feibus, K. (2008). Substance abuse in pregnant women: Making improved detection a good clinical outcome. *Clinical Pharmacology and Therapeutics*, *83*(4), 520–522.

Ashley, O., Marsden, M., & Brady, T. (2003). Effectiveness of substance abuse treatment programming for women: A review. *The American Journal of Drug and Alcohol Abuse*, *29*(1), 19–53.

Bauer, C. R., Shankaran, S., Bada, H. S., Lester, B., Wright, L. L., Krause-Steinrauf, H., et al (2002). The maternal lifestyle study: Drug exposure during pregnancy and short-term maternal outcomes. *American Journal of Obstetrics and Gynecology*, *186*, 487–495.

Bearer, C., Santiago, L., O'Riordan, M., Buck, K., Lee, S., & Singer, L. (2005). Fatty acid ethyl esters: Quantitative biomarkers for maternal alcohol consumption. *The Journal of Pediatrics*, *146*(6), 824–830.

Berenson, A. B., Stiglich, N. J., Wilkinson, G. S., & Anderson, G. D. (1991). Drug abuse and other risk factors for physical abuse among white non-Hispanic, black and Hispanic women. *American Journal of Obstetrics and Gynecology*, *164*(6 Pt 1), 1491–1499.

Breitbart, V., Chavkin, W., & Wise, P. (1994). The accessibility of drug treatment for pregnant women: A survey of programs in five cities. *American Journal of Public Health*, *84*(10), 1658–1661.

Buchi, K. F., & Varner, M. W. (1994). Prenatal substance use in a western urban community. *Western Journal of Medicine*, *161*(5), 483–486.

Burns, L., Mattick, R., Lim, K., & Wallce, C. (2006). Methadone in pregnancy: Treatment retention and neonatal outcomes. *Addiction*, *102*, 264–270.

Carroll, K., Chang, G., Behr, H., Clinton, B., & Kosten, T. (2005). Improving treatment outcome in pregnant, methadone-maintained women. *American Journal on Addictions*, *4*(1), 56–59.

Centers for Disease Control and Prevention (CDC) (1995). Update: Trends in fetal alcohol syndrome – United States, 1979–1993. *Morbidity and Mortality Weekly Report*, *44*(13), 249–251.

Chang, J., Dado, D., Frankel, R., Rodriguez, K., Zickmund, S., Ling, B., et al. (2008). When pregnant patients disclose substance use: Missed opportunities for behavioral change counseling. *Patient Education and Counseling*, *72*, 394–401.

Chang, G., Goetz, M. A., Wilkins-Haug, L., & Berman, S. (2000). A brief intervention for prenatal alcohol use: An in-depth look. *Journal of Substance Abuse Treatment*, *18*(4), 365–369.

Chang, G., McNamara, T., Orav, J., Koby, D., Lavigne, A., Ludman, B., et al. (2005). Brief intervention for prenatal alcohol use: A randomized trial. *Obstetrics & Gynecology*, *105*(5), 991–998.

Chang, G., McNamara, T., Oray, E., & Wilkins-Haug, L. (2006). Brief intervention for prenatal alcohol use: The role of drinking goal selection. *Journal of Substance Abuse Treatment*, *31*, 419–424.

Chang, G., Wilkins-Haug, L., Berman, S., Goetz, M. A., Behr, H., & Hiley, A. (1998). Alcohol use and pregnancy: Improving identification. *Obstetrics & Gynecology*, *91*(6), 892–898.

Chasnoff, I., Landress, H., & Barrett, M. (1990). The prevalence of illicit-drug or alcohol use during pregnancy and discrepancies in mandatory reporting in Pinellas County, Florida. *New England Journal of Medicine*, *322*, 1202–1206.

Chavkin, W. (1991) Mandatory treatment for drug use during pregnancy. *Journal of the American Medical Association 226*(11), 1556–1561.

Chazotte, C., Youchah, J., & Comerford Freda, M. (1995). Cocaine use during pregnancy and low birth weight: The impact of prenatal care and drug treatment. *Seminars in Perinatology*, *19*(4), 293–300.

Christensen, C. (2008). Management of chemical dependence in pregnancy. *Clinical Obstetrics and Gynecology*, *51*(2), 445–455.

Corse, S., & Smith, M. (1998). Reducing substance abuse during pregnancy: Discriminating among levels of response in a prenatal setting. *Journal of Substance Abuse Treatment*, *15*(5), 457–467.

Denny, C. H., Tsai, T., Floyd, R. L., & Green, P. P. (2009). Alcohol use among pregnant and non-pregnant women of childbearing age-United States. *Morbidity and Mortality Weekly Review*, *58*(19), 529–532(4).

Elk, R., Magnus, L., Rhodes, H., Andres, R., & Grabowski, J. (1998). Cessation of cocaine use during pregnancy: effects of contingency management interventions on maintaining abstinence and complying with prenatal care. *Addictive Behaviors*, *23*(1), 57–64.

Elk, R., Schmitz, J., Spiga, R., Rhoades, H., Andres, R., & Grabowski, J. (1995). Behavioral treatment of cocaine-dependent pregnant women and TB-exposed patients. *Addictive Behaviors*, *20*, 533–542.

Goler, N., Armstrong, M., Taillac, C., & Osejo, V. (2008). Substance abuse treatment linked with prenatal visits improves perinatal outcomes: a new standard. *Journal of Perinatology*, *28*, 597–603.

Greenfield, S., Brooks, A., Gordon, S., Green, C., Kropp, F., McHugh, K., et al. (2007). Substance abuse treatment entry, retention, and outcome in women: A review of the literature. *Drug and Alcohol Dependence*, *86*, 1–21.

Haller, D., Miles, D., & Dawson, K. (2003). Factors influencing treatment enrollment by pregnant substance abusers. *The American Journal of Drug and Alcohol Abuse*, *29*(1), 117–131.

Handler, A. S., Mason, E. D., Rosenberg, D. L., & Davis, F. G. (1993). The relationship between exposure during pregnancy to cigarette smoking and cocaine use and placenta previa. *American Journal of Obstetrics and Gynecology*, *170*(3), 884–889.

Handmaker, N., & Wilbourne, P. (2001). Motivational interventions in prenatal clinics. *Journal of the National Institute on Alcohol Abuse and Alcoholism, Alcohol Research and Health*, *25*(3), 219–299.

Hans, S. L. (1996). Prenatal drug exposure: Behavioral functioning in late childhood and adolescence. *Nida Research Monograph*, *164*, 261–276.

Hollinshead, W. H., Griffin, J. F., Scott, H. D., Burke, M. E., Coustan, D. R., & Vest, T. A. (1990). Statewide prevalence of illicit drug use by pregnant women – Rhode Island. *Morbidity and Mortality Weekly report*, *39*(14), 225–227.

Howell, E., Heiser, N., & Harrington, M. (1999). A review of recent findings on substance abuse treatment for pregnant women. *Journal of Substance Abuse Treatment*, *16*(3), 195–219.

Hser, Y., & Niv, N. (2006). Pregnant women in women-only and mixed-gender substance abuse treatment programs: A comparison of client characteristics and program services. *Journal of Behavioral Health Sciences & Research*, *33*(4), 431–442.

Hulse, G. K., English, D. R., Milne, E., Holman, C. D., & Bower, C. I. (1997). Maternal cocaine use and low birth weight newborns: A meta-analysis. *Addiction*, *92*(11), 1561–1570.

Jones, H., Haug, N., Stitzer, M., & Svikis, D (2000). Improving treatment outcomes for pregnant drug-dependent women using low magnitude voucher incentives. *Addictive Behaviors*, *25*(2), 263–267.

Jones, H., Kaltenbach, K., Coyle, M., Heil, S., O'Grady, K., Arria, A., et al. (2009). Rebuilding lives: Helping women recover from opioid addiction during pregnancy. *Counselor*, *10*(5), 10–19.

Jones, K. L., & Smith D. W. (1973). Recognition of the fetal alcohol syndrome in early infancy. *Lancet*,*2*, 999–1001.

Kakko, J., Heilig, M., & Sarman, I. (2008) Buprenorphine and methadone treatment of opiate dependence during pregnancy: Comparison of fetal growth and neonatal outcomes in two consecutive case series. *Drug and Alcohol Dependence*, *96*(1–2), 69–78.

Kerker, B. D., Leventhal, J. M., Schlesinger, M., & Horwitz, S. M. (2006). Racial and ethnic disparities in medical history taking: Detecting substance use among low-income pregnant women. *Ethnicity and Disease*, *16*, 28–34.

Knight, D., Logan, S., & Simpson, D. (2001). Predictors of program completion for women in residential substance abuse treatment. *American Journal of Drug and Alcohol Abuse*, *27*(1), 1–18.

Lejeune, C., Simmat-Durand, L., Gourarier, L., Aubisson, S., & Groupe d'Etudes Grossesse et Addictions (GEGA) (2006). Prospective multicenter observational study of 260 infants born to 259 opiate-dependent mothers on methadone or high-dose burprenorphine substitution. *Drug and Alcohol Dependence*, *82*(3), 250–257.

Lewis, L. (2004). Culturally appropriate substance abuse treatment for parenting African American women. *Issues in Mental Health Nursing*, *25*(451), 451–472.

Lui, S., Terplan, M., & Smith, E. (2008). Psychosocial interventions for women enrolled in alcohol treatment during pregnancy. *The Cochrane Library, 2.* Art. No.: CD006753. DOI: 10.1002/14651858.CD006753.pub2.

Macones, G. A., Sehdev, H. M., Parry, S., Morgan, M. A., & Berlin, J. A. (1997). The association between maternal cocaine use and placenta previa. *American Journal of Obstetrics and Gynecology*, *177*(5), 1097–1100.

March of Dimes Birth Defects Foundation. (1995). Substance Abuse Curriculum for Obstetricians and Gynecologists. In: R. K. Rosen, C. E. Dube, D. C. Lewis, A. C. Tsamas, J. Markowitz (Eds.). White Plains, New York: March of Dimes.

May, P., Miller, J., Goodhart, K., Maestas, O., Buckey, D., Trujillo, P., et al. (2007). Enhanced case management to prevent fetal alcohol spectrum disorders in northern plains communities. *Maternal and Child Health Journal*, *12*(6), 747–759.

McMurtrie, C., Rosenberg, K., Kerker, B., Kan, J., & Graham, E. (1999). A unique drug treatment program for pregnant and postpartum substance-using women in New York City: Results of a pilot project, 1990–1995. *American Journal of Drug and Alcohol Abuse*, *25*(4), 701–713.

Messinger, D. S., Bauer, C. R., Das, A., Seifer, R., Lester, B. M., Lagasse, L. L., et al. (2004). The maternal lifestyle study: Cognitive, motor, and behavioral outcomes of cocaine-exposed and opiate-exposed infants through three years of age. *Pediatrics*, *113*(6), 1677–1685.

Minozzi, S., Amato, L., Vecci, S., & Davoli, M. (2008). Maintenance agonist treatments for opiate dependent pregnant women. *The Cochrane Library, 2.* Art. No.: CD006318. DOI: 10.1002/14651858.CD006318.pub2.

Nalty, D. (1991). *South carolina prevalence study of drug use among women giving birth*. Columbia, SC: South Carolina Commission on Alcohol and Drug Abuse.

National Institute on Drug Abuse (2006). National Pregnancy and Health Survey: Drug Use Among Women Delivering Live Births, 1992. United States Department of Health and Human Services. National Institutes of Health. National Institute on Drug Abuse. National Pregnancy and Health Survey: Drug Use Among Women Delivering Live Births, 1992 [Computer file]. ICPSR02835-v2. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2008-07-31. doi:10.3886/ICPSR02835 http://dx.doi.org/10.3886/ICPSR02835

National Institute on Drug Abuse (2009). NIDA info facts: treatment approaches for drug addiction. Accessed October 2009 from http://www.drugabuse.gov/PDF/InfoFacts/IF_Treatment_Approaches_2009_to_NIDA_92209.pdf

Niv, N., & Hser, Y. (2007). Women-only and mixed-gender drug abuse treatment programs: Service needs, utilization and outcomes. *Drug and Alcohol Dependence*, *87*(2–3), 194–201.

O'Grady, M. J., Hopewell, J., & White, M. J. (2009). Management of neonatal abstinence syndrome: A national survey and review of practice. *Archives of Disease in Childhood*, *94*, F249–F252.

Ostrea, E. M., Brady, M., Gause S., Raymundo, A., & Stevens, M. (1992). Drug screening of newborns by meconium analysis: A large scale, prospective, epidemiologic study. *Pediatrics 89*(1), 107–113.

Phares, T. M., Morrow, B., Lansky, A., Barfield, W., Prince, C., Marchi, K., et al. (2004). Surveillance for disparities in maternal health-related behaviors-selected states, Pregnancy Risk Assessment Monitoring System (PRAMS), 2000–2001. *Morbidity and Mortality Weekly Report*, *53*(SS04), 1–13.

Project CHOICES Research Group (2002). Alcohol-exposed pregnancy: Characteristics associated with risk. *American Journal of Preventive Medicine*, *23*(3), 166–173.

Rayburn, W., & Bogenschutz, M. (2004). Pharmacotherapy for pregnant women with addictions. *American Journal of Obstetrics and Gynecology*, *191*(6), 1885–1897.

Russell, M., Martier, S. S., Sokol, R. J., Mudar, P., Jacobson, S., & Jacobson, J. (1996). Detecting risk drinking during pregnancy: A comparison of four screening questionnaires. *American Journal of Public Health*,*86*(10), 1435–1439.

Saiki, T., Lee, S., Hannam, S., & Greenough, A. (2010). Neonatal abstinence syndrome – postnatal ward versus neonatal unit management. *European Journal of Pediatrics*, *169*(1), 95–98.

Schempf, A. (2007). Illicit drug use and neonatal outcomes: A critical review. *Obstetrical and Gynecological Survey*, *62*(11), 749–757.

Schmidt, L., Greenfield, T., & Mulia, N. (2006). Racial and ethnic disparities in alcoholism treatment services. *Alcohol Treatment & Health*, *29*(1), 49–54.

Seracini, A., Nunes, E., Tross, S., & Spano, C. (1997). Achieving cocaine abstinence in perinatal cocaine dependent women: A contingency management approach. In L. S. Harris (Ed.), *Problems of Drug Dependence 1996: Proceedings of the 58th annual meeting* (NIH 97–4236, p. 261). Washington, DC: U.S. Government Printing Office.

Sidhu, J. S., & Floyd, R. L. (2002). Alcohol use among women of childbrearing age – United States, 1991–1999. *Morbidity and Mortality Weekly Report*, *51*(13), 263–276.

Silverman, K., Svikis, D., Robles, E., Stitzer, M., & Bigelow, G. (2001). A reinforcement-based therapeutic workplace for the treatment of drug abuse: Six-month abstinence outcomes. *Experimental and Clinical Psychopharmacology*, *9*(1), 14–23.

Sokol, R. J., Delaney-Black, V., & Nordstrom, B. (2003). Fetal alcohol spectrum disorder. *Journal of the American Medical Association*, *290*(22), 2996–2999.

Sokol, R. J., Martier, S. S., & Ager, J. W. (1989). The T-ACE questions: Practical prenatal detection of risk-drinking. *American Journal of Obstetrics and Gynecology*, *160*(4), 863–868.

Stratton, K., Howe, C., & Battaglia, F. (1996). *Fetal alcohol syndrome: Diagnosis, epidemiology, prevention and treatment* (pp. 82–99). Institute of Medicine. Washington DC: National Academies Press.

Substance Abuse and Mental Health Services Administration (1995). *TIP 2: Pregnant, substance using women*. Treatment Improvement Protocol (TIP) Series 2. Center for Substance Abuse Treatment, DHHS Publication No. (SMA) 95-3056, Rockville, MD. http://www.ncbi.nlm.nih.gov/bookshelf/br.fcgi?book=hssamhsatip&part=A22442

Substance Abuse and Mental Health Services Administration. (2005). *Results from the 2004 National Survey on Drug Use and Health*. Office of Applied Studies, NSUHD Series H-28, DHHS Publication No. SMA 05–4062. Rockville, MD. http://www.drugabusestatistics.samhsa.gov/nsduh/reports.htm#2k4

Substance Abuse and Mental Health Services Administration. (2006). *The Dasis Report: Facilities Offering Special Programs or Groups for Women: 2005*(35). http://www.oas.samhsa.gov/2k6/womenTx/womenTX.htm

Substance Abuse and Mental Health Services Administration. (2007). *Results from the 2006 National Survey on Drug Use and Health: National findings.* Office of Applied Studies, NSDUH –Series H-32, DHHS Publication No. SMA 07-4293. Rockville MD http://www.drugabusestatistics.samhsa.gov/nsduh/reports.htm#2k6

Substance Abuse and Mental Health Services Administration. (2008). *Results from the 2007 National Survey on Drug Use and Health*. Office of Applied Studies, NSDUH Series H-34 DHHS Publication No. SMA 08-4343, Rockville, MD. http://www.drugabusestatistics.samhsa.gov/nsduh/reports.htm#2k7

Substance Abuse and Mental Health Services Administration, Office of Applied Studies. (2009). *The NSDUH report: Substance Use among Women during Pregnancy and Following Childbirth*. Rockville MD. http://www.oas.samhsa.gov/2k9/135/PregWoSubUse.htm

Svikis, D., Golden, A. S., Huggins, G. R., Pickens, R. W., McCaul, M. E., Velez, M. L., et al. (1997). Cost-effectiveness of treatment for drug-abusing pregnant women. *Drug and Alcohol Dependence*, *45*, 105–113.

Svikis, S., Lee, J., Haug, N., & Stitzer, M. (1997). Attendance incentives for outpatient treatment: Effects in methadone and nonmethadone-maintained pregnant drug dependent women. *Drug and Alcohol Dependence*, *48*, 33–41.

Taillac, C., Goler, N., Armstrong, M., Haley, K., & Osejo, V. (2007). Early start: An integrated model of substance abuse intervention for pregnant women. *The Permanente Journal*, *11*(3), 5–11.

Tardiff, K., Marzuk, P. M., Leon, A. C., Hirsch, C. S., Stajic, M., & Portera, L. (1994). Homicide in New York City: Cocaine use and firearms. *Journal of the American Medical Association*, *272*(1), 43–46.

Terplan, M., & Lui, S. (2007). Psychosocial interventions for pregnant women in outpatient illicit drug treatment programs compared to other interventions. *Cochrane Database of Systemic Reviews, 4*, Art. No.: CD006037. DOI: 10.1002/14651858.CD006037.pub2.

Thompson, M. P., & Kingree, J. B. (1998). The frequency and impact of violent trauma among pregnant substance abusers. *Addictive Behaviors*, *23*(2), 257–262.

Thompson, B. L., Levitt, P., & Stanwood, G. D. (2009). Prenatal Expposure to drugs; effects on brain development and implications for policy and education. *Nature Reviews/Neuroscience*, *10*, 303–312.

United States Preventive Services Task Force. (2004). Revised USPSTF Screening and Behavioral Counseling Interventions in Primary Care to Reduce Alcohol Misuse: Recommendation Statement. *Annals of Internal Medicine*, *140*, 554–556.

United States Preventive Services Task Force. (2008). *Screening for illicit drug use*. January 2008; Agency for Healthcare Research and Quality, Rockville MD. http://ahrq.gov/clinic/uspstf/supsdrug.htm Counseling Interventions in Primary Care to Reduce Alcohol Misuse: Recommendation Statement (2004) *Annals of Internal Medicine*, *140*(7), 554–556.

Uziel-Miller, N., & Lyons, J. (2000). Specialized substance abuse treatment for women and their children: An analysis of program design. *Journal of Substance Abuse Treatment*, *19*, 355–367.

Vega, W. F., Kolody, B., Hwang, J., & Noble, A. (1993). Prevalence and magnitude of perinatal substance exposures in California. *New England Journal of Medicine*, *329*(12), 850–854.

Weber, M. K., Floyd, R. L., Riley, E. P., & Snider, D. E. (2002). National task force on fetal alcohol syndrome and effect: Defining the national agenda for fetal alcohol syndrome and other prenatal alcohol-related effects. *Morbidity and Mortality Weekly Report*, *51*(RR14), 9–12.

Weiss, S., Kung, H.-C., & Pearson, J. (2003). Emerging issues in gender and ethnic differences in substance abuse and treatment. *Current Women's Health Reports*, *3*, 245–253.

Winklebaur, B., Kopf, N., Ebner, N., Jung, E., Thau, T., & Fischer, G. (2008). Treating pregnant women dependent on opioids is not the same as treating pregnancy and opioid dependence: A knowledge synthesis for better treatment for women and neonates. *Addiction*, *103*, 1429–1440.

Wolfe, E. L., Davis, T., Guydish, J., & Delucchi, K. L. (2005). Mortality risk associated with perinatal drug and alcohol use in California. *Journal of Perinatology*, *25*, 93–100.

Wolfe, E., Guydish, J., Santos, A., Delucchi, K., & Gleghorn, A. (2007). Drug treatment utilization before, during and after pregnancy. *Journal of Substance Use*, *12*(1), 27–38.

Chapter 13
The Evidence for Perinatal Depression Screening and Treatment

Laura J. Miller, Renu Gupta, and Andrea M. Scremin

Depression is among the most prevalent and high-risk perinatal health problems. From studies using objective measures of depression, it is estimated that in the United States, the period prevalence of major depression during pregnancy ranges from 9.4 to 12.7%, and that 21.9% of women have a major depressive episode during the first year postpartum (Gaynes et al., 2005). Peripartum depression may differentially affect women in various cultural, ethnic, and/or socioeconomic groups. A review of 143 studies of postpartum depression in 40 counties (Halbreich & Karkun, 2006) showed a wide range of prevalence rates in different countries, from nearly 0 to nearly 60%. Within the United States, most studies directly comparing rates of perinatal depression in different racial/ethnic groups have found increased prevalence in African-American and Hispanic women as compared to white women (Howell, Mora, Horowitz, & Leventhal, 2005; Orr, Blazer, & James, 2006). Factors posited to account for these differences include variations in the types and severity of risk factors (e.g., poverty, stress), protective factors (e.g., social support, enriched nutrition), ways of communicating symptoms or distress, conceptualizations of mental illness, levels of stigma, and biological vulnerabilities. When only studies employing structured clinical interviews (as opposed to self-report) are analyzed, the prevalence of major depression is comparable among difference socioeconomic groups, but the prevalence of minor (subsyndromal) depression is more prevalent among women with lower socioeconomic status (Gavin et al., 2005).

Untreated perinatal depression poses risks to mothers and their offspring (Bonari et al., 2004). An estimated 5–14% of women have thoughts of self-harm during pregnancy or postpartum, with suicide accounting for up to 20% of postpartum maternal deaths (Lindahl, Pearson, & Colpe, 2005). Symptoms of depression during pregnancy have been associated with increased risk of low birth weight (Hoffman & Hatch, 2000; Federenko & Wadhwa, 2004; Field, Diego, & Hernandez-Reif, 2006a), preterm delivery (Dayan et al., 2002; Orr, James, & Prince, 2002; Dayan et al., 2006; Field et al., 2006a; Alder, Fink, Bitzer, Hosli, & Holzgreve 2007), and neonatal irritability (Zuckerman, Bauchner, Parker, & Cabral, 1990). Less well established are possible links between untreated antenatal depression and spontaneous abortion (Arck et al., 2001), bleeding (Preti et al., 2000), increased uterine artery resistance (Teixeira, Fisk, & Glover, 1999), low Apgar scores (Zax, Sameroff, & Babigian, 1977), maternal hypertension (Paarlberg, Vingerhoets, Passchier, Dekkar, & Van Geijn, 1995), preeclampsia (Kurki, Hiilesmaa, Raitasalo, Mattila, & Ylikorkala, 2000), poor neonatal adaptation (Misri et al., 2004), and admission to neonatal intensive care units (Chung, Lau, Yip, Chiu, & Lee, 2001). Several studies also suggest long-term effects on offspring of their mothers' antenatal depression, including poorer growth, increased risk of infection, and more difficult temperaments (Huot, Brennan, Stowe, Plotsky, & Walker, 2004; Rahman, Iqbal, Bunn, Lovel, & Harrington, 2004). Untreated postpartum depression

L.J. Miller (✉)
Women's Mental Health Division, Department of Psychiatry,
Brigham and Women's Hospital, Harvard University, 75 Francis St. Boston, MA 02115, USA
e-mail: lmiller23@partners.org

A. Handler et al. (eds.), *Reducing Racial/Ethnic Disparities in Reproductive and Perinatal Outcomes: The Evidence from Population-Based Interventions*, DOI 10.1007/978-1-4419-1499-6_13,

can adversely affect the quality of the mother-child interaction, leading to more negative dyadic interactions, reduced likelihood of breast-feeding, less healthy feeding and sleeping practices, and fewer positive enrichment activities such as reading, singing, and storytelling (Cohn, Campbell, Matias, & Hopkins, 1990; Hatton et al., 2005; Paulson, Dauber, & Leiferman, 2006; McLearn, Minkovitz, Strobino, Marks, & Hou, 2006). This adversely affects the subsequent emotional, behavioral and cognitive development of the child (Murray, Cooper, & Hipwell, 2003).

Perinatal depressive symptoms seem to be a more significant risk factor for poor pregnancy outcomes in women already at higher risk. For example, in one study, there was no significant association between depression scores and fetal growth in the overall study sample, but among study participants from households with low occupational status, higher depression scores correlated significantly with lower birth weights (Hoffman & Hatch, 2000). Similarly, another study found a significant association between depression scores and preterm birth in African-American, but not white, study participants (Orr & Miller, 1995).

Given the high prevalence and associated risks, a key public health question is whether or not universal screening for depressive symptoms in pregnant and postpartum women will make a positive difference in reducing disparities in adverse pregnancy outcomes. This chapter will review validation data for available screening instruments and discuss their usability in different contexts. Since the utility of screening is linked with the availability of effective interventions, the chapter will also review the emerging evidence base for perinatal depression treatments.

Screening for Perinatal Depression

Overview: Should We Routinely Screen for Perinatal Depression?

When considering whether or not it makes sense to screen for a particular symptom cluster or disorder, criteria established by the United States Preventive Services Task Force (USPSTF) (Pignone, Gaynes, & Rushton, 2002) and the United Kingdom National Screening Committee (Buist et al., 2002) provide useful guidelines:

- *Guideline #1*: The condition should be an important health issue, based on prevalence and associated risks if untreated. In the case of perinatal depression, this includes risks to offspring (McLennan & Offord, 2002).
- *Guideline #2*: The screening tool should be valid, acceptable to users and cost-effective.
- *Guideline #3*: Treatment should be effective and available.

Regarding Guideline #1, prevalence data support universal screening for perinatal depression. Prenatal care includes routine screening for conditions with lower prevalence rates, such as gestational diabetes (4.8% of pregnant women) (Ferrara, Hedderson, Quesenberry, & Selby, 2002) and hypertension (5.0% of pregnant women) (Haddad & Sibai, 1999). It is also clear that when untreated, perinatal depression is associated with serious maternal morbidity and mortality, the latter due mostly to suicide. There is more controversy about whether perinatal depression directly causes adverse effects in offspring, due to a paucity of data, potential confounds, and inconsistent findings in studies to date. What does consistently emerge from existing data is that maternal depression is an additive risk factor for adverse effects on offspring in vulnerable families (McLennan & Offord, 2002).

In summary, the aggregate of evidence supports the idea that perinatal depression warrants universal screening, provided that there is a valid, acceptable, cost-effective screening tool (Guideline #2), and effective treatment (Guideline #3). This chapter examines validity data for relevant screening tools, as well as emerging data about treatment for perinatal depression.

Outcomes

The primary outcomes of interest in assessing the validity of a depression screening tool are its sensitivity (proportion of "true positives" or women scoring above the cut-off who are depressed), and specificity (proportion of "true negatives" or women scoring below the cut-off who are not depressed). In order to study these outcome variables, "depression" must be defined using consensus criteria, such as the Diagnostic and Statistical Manual of Mental Disorders (DSM), International Classification of Diseases (ICD), or Research and Diagnostic Criteria (RDC). Studies must specify whether they are screening for only major depression, or for both major and minor depression. Assessing validity also requires a reference standard, usually a comprehensive semistructured psychiatric interview tool. Examples of such reference standards for depression are the Structured Clinical Interview for DSM-IV (SCID), the Schedule for Affective Disorders and Schizophrenia (SADS), the Present State Examination (PSE), and the Diagnostic Interview Schedule (DIS).

Several depression screening tools including the Beck Depression Inventory (BDI), the Center for Epidemiologic Studies Depression Screen (CES-D), and the Zung Self-Assessment Depression Scale (SDS) have been validated in primary care settings (Williams, Pignone, Ramirez, & Perez Stellato, 2002). However, in peripartum populations, normal pregnancy-related physical changes can elevate scores on those instruments, such that low-to-medium range scores are not reliably predictive of clinically significant depression (Salamero, Marcos, Gutierrez, & Rebull, 1994). For this reason, specific peripartum depression screening tools have been developed, including the Edinburgh Postnatal Depression Scale (EPDS) (Cox, Holden, & Sagovsky, 1987) and the Postpartum Depression Screening Scale (PDSS) (Beck & Gable, 2001). Below, we will review validation studies of both specific and general scales utilized for perinatal depression screening.

Overview of the Evidence

A review of depression screening tools was conducted from 11/06 to 12/06. The 16 studies included in this review met the following criteria:

- Prospective validation study of depression screening tool(s).
- Study participants pregnant and/or within one year postpartum.
- Study specifies whether screening is for major depression, minor depression or both, as defined by specified consensus criteria.
- Study includes a valid reference standard.
- Publication date January 1985 through August 2006.

Quality of study methodology was evaluated and scored according to criteria utilized by the RTI-University of North Carolina Evidence-Based Practice Center in its Evidence Report on Perinatal Depression: Prevalence, Screening Accuracy, and Screening Outcomes (Gaynes et al., 2005).[1] These criteria include:

- Reporting (whether study aims, depression assessment, potential confounds and procedures are described).
- External validity (to what degree the study population, setting and clinicians are representative).
- Internal validity (to what degree there is potential bias in the use of the screening tool and the reference standard).

[1] The quality of study methodology scoring criteria were different than those used in other chapters in this book as they were based on a previously completed report of perinatal depression screening produced by the Agency for Healthcare Research and Quality.

Studies were scored on individual items in each category, and were given a total quality score. The maximum score using this system is 21, with 15–21 considered good, 8–14 fair, and 0–7 poor. Table 13.1 summarizes study outcomes, and Table 13.2 reports quality scores. All studies that met criteria for inclusion in this review had good to fair quality ratings. Studies previously included in the evidence report cited above (Gaynes et al. 2005) are noted with the scores from that report. Studies published since that report was generated were scored using the same criteria as the Gaynes et al. report.

Most (13) of the studies used the EPDS. Other screening tools studied in postpartum women only (not pregnant women) were the BDI (3 studies), the PDSS (2 studies), and the CES-D (1 study). Table 13.3 summarizes key characteristics of each of these screening tools, including advantages, disadvantages and optimal settings for use.

In addition to the validation data summarized in Table 13.1, some studies have directly compared the performance of tools designed specifically for perinatal depression screening with general depression screening tools in perinatal populations. Overall, the two tools designed specifically for perinatal screening (EPDS and PDSS) have been shown to have greater sensitivity (0.75–1.0 range) than the BDI (0.32–0.68) in peripartum populations (Gaynes et al., 2005). However, a study comparing the EPDS with the CES-D (Guedeney, Fermanian, Guelfi, & Kumar, 2000) found that while the EPDS was better at identifying depression in women with anhedonia (inability to feel pleasure) and anxious symptoms, it was worse at identifying psychomotor retardation (slowing of thought processes and movements).

A key question is which cut-off score will be considered a "positive" screen and result in further intervention. For clinical purposes, a false negative (missing a case of major depression because the tool's sensitivity is too low) is more problematic than a false positive. However, too many false positives can tax the resources of patients and the clinics that serve them. While available validation studies do not provide a definitive answer about optimal cut-off scores, the EPDS studies provide enough information to yield useful guidelines. For example, suppose a clinic decides to aim for a sensitivity of at least 0.8 and a specificity of at least 0.7 in detecting peripartum major depression. Data from the studies summarized in Table 13.1 indicate that EPDS cut-off scores between 10 and 12 have consistently yielded sensitivity and specificity scores in that range, while cut-off scores above 12 have not been sensitive enough in some studies.

Review of the Evidence

After reviewing the evidence, the EPDS emerges as the peripartum depression screening tool with the most data exploring its validity for use during pregnancy and postpartum, with a cut-off score between 10 and 12 yielding sensitivity and specificity outcomes that are appropriate for most clinical settings. While some general depression screening tools appear to have acceptable validity in postpartum populations, they have not been validated for use in pregnant women, and available studies suggest that when compared directly with the EPDS, the latter performs better in perinatal populations. Further, EPDS validation studies have been carried out in diverse cultural groups and in many different languages (Cox & Holden, 2003).

A key disadvantage of the EPDS for practical clinical use is that its items are not linked to diagnostic criteria for depression from the Diagnostic and Statistical Manual of Mental Disorders (DSM). Scoring above a cut-off on the EPDS indicates that a woman is at risk of having clinically significant depression, but it does not indicate whether she meets DSM criteria for depression. This contrasts with the 9-item Patient Health Questionnaire (PHQ-9), a tool commonly used to screen for depression in primary care clinics. The PHQ-9, having been derived from DSM criteria, can be used not only for screening but for diagnostic assessment. PHQ-9 scores indicate whether a woman meets DSM criteria for major depression, and if so, the level of severity of depression. Unlike the

Table 13.1 Major outcomes (validity) associated with studies of perinatal depression screening tools

Author, year	Study design/ study type	Description of screening tool	Populations studied and sample size	Address disparities (yes/no)	Key findings related to intervention effectiveness			Caveats/biases	Findings support validity? yes/no for which populations?
					Cutoff Score	Sensitivity	Specificity		
Whiffen, 1988	Validation study/ published article	BDI, Ontario, Canada	120 postpartum women, mean age 28 years, race/ethnicity not reported, major or minor depression	No	10	0.48	0.86		Yes (reference standard was SADS)
Harris et al., 1989	Validation study/ published article	EPDS and BDI, Wales	147 postpartum women, mean age 24.6 years, race/ethnicity not reported, major depression	No	EPDS:				Yes (reference standards were Raskin 3 Area Scale for Depression and MADRS)
					10	1.0	0.82		
					13	0.95	0.93		
					BDI:				
					11	0.68	0.88		
					13	0.63	0.92		
					21	0.32	0.99		
Murray & Carothers, 1990	Validation study/ published article	EPDS, England	646 postpartum women, age not reported, race/ethnicity not reported, major or minor depression	No	12	0.88	0.925		Yes (reference standard was SPI)
					13	0.811	0.957		
					14	0.731	0.975		
Murray & Cox, 1990	Validation study/ published article	EPDS, UK	100 antepartum women, mean age 24.6 years, race/ethnicity not reported, major and major or minor depression	No	Major depression:				Yes (reference standard was SPI)
					12	1.0	0.79		
					13	1.0	0.87		
					14	1.0	0.94		
					15	1.0	0.96		
					Major or minor depression:				
					11	0.71	0.72		
					12	0.64	0.80		
					13	0.64	0.90		
					14	0.57	0.95		
Campbell & Cohn, 1991	Validation study/ published article	CES-D, USA	1007 postpartum women, age not reported, 100% white, major or minor depression	No	16	0.60	0.92		Yes (reference standard was modified SADS)
					21	0.43	0.97		

(continued)

Author, year	Study design/ study type	Description of screening tool	Populations studied and sample size	Address disparities (yes/no)	Key findings related to intervention effectiveness			Caveats/biases	Findings support validity? yes/no for which populations?
					Cutoff Score	Sensitivity	Specificity		
Boyce et al., 1993	Validation study/ published article	EPDS, Australia	103 postpartum women, mean age 28.4 years, race/ethnicity not reported, major depression	No	10	1.0	0.89		Yes (reference standard was DIS)
					13	1.0	0.96		
Ballard et al., 1994	Validation study/ Published article	EPDS, UK	200 postpartum women, mean age 28.8 years, race/ethnicity not reported, major depression	No	13	0.96	0.70		Yes (reference standard was adaptation of PSE)
Cox et al., 1996	Validation study/ published article	EPDS, UK	128 postpartum women, mean age 27.2 years, race/ethnicity not reported, major and major or minor depression	No	Major depression:				Yes (reference standard was SPI)
					10	0.88	0.71		
					12	0.88	0.76		
					13	0.75	0.84		
					Major or minor depression:				
					10	0.81	0.77		
					12	0.76	0.81		
					13	0.62	0.89		
Leverton & Elliott, 2000	Validation study/ published article	EPDS, England	199 postpartum women, age not reported, race/ ethnicity not reported, major or minor depression	No	10	0.69	0.85		Yes (reference standard was PSE)
					13	0.44	0.92		
Beck & Gable, 2001	Validation study/ published article	EPDS, BDI, and PDSS, USA	150 postpartum women, mean age 31 years (range 18 to 46 years), 87% white, 8% Black, 4% Hispanic, 1% Asian, major and major or minor depression	No	Major depression:				Yes (reference standard was SCID)
					EPDS 13	0.78	0.99		
					BDI 21	0.56	1.0		
					PDSS 81	0.94	0.98		
					Major or minor depression:				
					EPDS 10	0.59	0.86		
					BDI 15	0.57	0.97		
					PDSS 61	0.91	0.72		

Aydin et al., 2004	Validation study/ published article	EPDS (in Turkish), Erzurum, Turkey	341 postpartum women, mean age 26.6 years, all subjects were Turkish speaking, major or minor depression	Yes	9.5 10.5 11.5 12.5 13.5	0.96 0.898 0.796 0.755 0.612	0.47 0.592 0.658 0.715 0.774	High rate of postpartum depression found in this sample (35.8%)	Yes (reference standard was SCID)
Adouard et al., 2005	Validation study/ published article	EPDS (in French), France	60 antepartum women, mean age 31.5 years (range 23 to 46 years), place of birth reported – 68% France, 9% other Europe, 15% Africa, 8% other, major depression	Yes	9.5 10.5 11.5 12.5	0.87 0.80 0.80 0.73	0.71 0.73 0.80 0.82	Most women in this sample had a high-risk pregnancy	Yes (reference standard was DSM-IV diagnosis)
Beck & Gable, 2005	Validation study/ published article	PDSS (in Spanish), Connecticut and Texas	150 postpartum women, mean age 25.75 (range 16 to 44 years), 43% Puerto Rican, 43% Mexican, 12% Central/ South American, 2% Caribbean, major or minor depression	Yes	50 55 60 65	0.96 0.89 0.84 0.75	0.68 0.78 0.84 0.87	High rate of postpartum depression found in this sample (37%)	Yes (reference standard was SCID)
Felice et al., 2006	Validation study/ published article	EPDS (in Maltese), Malta	239 antepartum women, mean age 27.1 years (range 15 to 34 years), race/ethnicity not reported, major and minor depression	Yes	8.5 9.5 10.5 11.5 12.5 13.5 14.5 15.5	1.00 0.906 0.875 0.813 0.781 0.75 0.656 0.50	0.733 0.801 0.843 0.874 0.895 0.958 0.974 0.989		Yes [reference standard was revised Clinical Interview Schedule (CIS-R) with ICD-10 diagnosis]
Jardri et al., 2006	Validation study/ published article	EPDS (in French), Lille, France	815 postpartum women, age not reported, all subjects were French speaking, major or minor depression	Yes	6.5 7.5 8.5 9.5 10.5 11.5 12.5	0.90 0.89 0.87 0.82 0.70 0.60 0.50	0.43 0.49 0.60 0.68 0.74 0.80 0.85		Yes (reference standard was the Mini International Neuropsychiatric Interview)

Table 13.1 (continued)

Author, year	Study design/ study type	Description of screening tool	Populations studied and sample size	Address disparities (yes/no)	Key findings related to intervention effectiveness			Caveats/biases	Findings support validity? yes/no for which populations?
					Cutoff Score	Sensitivity	Specificity		
Werrett & Clifford, 2006	Validation study/ published article	EPDS (in English and in Punjabi), UK	23 postpartum women, mean age 28.6 years (range 23 to 40 years), 74% born in the UK, 26% born outside of the UK	Yes	7.5	1.00	0.625	Small sample size. Needs to be re-evaluated with subjects who are primarily Punjabi-speaking rather than bilingual	Yes [reference standard was the composite international diagnostic interview (CIDI)]
					8.5	1.00	0.625		
					9.5	0.714	0.687		
					10.5	0.714	0.75		
					11.5	0.714	0.875		
					12.5	0.714	0.937		
					13.5	0.714	0.937		
					14.5	0.714	0.937		
					15.5	0.571	0.937		

Abbreviations: *BDI* Beck Depression Inventory, *CES-D* Center for Epidemiologic Studies – Depression Scale, *CIDI* Composite International Diagnostic Interview, *CIS-R* Clinical Interview Schedule, Revised, *DIS* Diagnostic Interview Schedule, *DSM-IV* Diagnostic and Statistical Manual of Mental Disorders: Fourth Edition, *EPDS* Edinburgh Postnatal Depression Scale, *ICD*-10 International Classification of Diseases, Tenth Edition, *MADRS* Montgomery-Asberg Depression Rating Scale, *PDSS* Postpartum Depression Screening Scale, *PSE* Present State Examination, *SADS* Schedule for Affective Disorders and Schizophrenia, *SCID* Structured Clinical Interview for Diagnosis, SPI Standardized Psychiatric Interview, *UK* United Kingdom, *USA* United States of America

Table 13.2 Quality rating of validation studies of perinatal depression screening tools

Author, year	Reporting	External validity	Internal validity	Power	Total quality score (0–21) <14=poor, 15–19=fair, >20=good
*Whiffen, 1988	–	–	–	–	10
*Harris et al., 1989	–	–	–	–	13
*Murray & Carothers, 1990	–	–	–	–	15
*Murray & Cox, 1990	–	–	–	–	16
*Campbell & Cohn, 1991	–	–	–	–	19
*Boyce et al., 1993	–	–	–	–	16
*Ballard et al., 1994	–	–	–	–	18
*Cox et al., 1996	-	–	–	–	13
*Leverton & Elliott, 2000	–	–	–	–	12
*Beck & Gable, 2001	–	–	–	–	15
Aydin et al., 2004	6	2	8	0	13
Adouard et al., 2005	6	1	4	0	15
Beck & Gable, 2005	6	2	8	0	16
Felice et al., 2006	3	3	8	0	14
Jardri et al., 2006	7	3	6	0	16
Werrett & Clifford, 2006	6	1	8	0	15

All studies published since the Agency for Healthcare Research and Quality (AHRQ) *Evidence Report on Perinatal Depression: Prevalence, Screening Accuracy, and Screening Outcomes* (Gaynes et al., 2005) that meet criteria noted above are rated here per the Quality Checklist for Studies of Screening Instruments/Procedures. For studies included in the AHRQ report (marked with *), the total quality score is reported here (subdivisions of the total quality score were not given)

EPDS, there are data to support the use of PHQ-9 scores for tracking treatment response (Lowe, Unutzer, Callahan, Perkins, & Kroenke, 2004), and the clinical relevance of changes in scores is clear. Finding a single tool to screen, assess and track treatment response is especially useful in Depression Disease Management Models, which are comprehensive systems of detecting and treating depression in primary care settings (Neumayer-Gromen, Lampert, Stark, & Kallischnigg, 2004). While the PHQ-9 has not yet been formally validated for use in pregnant or postpartum women, two studies (Kelly, Zatzick, & Anders, 2001; Spitzer, Williams, Kroenke, Hornvak, & McMurray, 2000) have piloted its use in peripartum clinic populations.

What remains to be clearly established is whether implementing a peripartum depression screening program leads to demonstrable alleviation of depression. Available studies suggest that the ultimate effectiveness of screening may depend greatly on whether there are viable systems in place to assess and treat women who score above a cut-off on a screening tool ("positive screens"). Referring all women with positive screens to mental health professionals may not be feasible due to insufficient mental health resources, prohibitive expense, and logistical and cultural obstacles. Research to date suggests that acceptance of perinatal depression screening is high, but acceptance of assessment by a mental health professional after a positive screen is low. In one study, for example (Carter et al., 2005), 93% of pregnant women presenting for routine ultrasound examinations agreed to complete the EPDS. However, among women who scored above the EPDS cut-off in that study, only 30.6% agreed to assessment by a mental health professional, and less than half of those actually attended the assessment interview.

Implementation of the Evidence

Overall, the evidence is compelling that universal screening for depression in pregnant and postpartum women would greatly improve detection of this highly prevalent, high-risk condition (Powers, Zahorik, & Morrow, 1993; Evins, Theofrastous, & Galvin, 2000; Marcus, Flynn, Blow, & Barry, 2003;

Table 13.3 Depression screening tools: use in perinatal populations

Screening Tool	Description	Advantages	Disadvantages	Recommended Use
Beck Depression Inventory II	21-item self-report designed for use in clinic populations	Linked to DSM-IV Tracks response to treatment	Few validation data postpartum; none antepartum Potential somatic confounds	Primary care clinics that deliver perinatal care
Center for Epidemiologic Studies – Depression Scale	20-item self-report designed for community use	Better than EPDS at identifying psychomotor retardation in one study	Few validation data postpartum; none antepartum Potential somatic confounds	Epidemiologic & community studies
Edinburgh Postnatal Depression Scale	10-item self-report designed for peripartum use	Brief, easy to use Well validated in many cultures & settings Reduces somatic confounds	Not validated for tracking treatment response (clinical significance of score reduction unclear)	OB clinics Doulas Pediatric clinics
Patient Health Questionnaire	9-item self-report designed for primary care use	Brief, easy to use Linked to DSM-IV Tracks response to treatment	Piloted but not yet validated peripartum	Primary care clinics that deliver perinatal care
Postpartum Depression Screening Scale	35-item self-report designed for post-partum use	Greater symptom detail Reduces somatic confounds	Time-consuming Few validation data postpartum; none antepartum Not linked to DSM-IV	Psychotherapists and counselors

Abbreviations: *BDI* Beck Depression Inventory, *CES-D* Center for Epidemiologic Studies – Depression Scale, *DSM-IV* Diagnostic and Statistical Manual of Mental Disorders – Fourth Edition, *EPDS* Edinburgh Postnatal Depression Scale, *OB* obstetric, *PDSS* postpartum depression screening scale, *PHQ-9* patient health questionnaire – 9-item version

Morris-Rush, Freda, & Bernstein, 2003; Smith et al., 2004). However, in order to produce clinical improvement, screening must happen within a context where it can reliably lead to assessment, treatment and follow-up for women who screen positive.

Both the EPDS and the PHQ-9 can serve as useful tools in different contexts. Clinical and public health initiatives that focus exclusively on peripartum populations, including multiethnic populations, can employ the EPDS with confidence due to its well-established validity, cross-cultural applicability, and availability in many languages. By contrast, a primary care clinic that serves peripartum women as a sub-population may choose to use the PHQ-9 for general depression screening rather than incur the extra costs and logistical confusions of using two different screening tools for different patient populations. Also, a prenatal or primary care clinic that aims to conduct a substantial amount of assessment and treatment of women with positive screens (as opposed to referring all women with positive screens to an external mental health service) may find that the PHQ-9 streamlines assessment and tracking of treatment response. However, validation studies of the PHQ-9 in perinatal populations are needed to ensure that it performs adequately.

An important limitation of all depression screening tools is that they do not rule out bipolar disorder, which has been found to be common and under-recognized in primary care settings (Das et al., 2005). Treating patients with antidepressants who have bipolar disorder increases the risk of mania and rapid cycling (very frequent episodes of mania and/or depression) (Mundo, Cattaneo,

Russo, & Altamura, 2006). Some data suggest that a history of bipolar disorder is an even more potent risk factor than a history of unipolar depression for developing depression during pregnancy or postpartum (Viguera, 2005). Therefore, clinical assessment after using any depression screening tool must include assessment for bipolar disorder.

Any clinic launching a perinatal depression screening program also needs a system in place for assessing level of suicide risk. Although some data suggest that suicide rates are lower during pregnancy than at other times (Marzuk et al., 1997), suicide remains an important cause of death in peripartum women (Shadigian & Bauer, 2005). In a study of women receiving prenatal care in public-sector clinics, 23% screened positive for a current depressive disorder, and of those, 19% endorsed current thoughts of harming themselves (Smith et al., 2004). Fewer than 12% of the women endorsing suicidal thoughts were detected or referred for treatment by their obstetric clinicians. Similarly, in a study of soldiers screened with the EPDS during pregnancy and postpartum, 11% endorsed suicidal thoughts (O'Boyle, Magann, Ricks, Doyle, & Morrison, 2005). Most depression screening tools, including the EPDS, contain single items asking about suicidal thoughts. A system for "flagging" screens endorsing those items for urgent assessment and intervention is an essential element of a screening program.

Treatment of Perinatal Depression

Overview of Treatment Approaches

There is ample evidence that major depression is a treatable condition. Numerous studies have demonstrated efficacy for antidepressant medication, electroconvulsive therapy, and certain forms of psychotherapy, such as cognitive-behavior therapy (CBT) and interpersonal psychotherapy (IPT). Pilot studies suggest efficacy for emerging and/or adjunctive treatments for depression, such as phototherapy and aerobic exercise.

There is a relative paucity of research on the efficacy of treatments specifically for perinatal depression, as it is often assumed that treatments for major depression in general will be effective for pregnant or postpartum women. However, there are reasons to posit that treatment response could differ for women who are pregnant or who have recently given birth. Response to biologically based treatments, such as medications, may differ for depressive symptoms that are influenced by hormonal flux. Abrupt changes in estrogen levels appear to affect serotonin more than they affect other neurotransmitters, so antidepressants that boost serotonergic activity may be more effective for postpartum depression than antidepressants that affect other neurotransmitter systems (Payne, 2003). In addition, social role transitions and a lack of social support may play a key role in increasing vulnerability to perinatal depression (Spinelli & Endicott, 2003). This suggests the possibility that perinatal depression might respond preferentially to IPT as compared to other forms of psychotherapy.

Outcomes

The primary outcomes of interest in assessing the efficacy of a treatment intervention for peripartum depression are: (1) whether it significantly decreases mean group scores on a depression scale (continuous approach), or (2) whether it leads to remission of major depression in significantly more women (binary approach). The latter is more reliably clinically significant. By contrast, statistically comparing mean group depression scores between two samples of women does NOT yield information on how many participants show adequate clinical improvement (Matthey, 2004).

In either case, the intervention needs to be compared to a control condition without major confounds. For example, if an intervention involves extensive patient-provider contact and the control condition does not, the extensive contact might explain any differences in outcome, regardless of the nature of the intervention.

Overview of the Evidence

Due to the early stage of development of this line of research, there are only a small number of studies that have systematically examined treatment interventions for peripartum depression. Therefore, the review presented here uses less stringent criteria than for studies related to depression screening for study inclusion, as follows:

- Prospective randomized controlled trial of a treatment intervention for depression.
- Use of a standardized outcome measure for depression.
- Study participants pregnant and/or within one year postpartum.
- Publication date January 1985 through August 2006.

Studies meeting these criteria were included even if they were under-powered, failed to rule out key confounds, failed to adequately describe the nature of the intervention, did not demonstrate uniformity in the intervention, and/or were not blinded.

Table 13.4 summarizes study outcomes, and Table 13.5 shows quality ratings. The studies are subdivided into three categories of intervention type:

- Psychotherapies – eight studies examined the efficacy of specific types of psychotherapy, sometimes alone and sometimes combined with other interventions.
- Psycho-educational/social support – five studies investigated the efficacy of strategies to educate about peripartum depression and/or enhance social support.
- Biological modalities – five studies examined biological interventions, such as antidepressant medications, phototherapy, and estrogen therapy.

A key overall limitation is that most of the studies have small sample sizes, high rates of participant decline and intervention attrition, leading to limited statistical power. In addition, most of the studies rely on statistically significant changes in mean group scores on depression scales as the sole outcome measure. This does not allow for determination of whether the intervention leads to clinically significant improvement in individual women. Only three studies used symptom remission measures as outcome criteria; two of those (O'Hara, Stuart, Gorman, & Wenzel, 2000; Spinelli & Endicott, 2003) evaluated the efficacy of interpersonal psychotherapy, and one (Wisner, Hanusa, & Perel, 2006) compared the efficacy of two antidepressant medications, but without placebo controls.

Review of the Evidence

Due to the methodological limitations of studies to date, the current data base can best be used to point to promising interventions rather than to draw definitive conclusions about effective treatments for peripartum depression. Among psychotherapy studies, the data are strongest for interpersonal psychotherapy, with two studies (O'Hara et al., 2000; Spinelli & Endicott, 2003) showing clinically and statistically significant increases in remission from depression in women receiving IPT versus women in control conditions. The four studies attempting to evaluate the efficacy of

Table 13.4 Major outcomes associated with studies of interventions for perinatal depression

Author, year	Study design/ study type	Description of intervention	Populations studied and sample size	Address disparities (yes/no)	Key findings related to intervention effectiveness	Caveats/biases	Findings support the intervention? yes/no for which populations?
Intervention type #1: studies of psychotherapy							
Holden, Sagovsky, & Cox, 1989	RCT/ published article	– Rogerian therapy – 8 weekly sessions of Rogerian therapy – Control: routine primary care	– Major or minor depression per RDC – Postpartum – Race/ethnicity, age, SES not reported – *N*=50	No	Significantly more EPDS score reduction with Rogerian therapy vs. control condition	12 women in the intervention group took antidepressant medication	Yes – for postpartum women
Meager & Milgrom, 1996	RT/published article	– CBT/psych-educational group therapy – 10 weekly CBT/ psych-educational group therapy sessions – Control: wait list	– Depression per EPDS, BDI and Profile of Mood States (POMS) – Postpartum – Race/ethnicity, SES not reported – *N*=20	No	Significantly greater EPDS & BDI score reduction with intervention vs. control condition	– Small sample size – Intervention not defined enough to be replicated	Yes – for postpartum women
Appleby et al., 1997	RCT/published article	– Fluoxetine pharma-cotherapy with CBT – 4 groups: • 1 session CBT + med • 6 session CBT + med • 1 session CBT + placebo • 6 session CBT + placebo	– Major or minor depression per Research Diagnostic Criteria (RDC) with EPDS>9 & modified Clinical Interview Schedule (CIS) >11 – Postpartum – Mean age 26.3 – Race/ethnicity, SES not reported – *N*=87	No	– Improvement in women receiving fluoxetine was significantly greater than in those receiving placebo. – Significantly more reduction in modified CIS & Hamilton Rating Scale – Depression (HRSD) scores with 6 versus 1 session of CBT – No improvement by combining fluoxetine & CBT	– Small N within cells – High degree of individual variation in treatment response among study participants – Insufficient reporting of the type of CBT, short duration of CBT	Yes – for non-psychotic depression in postpartum women

(continued)

Table 13.4 (continued)

Author, year	Study design/ study type	Description of intervention	Populations studied and sample size	Address disparities (yes/no)	Key findings related to intervention effectiveness	Caveats/biases	Findings support the intervention? yes/no for which populations?
Brugha et al., 2000	RCT/published article	– CBT+social support – Six 2-h weekly classes based on CBT plus social support – Control: routine prenatal care	– Depression per General Health Questionnaire (GHQ-D) score>1 – Postpartum, primiparous – Median age 19 – Intervention group 72.8% European; control group 72.6% Asian – *N*=190	No	No significant difference in depressive symptom improvement (per GHQ-D scores) with intervention vs. controls	55% of participants in intervention group attended insufficient number of sessions to receive hypothesized benefit	No
O'Hara et al., 2000	RCT/published article	– IPT – 12 weekly 60-min IPT sessions – Control: wait list	– Major depression per DSM-IV – At least age 18 – Almost 100% white – SES not reported – *N*=120	No	Significantly more remission with IPT (p=0.001–0.007 depending on scale used)	– Interviewers not blinded – A majority of eligible women declined participation, and about 20% withdrew from treatment – Study participants almost all well educated and in stable relationships, so findings cannot be generalized	Yes – for white women with major depression

Clark et al., 2003	RCT/published article	– IPT & mother-infant dyadic therapy – 3 groups: • Mother-infant dyadic therapy (12 weekly 1-h sessions) • IPT (12 weekly 1-h sessions) • Control: wait list	– Major depression per DSM-IV – Postpartum – Mean age 31.4 – 97% white – Mean family income $33,353 – *N*=39	No	Significantly more reduction in BDI and CES-D score in either intervention group vs. control group	– Severity of depression differed at baseline among groups – Antidepressant medication used by more women in IPT than dyadic group	Yes – for white postpartum women
Cooper et al., 2003	RCT/published article	– CBT & psycho-dynamic psycho-therapy – 4 groups (10 weekly sessions each): • CBT • Psycho-dynamic psycho-therapy • Nondirective counseling • Control: routine primary care	– EPDS score >11 – Postpartum, primiparous – Mean age 28 – Race/ethnicity, SES not reported – *N*=193	No	– Of the interventions, only psychodynamic therapy produced more symptom reduction than controls per Structured Clinical Interview for DSM III-R (SCID) – Differences not sustained at 9 months postpartum; treatment did not ameliorate subsequent depressive episodes	– Nature & quality of interventions not adequately described – Underpowered to detect between-group differences	Yes – for postpartum women
Spinelli & Endicott, 2003	RCT/published article	– IPT – 16 weekly 45-min IPT sessions – Control: 16 weekly 45-min parenting education program	– Major depression per DSM-IV; HRSD >11 – Pregnant – Ages 18–45 – 66% Latina, 29% white, 5% African-American – Majority low SES – *N*=38	No	– More symptom reduction with IPT ($p<0.03$) – More remission with IPT on Clinical Global Impression (CGI) ($p<0.02$), but not HRSD ($p=0.40$)	HRSD contains more items that could be artificially elevated by normal pregnancy changes	Yes – for women with major depression

(continued)

Table 13.4 (continued)

Author, year	Study design/ study type	Description of intervention	Populations studied and sample size	Address disparities (yes/no)	Key findings related to intervention effectiveness	Caveats/biases	Findings support the intervention? yes/no for which populations?
Intervention type #2: supportive/educational interventions							
Wickberg & Hwang, 1996	RCT/ published article	– Health visitor counseling – Six 1-h non-directive counseling sessions – Control: routine primary care	– EPDS score >12, major depression per DSM-III and Montgomery Asberg Depression Rating Scale (MADRS)>9 – Postpartum – Mean age 28.4 – Race, SES not reported – Swedish – *N*=41	No	Significantly greater reduction of depression (as measured by MADRS score and DSM-III criteria) with intervention versus control ($p<0.01$)	– Nature and quality of intervention not adequately described and not manualized	Yes – for Swedish women with major depression
Misri et al., 2000	RCT/published article	– Partner support – 7 psycho-educational visits, 4 of which included partners – Control: 7 psycho-educational visits without partners	– Major depression with postpartum onset per DSM-IV – Majority white – Age, SES not reported – *N*=29	No	Significant decrease in EPDS & Mini International Neuropsychiatric Interview (MINI) with intervention vs. control condition	– Marital satisfaction differed at baseline between groups – Intervention not manualized – Underpowered	Yes – for women with major depression with postpartum onset
Honey et al., 2002	RCT/published article	– Group psycho-education – 8 weekly 2-h psycho-educational group sessions – Control: routine primary care	– EPDS score >11 – Postpartum – Mean age 29.30 – Race/ethnicity, SES not reported – *N*=45	No	– Significantly greater EPDS score reduction with intervention vs. control condition ($p=0.01$) – Effects maintained for 6 months	– Significantly more women in intervention group were married or cohabitating – Intervention not manualized	Yes – for postpartum women

Reid et al., 2002	RCT Published Article	– Support group – Weekly 2-h support group – Control: routine prenatal care	– EPDS score>11 – Postpartum, primiparous – Race/ethnicity not reported – Mean age 26.5 – SES 8.2–9.7% deprived; 87–89% intermediate to affluent social class – *N*=717	No	No significant difference in EPDS score change between intervention and control groups	Only 18% of qualified participants assigned to the support group intervention attended a support group session	No
Dennis, 2003	RCT (Pilot)/ published article	– Peer Support – Telephone based peer support – Control: standard community postpartum services	– EPDS score>9 – Postpartum, primiparous – Race/ethnicity not reported – SES 82%<$80,000 (control); 94%<$80,000 (experimental) – *N*=42	No	– Significantly greater EPDS score reduction at 4 and 8 weeks (p=0.02; p=0.01 respectively) in intervention vs. control group	– Frequency and duration of peer interactions not standardized – Intervention not manualized	Yes – for postpartum women
Intervention type #3: other modalities							
Gregoire et al., 1996	RCT/ published article	– Estrogen therapy – 17-beta estradiol x 12 weeks; dedrogesterone added for 12 days each month – Control: placebo	– Depression per EPDS >14 – Postpartum – Race/ethnicity, SES not reported – *N*=61	No	Estrogen improved EPDS scores significantly more than placebo	Antidepressant medication use by 47% of women in intervention group and 37% of women in control group	Yes – for postpartum women

(continued)

Table 13.4 (continued)

Author, year	Study design/ study type	Description of intervention	Populations studied and sample size	Address disparities (yes/no)	Key findings related to intervention effectiveness	Caveats/biases	Findings support the intervention? yes/no for which populations?
Appleby et al., 1997	RCT/published article	– Fluoxetine pharmacotherapy with CBT – 4 groups: • 1 session CBT + med • 6 session CBT + med • 1 session CBT + placebo • 6 session CBT + placebo	– Major or minor depression per Research Diagnostic Criteria (RDC) with EPDS>9 & modified Clinical Interview Schedule (CIS)>11 – Postpartum – Mean age 26.3 – Race/ethnicity, SES not reported – *N*=87	No	– Improvement in women receiving fluoxetine was significantly greater than in those receiving placebo. – Significantly more reduction in modified CIS & Hamilton Rating Scale Depression (HRSD) scores with 6 vs. 1 session of CBT – No improvement by combining fluoxetine & CBT	– Small N within cells – High degree of individual variation in treatment response among study participants – Insufficient reporting of the type of CBT, short duration of CBT	Yes – for non-psychotic depression in postpartum women
Logsdon et al., 2003	RCT/published article	– Tricyclic vs. SSRI antidepressants – Nortriptyline increased stepwise to 150 mg/day – Sertraline increased stepwise to 200 mg/day	– Major depression per DSM-IV & HRSD score >18 – Postpartum, primiparous – 78% white – 57% middle to upper-middle class – *N*=61	No	HRDS, EPDS & Clinical Global Impressions (CGI) scores & certain facets of role functioning improved in both groups, with no significant difference between groups	Two serotonergic agents were compared instead of serotonergic and non-serotonergic	Yes
Epperson et al., 2004	RCT/published article	– Phototherapy – 7000 lux light box (5 weeks) – Control: 500 lux light box (5 weeks)	– Major Depression per DSM-IV – Pregnant – Mean age 32.1 – 80% white, 10% African-American, 10% Hispanic – *N*=10	No	No significant differences in symptom reduction with intervention versus control (p=0.88)	– Study under-powered – Promising trend toward improvement with active light for patients continuing beyond 5 weeks	No

Wisner et al., 2006	RCT/published article	– Tricyclic (nortriptyline) vs. SSRI (sertraline) antidepressants – 8 weeks comparative trial – Fixed dosing strategy	– Major depression with postpartum onset per DSM-IV – HRSD>17 – Age 15–45 – Race/ethnicity, SES not reported – *N*=95	No	No differences in remission (as measured by 50% reduction in HRSD & CGI scores), psychosocial functioning improvement, or side effect burden (as measured by the Asberg Side Effects Rating Scale)	– No placebo group – More minority women assigned to sertraline group	No

Abbreviations: *BDI* beck depression inventory, *CGI* clinical global impression, *CIS* clinical interview schedule, *CBT* cognitive-behavioral therapy, *DSM-III* Diagnostic and Statistical Manual of Mental Disorders – Third Edition, *DSM-IV* Diagnostic and Statistical Manual of Mental Disorders – Fourth Edition, *EPDS* Edinburgh Postnatal Depression Scale, *GHQ-D* General Health Questionnaire – Depression, *HRSD* Hamilton Rating Scale – Depression, *IPT* Interpersonal Psychotherapy, *MADRS* Montgomery-Asberg Depression Rating Scale, *MINI* Mini International Neuropsychiatric Interview, *N* study participant number, *POMS* Profile of Mood States, *RCT* randomized controlled trial, *RDC* Research Diagnostic Criteria, *SES* socioeconomic status, *SSRI* Serotonin-selective reuptake inhibitor

Table 13.5 Quality ratings of studies associated with treatment for perinatal depression

Treatment type	Author, year	Reporting	External validity	Internal validity-bias	Internal validity-confounding	Power	Total quality score ≤14=poor, 15–19=fair, ≥20=good	Suitability of study to assess effectiveness: greatest, moderate, least
Psychotherapy	Holden, Sagovsky, & Cox, 1989	8	2	5	5	0	20 (good)	Greatest
Psychotherapy	Meager & Milgrom, 1996	9	3	3	4	0	19 (fair)	Greatest
Psychotherapy	Appleby et al., 1997	11	4	5	5	0	25 (good)	Greatest
Psychotherapy	Brugha et al., 2000	10	4	5	5	1	25 (good)	Greatest
Psychotherapy	O'Hara et al., 2000	10	4	5	5	1	25 (good)	Greatest
Psychotherapy	Clark et al., 2003	9	2	4	1	0	16 (fair)	Greatest
Psychotherapy	Cooper et al., 2003	10	4	4	4	1	23 (good)	Greatest
Psychotherapy	Spinelli et al., 2003	9	2	3	4	0	18 (fair)	Greatest
Supportive/educational	Wickberg et al., 1996	7	4	3	4	0	18 (fair)	Greatest
Supportive/educational	Misri et al., 2000	11	3	3	4	0	21 (good)	Greatest
Supportive/educational	Honey et al., 2002	7	1	3	3	0	14 (poor)	Greatest
Supportive/educational	Reid et al., 2002	11	4	4	5	1	25 (good)	Greatest
Supportive/educational	Dennis, 2003	11	3	4	3	1	22 (good)	Greatest
Combined pharmacotherapy & CBT	Appleby et al., 1997	11	4	5	5	0	25 (good)	Greatest
Phototherapy	Epperson et al., 2004	9	1	4	4	0	18 (fair)	Greatest
Hormone therapy	Gregoire et al., 1996	10	2	5	5	0	22 (good)	Greatest
Pharmacotherapy	Logsdon et al., 2003	7	1	4	4	0	16 (fair)	Greatest
Pharmacotherapy	Wisner et al., 2006	12	4	5	6	1	28 (good)	Greatest

cognitive-behavior therapy show mixed results. Two (Appleby, Warner, Whitton, & Faragher, 1997; Meager & Milgrom, 1996) demonstrated statistically significant symptom reduction with CBT, while two (Brugha et al., 2000; Cooper, Murray, Wilson, & Romaniuk, 2003) showed no statistically significant difference between women undergoing CBT versus women in control groups. Of note, the study by Cooper and colleagues (2003) was underpowered, and the majority of women in the intervention arm in the study by Brugha and colleagues (2000) did not attend enough of the six intervention sessions to achieve the expected benefit.

The findings from psychoeducational/social support intervention studies indicate that this is a promising approach, as predicted by studies showing that lack of social support is a risk factor for perinatal depression (Surkan, Peterson, Hughes, & Gottlieb, 2006). Of the five studies in this category, four showed statistically significant improvement with the intervention as compared to control conditions. However, in two of the four positive studies (Misri, Kostaras, Fox, & Kostaras, 2000; Honey, Bennett, & Morgan, 2002) a major confound is that the women in the intervention groups had better social support at baseline than women in the control groups.

Among the five studies of biological interventions, three studied antidepressant medication. Due to ethical considerations, only the oldest study of those three (Appleby et al., 1997) included a placebo control group; as expected, the antidepressant (fluoxetine) led to significantly more symptom reduction than placebo. The more relevant question is whether certain antidepressants perform better than others for perinatal depression, as is the case for premenstrual dysphoric disorder (PMDD). Unfortunately, the two studies that compare antidepressants to one another (Logsdon, Wisner, Hanusa, & Philips, 2003; Wisner et al., 2006) did not choose agents that would best address this question. PMDD studies consistently show that serotonergic antidepressants outperform non-serotonergic antidepressants, so an optimal postpartum study would compare a serotonergic with a non-serotonergic antidepressant. While the two studies reviewed here compare medications of different classes (a tricyclic and a serotonin-selective reuptake inhibitor), both of the medications studied affect serotonin, and therefore would not be hypothesized to behave differently, as indeed they do not.

The two other studies of biological modalities are small pilot studies of estrogen therapy (Gregoire, Kumar, Everitt, Henderson, & Studd, 1996) and phototherapy (Epperson et al., 2004). The study of phototherapy is underpowered, which may explain the lack of a statistically significant difference between the intervention and control groups despite a promising trend toward greater improvement with phototherapy. A key confound in the estrogen study is that many study participants were simultaneously taking antidepressant medication, more so in the intervention group than in the control group.

The only study (Appleby et al., 1997) specifically comparing combined treatment modalities (medication plus CBT) to single modalities did not find that combining interventions conferred added benefits compared to each separate modality. Several methodological limitations of this study (small sample size, highly variable responses, insufficient reporting of the type of CBT, short duration of CBT) preclude definite conclusions about the efficacy of combined psychotherapy and medication.

Implementation of the Evidence

A key practical issue that emerges from these treatment studies is the difficulty with acceptance of treatment by women who are depressed during pregnancy and postpartum. Indirect evidence suggests that the level of treatment burden (e.g., time and resource commitment) strongly influences whether patients accept treatment. For example, a trial of cognitive-behavior therapy for pregnant women with depressive symptoms found that only one of 49 study participants who screened

positive for depression completed a course of CBT (Carter et al., 2005). In another study of an intended six-session CBT intervention for women with peripartum depression (Brugha et al., 2000), 55% of women in the intervention group did not attend enough sessions to have any expected benefit. In a support group intervention (Reid, Glazener, Murray, & Taylor, 2002), only 18% of eligible women attended. Similarly, in a study of IPT for postpartum depression (O'Hara et al., 2000), a majority of eligible women declined participation, and about 20% withdrew from treatment. By contrast, studies of a lower-burden intervention (weekly telephone calls from a volunteer or companion) found much higher acceptance rates (Wolman, Chalmers, Hofmeyr, & Nikodem, 1993; Stamp, Williams, & Crowther, 1995). Unfortunately, evidence thus far does not support the effectiveness of these low-burden non-pharmacologic interventions in treating major depression.

An optimistic note is sounded by a study (Birndorf, Madden, Portera, & Leon, 2001) that underscores the role of direct involvement of prenatal care providers in promoting acceptance of mental health treatment. In that study, 100% of depressed pregnant women reported a willingness to see a mental health professional if their obstetrician referred them.

Future Directions and Public Health Challenges

There are two central challenges that need to be addressed for health care systems to more effectively detect and treat perinatal depression. The first is incorporating existing data about screening and treatment into widespread, cost-effective systems of care. The second is generating new data about the relationship between health disparities and perinatal depression, to guide policies that can effectively reduce those disparities.

Incorporating Current Knowledge into Public Policy

Available data provide a compelling rationale for incorporating systematic perinatal depression screening into programs and clinics that serve large numbers of pregnant and postpartum women. Perinatal depression screening is being promoted by several national initiatives, including Healthy Start and Healthy Beginnings. In addition, several states, including New Jersey, Texas, and Illinois, have passed legislation promoting and/or mandating educating women about perinatal depression and screening for its presence.

Public policy initiatives that show promise in supporting screening efforts include:

- *Reimbursing health care providers for screening*. Traditional health insurance policies rarely reimburse for screening. Without reimbursement, clinics and programs lack resources to administer and score screens and address the needs of women with positive screens. The Illinois Department of Healthcare and Family Services (HFS), the state Medicaid agency, piloted such reimbursement, and found a 96.32% increase in the number of screens in the first year of the pilot project (Maram & Murphy, 2006).
- *Incorporating perinatal depression screening as a quality performance measure*. Organized health care systems, such as managed care organizations and Medicaid provider networks, can specify perinatal depression screening as a component of quality perinatal care. This can be supported with chart audits for screening, feedback to health care providers about their performance on this measure, and pay-for-performance systems. Health Plan Employer Data and Information Set (HEDIS) performance measures and Early and Periodic Screening, Diagnosis and Treatment (EPSDT)

guidelines are examples of relevant systems/programs that could include perinatal depression screening.

- *Providing support for clinicians to address women with positive screens*. Stand-alone screening programs run the risk of identifying women with depression but not being able to provide them with interventions. By contrast, a screening program that includes ready access to mental health consultation improves the capacity of primary and prenatal care providers to offer effective treatment to women who screen positive. An example is the Illinois Perinatal Mental Health Consultation Service, which offers expert mental health consultation to all health care providers statewide via toll-free telephone or on line.

Challenges Regarding Health Disparities and Perinatal Depression

Available data suggest that women who are unmarried, have lower educational levels, lower socioeconomic status, more stress, and more unwanted pregnancies are more vulnerable to perinatal depression (Field, Hernandez-Reif, & Diego, 2006b). Studies also suggest that the risks of untreated perinatal depression may be greater in women with additional vulnerabilities (Orr & Miller, 1995; Hoffman & Hatch, 2000). Initiatives addressing perinatal health care disparities, such as Healthy Start, also underscore the profound dearth of mental health care received by high-risk women. For example, in high-risk Chicago neighborhoods, only 0.3–0.6% of pregnant and postpartum women receive mental health services (Saunders, 2006). However, a systematic review of studies to date on perinatal depression shows that many did not report full demographic information, and among those that did, study participants were predominantly white, partnered, and of moderate to high socioeconomic status (Ross, Campbell, Dennis, & Blackmore, 2006). In view of data suggesting that depression may be a treatable component of economic, racial and ethnic disparities in poor birth outcomes, there is a compelling need for future studies to address the following issues:

- How does depression interact with other vulnerability factors to influence outcomes for women and their offspring?
- Are there differences in acceptance and access to different forms of perinatal depression treatment (e.g., medication, psychotherapy) in women of different cultural, ethnic and socioeconomic groups?
- Can perinatal disease management models be designed and implemented that are culturally sensitive and effective for a wide variety of vulnerable populations of women?

Conclusions

Promoting the mental health of new mothers should be a central public health priority, given the high prevalence of perinatal depression and the substantial risks to mothers and their infants when symptoms remain undetected and untreated. There are cost-effective, well-validated, achievable ways of improving the detection, diagnosis, and treatment of perinatal depression. Formal screening clearly improves detection, but can only be effective in improving outcomes when followed by assessment and provision of treatment. Adapting stepped-care disease management models to perinatal care settings is a promising way to achieve this. These would allow for on-site assessment and treatment by prenatal care providers, which is more cost effective and better accepted by patients than routine referral to outside mental health services. Creating a national Perinatal Mental Health Consult Service would be a cost-effective way of ensuring that prenatal health care providers

nationwide could deliver quality care to their patients with adequate training, tools and access to expert consultation.

Depression is one of the most treatable contributory factors to adverse obstetric outcomes. Widespread attention to reducing perinatal depression may be one of the most effective ways to reduce disparities in these outcomes. While depression screening and treatment is essential during the perinatal period, increasing the availability of mental health screening and treatment prior to pregnancy should be explored as a strategy for reducing the burden of mental illness during the peripartum period. Furthermore, increased attention to prevention, the identification and amelioration of the community level and individual level factors associated with the onset of depression, is clearly warranted.

References

Adouard F., Glangeaud-Freudenthal, N. M. C., & Golse, B. (2005). Validation of the Edinburgh postnatal depression scale (EPDS) in a sample of women with high-risk pregnancies in France. *Archives of Womens Mental Health*, *8*, 89–95.

Alder J., Fink, N., Bitzer, J., Hosli, I., & Holzgreve, W. (2007). Depression and anxiety during pregnancy: A risk factor for obstetric, fetal and neonatal outcome? A critical review of the literature. *Journal of Maternal Fetal & Neonatal Medicine*, *20*, 189–209.

Appleby, L., Warner, R., Whitton, A., & Faragher, B. (1997). A controlled study of fluoxetine and cognitive-behavioral counseling in the treatment of postpartum depression. *BMJ*, *314*, 932–936.

Arck, P. C., Rose, M., Hertwig, K., Hagen, E., Hildebrandt, M., & Klapp, B. F. (2001). Stress and immune mediators in miscarriage. *Human Reproduction*, *16*, 505–511.

Aydin, N., Inandi, T., Yigit, A., & Hodoglugil, N. N. (2004).Validation of the Turkish version of the Edinburgh Postnatal Depression Scale. *Social Psychiatry and Psychiatric Epidemiology*, *39*, 483–486.

Ballard, C. G., Davis, R., Cullen, P. C., Mohan, R. N., & Dean, C. (1994). Prevalence of postnatal psychiatric morbidity in mothers and fathers. *British Journal of Psychiatry*, *164*, 782–788.

Beck, C. T., & Gable, R. K. (2001). Comparative analysis of the performance of the Postpartum Depression Screening Scale with two other depression instruments. *Nursing Research*, *50*, 242–250.

Beck, C. T., & Gable, R. K. (2005). Screening performance of the Postpartum Depression Screening Scale – Spanish Version. *Journal of Transcultural Nursing*, *16*, 331–338.

Birndorf, C. A., Madden, A., Portera, L., & Leon, A. C. (2001). Psychiatric symptoms, functional impairment, and receptivity toward mental health treatment among obstetrical patients. *International Journal of Psychiatry in Medicine*, *31*, 355–365.

Bonari, L., Pinto, N., Ahn, E., Einarson, A., Steiner, M., & Koren, G. (2004). Perinatal risks of untreated depression during pregnancy. *Canadian Journal of Psychiatry*, *49*(11), 726–735.

Boyce, P., Stubbs, J., & Todd, A. (1993). The Edinburgh Postnatal Depression Scale: Validation for an Australian sample. *Australian & New Zealand Journal of Psychiatry*, *27*, 472–476.

Brugha, T. S., Wheatley, S., Taub, N. A., Culverwell, A., Friedman, T., Kirway, P., et al. (2000). Pragmatic randomized trial of antenatal intervention to prevent postnatal depression by reducing psychosocial risk factors. *Psychological Medicine*, *30*, 1273–1281.

Buist, A. E., Bryanne, E. W., Milgrom, J., Pope, S., Condon, J. T., Ellwood, D. A., & Boyce, P. M. (2002). To screen or not to screen – that is the question in perinatal depression. *Medical Journal of Australia*, *177*(Suppl), S101–S105.

Campbell, S. B., & Cohn, J. F. (1991). Prevalence and correlates of postpartum depression in first-time mothers. *Journal of Abnormal Psychology*, *100*, 594–599.

Carter, F. A., Carter, J. D., Luty, S. W., Wilson, D. A., Frampton, C. M., & Joyce, P. R. (2005). Screening and treatment for depression during pregnancy: A cautionary note. *Australian & New Zealand Journal of Psychiatry*, *39*, 255–261.

Chung, T. K., Lau, T. K., Yip A. S., Chiu, H. E., & Lee, D. T. (2001). Antepartum depressive symptomatology is associated with adverse obstetric and neonatal outcomes. *Psychosomatic Medicine*, *63*(5), 830–834.

Clark, R., Tluczek, A., & Wenzel, A. (2003). Psychotherapy for postpartum depression: A preliminary report. *American Journal of Orthopsychiatry*, *73*, 441–454.

Cohn, J. F., Campbell, S. B., Matias, R., & Hopkins, J. (1990). Face-to-face interactions of postpartum depressed and nondepressed mother-infant pairs at 2 months. *Developmental Psychology*, *26*, 15–23.

Cooper, P., Murray, L., Wilson, A., & Romaniuk, H. (2003). Controlled trial of the short-and long-term effect of psychological treatment of post-partum depression. *British Journal of Psychiatry*, *182*, 412–419.

Cox, J., Chapman, G., Murray, D., & Jones, P. (1996). Validation of the Edinburgh Postnatal Depression Scale (EPDS) in non-postnatal women. *Journal of Affective Disorders*, *39*, 185–189.

Cox, J., & Holden, J. (2003). *Perinatal mental health: A guide to the Edinburgh Postnatal Depression Scale*. London: Gaskell.

Cox, J. L., Holden, J. M., & Sagovsky, R. (1987). Detection of postnatal depression: development of the 10-item Edinburgh postnatal depression scale. *British Journal of Psychiatry*, *150*, 782–786.

Das, A. K., Olfson, M., Gameroff, M. J., Pilowsky, D. J., Blanco, C., Feder, et al. (2005). Screening for bipolar disorder in a primary care practice. *JAMA*, *293*, 956–963.

Dayan, J., Creveuil, C., Herlicoviez, M., Herbel, C., Baranger, E., Savoye, C., et al. (2002). Role of anxiety and depression in the onset of spontaneous preterm labor. *American Journal of Epidemiology*, *155*, 293–301.

Dayan, J., Creveuil, C., Marks, M. N., Conroy, S., Herlicoviez, M., Dreyfus, M., et al. (2006). Prenatal depression, prenatal anxiety, and spontaneous preterm birth: a prospective cohort study among women with early and regular care. *Psychosomatic Medicine*, *68*, 938–946.

Dennis, C. L. (2003). The effect of peer support on postpartum depression: A pilot randomized controlled trial. *Canadian Journal of Psychiatry*, *48*, 115–124.

Epperson, C. N., Terman, M., Terman, J. S., Hanusa, B. H., Oren, D. A., Peindl, K. S., et al. (2004). Randomized clinical trial of bright light therapy for antepartum depression: Preliminary findings. *Journal of Clinical Psychiatry*, *65*, 421–425.

Evins, G. G., Theofrastous, J. P., & Galvin, S. L. (2000). Postpartum depression: a comparison of screening and routine clinical evaluation. *American Journal of Obstetrics and Gynecology*, *182*, 1080–1082.

Federenko, I. S., & Wadhwa, P. D. (2004). Women's mental health during pregnancy influences fetal and infant developmental and health outcomes. *CNS Spectrums*, *9*, 198–206.

Felice, E., Saliba, J., Grech, V., & Cox, J. (2006). Validation of the Maltese version of the Edinburgh Postnatal Depression Scale (EPDS). *Archives of Womens' Mental Health*, *9*, 75–80.

Ferrara, A., Hedderson, M. M., Quesenberry, C. P., & Selby, J. V. (2002). Prevalence of gestational diabetes mellitus detected by the national diabetes data group or the carpenter and coustan plasma glucose thresholds. *Diabetes Care*, *25*, 1625–1630.

Field, T., Diego, M., & Hernandez-Reif, M. (2006a). Prenatal depression effects on the fetus and newborn: a review. *Infant Behavior & Development*, *29*, 445–455.

Field, T., Hernandez-Reif, M., & Diego, M. (2006b). Risk factors and stress variables that differentiate depressed from nondepressed pregnant women. *Infant Behavior & Development*, *29*, 169–174.

Gavin, N. I., Gaynes, B. N., Lohr, K. N., Meltzer-Brody, S., Gartlehner, G., & Swinson, T. (2005). Perinatal depression: a systematic review of prevalence and incidence. *Obstetrics and Gynecology*, *106*, 1071–1083.

Gaynes, B. N., Gavin, N., Meltzer-Brody, S., Lohr, K. N., Swinson, T., & Gartlehner, G. (2005). *Perinatal depression: Prevalence, screening accuracy, and screening outcomes*. Rockville MD: Agency for Healthcare Research and Quality.

Gregoire, A. J., Kumar, R., Everitt, B., Henderson, A. F., & Studd, J. W. (1996). Transdermal oestrogen for treatment of severe postnatal depression. *Lancet*, *347*, 930–933.

Guedeney, N., Fermanian, J., Guelfi, J. D., & Kumar, R. C. (2000). The Edinburgh Postnatal Depression Scale (EPDS) and the detection of major depressive disorders in early postpartum: some concerns about false negatives. *Journal of Affective Disorders*, *61*(1–2), 107–112.

Haddad, B., & Sibai, B. M. (1999). Chronic hypertension in pregnancy. *Annals of Medicine*, *31*, 246–252.

Halbreich, U., & Karkun, S. (2006). Cross-cultural and social diversity of prevalence of postpartum depression and depressive symptoms. *Journal of Affective Disorders*, *91*(2–3), 97–111.

Harris, B., Huckle, P., Thomas, R., Johns, S., & Fung, H. (1989).The use of rating scales to identify post-natal depression. *British Journal Psychiatry*, *154*, 813–817.

Hatton, D. C., Harrison-Hohner, J., Coste, S., Dorato, V., Curet, L. B., & McCarron, D. (2005). Symptoms of postpartum depression and breastfeeding. *Journal of Human Lactation*, *21*, 444–449.

Hoffman, S., & Hatch, M. C. (2000). Depressive symptomatology during pregnancy: evidence for an association with decreased fetal growth in pregnancies of lower social class women. *Health Psychology*, *19*, 535–543.

Holden, J. M., Sagovsky, R., & Cox, J. L. (1989). Counseling in a general practice setting: controlled study of health visitor intervention in treatment of postnatal depression. *BMJ*, *298*, 223–226.

Honey, K. L., Bennett, P., & Morgan, M. (2002). A brief psycho-educational group intervention for postnatal depression. *British Journal of Clinical Psychology*, *41*, 405–409.

Howell, E., Mora, P., Horowitz, C., & Leventhal, H. (2005). Racial and ethnic differences in factors associated with early postpartum depressive symptoms. *Obstetrics & Gynecology*, *105*, 1442–1450.

Huot, R. L., Brennan, P. A., Stowe, Z. N., Plotsky, P. M., & Walker, E. F. (2004). Negative affect in offspring of depressed mothers is predicted by infant cortisol levels at 6 months and maternal depression during pregnancy, but not postpartum. *Annals of the New York Academy of Sciences*, *1032*, 234–236.

Jardri, R., Pelta, J., Maron, M., Thomas, P., Delion, P., Codaccioni, X., et al. (2006). Predictive validation study of the Edinburgh Postnatal Depression Scale in the first week after delivery and risk analysis for postnatal depression. *Journal of Affective Disorders*, *93*, 169–176.

Kelly, R. H., Zatzick, D. F., & Anders, T. F. (2001). The detection and treatment of psychiatric disorders and substance use among pregnant women cared for in obstetrics. *American Journal of Psychiatry*, *158*, 213–219.

Kurki, T., Hiilesmaa, V., Raitasalo, R., Mattila, H., & Ylikorkala, O. (2000). Depression and anxiety in early pregnancy and risk for preeclampsia. *Obstetrics & Gynecology*, *95*, 487–490.

Leverton, T. J., & Elliott, S. A. (2000). Is the EPDS a magic wand? A comparison of the Edinburgh Postnatal Depression Scale and health visitor report as predictors of diagnosis on the Present State Examination. *Journal of Reproductive and Infant Psychology*, *18*, 279–296.

Lindahl, V., Pearson, J. L., & Colpe, L. (2005). Prevalence of suicidality during pregnancy and the postpartum. *Archives of Women's Mental Health*, *8*, 77–87.

Logsdon, M. C., Wisner, K., Hanusa, B. H., & Phillips, A. (2003). Role functioning and symptoms remission in women with postpartum depression after antidepressant treatment. *Archives of Psychiatric Nursing*, *6*, 276–283.

Lowe, B., Unutzer, J., Callahan, C. M., Perkins, A. J., & Kroenke, K. (2004). Monitoring depression treatment outcomes with the patient health questionnaire-9. *Medicare Care*, *42*, 1194–1201.

Maram, B. S., & Murphy, A. M. (2006). Illinois Department of Healthcare and Family Services: Report to the General Assembly, Public Act 93–0536.

Marcus, S. M., Flynn, H. A., Blow, F. C., & Barry, K. L. (2003). Depressive symptoms among pregnant women screened in obstetrics settings. *Journal of Women's Health*, *12*, 373–380.

Marzuk, P. M., Tardiff, K., Leon, A. C., Hirsch, C. S., Portera, L., Hartwell, N., et al. (1997). Lower risk of suicide during pregnancy. *American Journal of Psychiatry*, *154*, 122–123.

Matthey, S. (2004). Calculating clinically significant change in postnatal depression studies using the Edinburgh Postnatal Depression Scale. *Journal of Affective Disorders*, *78*, 269–272.

McLearn, K. T., Minkovitz, C. S., Strobino, D. M., Marks, E., & Hou, W. (2006). Maternal depressive symptoms at 2 to 4 months post partum and early parenting practices. *Archives of Pediatric & Adolescent Medicine*, *160*, 279–284.

McLennan, J. D., & Offord, D. R. (2002). Should postpartum depression be targeted to improve child mental health? *Journal of the American Academy of Child & Adolescent Psychiatry*, *41*, 28–35.

Meager, I., & Milgrom, J. (1996). Group treatment for postpartum depression: A pilot study. *Australian and New Zealand Journal of Psychiatry*, *30*, 852–860.

Misri, S., Kostaras, X., Fox, D., & Kostaras, D. (2000). The impact of partner support in the treatment of postpartum depression. *Canadian Journal of Psychiatry*, *45*, 554–558.

Misri, S., Oberlander, T. F., Fairbrother, N., Carter, D., Ryan, D., Kuan, A. J., et al. (2004). Relation between prenatal maternal mood and anxiety and neonatal health. *Canadian Journal of Psychiatry*, *49*, 684–689.

Morris-Rush, J. K., Freda, M. C., & Bernstein, P. S. (2003). Screening for postpartum depression in an inner-city population. *American Journal of Obstetrics & Gynecology*, *188*, 1217–1219.

Mundo, E., Cattaneo, E., Russo, M., & Altamura, A. C. (2006). Clinical variables related to antidepressant-induced mania in bipolar disorder. *Journal of Affective Disorders*, *92*, 227–230.

Murray, L., & Carothers, A. (1990). The validation of the Edinburgh Post-natal Depression Scale on a community sample. *British Journal of Psychiatry*, *157*, 288–290.

Murray, L., Cooper, P., & Hipwell, A. (2003). Mental health of parents caring for infants. *Archives of Womens Mental Health*, *6*(Suppl 2), s71–s77.

Murray, D., & Cox, J. L. (1990). Screening for depression during pregnancy with the Edinburgh Depression Scale (EPDS). *Journal of Reproductive and Infant Psychology*, *8*, 99–107.

Neumayer-Gromen, A., Lampert, T., Stark, K., & Kallischnigg, G. (2004). Disease Management Programs for depression: a systematic review and meta-analysis of randomized controlled trials. *Medical Care*, *42*, 1211–1221.

O'Boyle, A. L., Magann, E. F., Ricks, R. E., Doyle, M., & Morrison, J.C. (2005). Depression screening in the Pregnant Soldier Wellness Program. *Southern Medical Journal*, *98*, 416–418.

O'Hara, M. W., Stuart, S., Gorman, L. L., & Wenzel, A. (2000). Efficacy of interpersonal psychotherapy for postpartum depression. *Archives of General Psychiatry*, *57*, 1039–1045.

Orr, S. T., Blazer, D. G., & James, S. A. (2006). Racial disparities in elevated prenatal depressive symptoms among black and white women in eastern North Carolina. *Annals of Epidemiology*, *16*, 463–468.

Orr, S. T., James, S. A., & Prince, C. B. (2002). Maternal prenatal depressive symptoms and spontaneous preterm births among African American women in Baltimore, Maryland. *American Journal of Epidemiology*, *156*, 797–802.

Orr, S. T., & Miller, C. A. (1995). Maternal depressive symptoms and the risk of poor pregnancy outcome: review of the literature and preliminary findings. *Epidemiologic Reviews*, *17*, 165–171.

Paarlberg, K. M., Vingerhoets, A. J., Passchier, J., Dekkar, G. A., & Van Geijn, H. P. (1995). Psychosocial factors and pregnancy outcome: A review with emphasis on methodological issues. *Journal of Psychosomatic Research*, *39*, 563–595.

Paulson, J. F., Dauber, S., & Leiferman, J. A. (2006). Individual and combined effects of postpartum depression in mothers and fathers on parenting behavior. *Pediatrics*, *188*, 659–668.

Payne, J. L. (2003). The role of estrogen in mood disorders in women. *International Review of Psychiatry*, *15*, 280–290.
Pignone, M. P., Gaynes, B. N., & Rushton, J. L. (2002). U.S. Preventive Services Task Force. Screening for depression in adults: Summary of the evidence. *Annals of Internal Medicine*, *136*, 765–776.
Powers, Z. L., Zahorik, P., & Morrow, B. (1993). An evaluation of residents' recognition of depressive symptoms in obstetrical patients. *Journal of the Tennessee Medical Association*, *86*, 147–149.
Preti, A., Cardascia, I., Zen, T., Pellizzari, P., Marchetti, M., & Favaretto, G. (2000). Obstetric complications in patients with depression-a population based case control study. *Journal of Affective Disorders*, *61*, 101–106.
Rahman, A., Iqbal, Z., Bunn, J., Lovel, H., & Harrington, R. (2004). Impact of maternal depression on infant nutritional status and illness: a cohort study. *Archives of General Psychiatry*, *61*, 946–952.
Reid, M., Glazener, C., Murray, G. D., & Taylor, G. (2002). A two-centered pragmatic randomised controlled trial of two interventions of postnatal support. *British Journal of Obstetrics & Gynecology*, *109*, 1164–1170.
Ross, L. E., Campbell, V. L., Dennis, C. L., & Blackmore, E. R. (2006). Demographic characteristics of participants in studies of risk factors, prevention and treatment of postpartum depression. *Canadian Journal of Psychiatry*, *51*, 704–710.
Salamero, M., Marcos, T., Gutierrez, F., & Rebull, E. (1994). Factorial study of the BDI in pregnant women. *Psychological Medicine*, *24*, 1031–1035.
Saunders, S. (2006). *Closing the gap: Disparities in birth outcomes*. Illinois Departments of Healthcare and Family Services and Human Services.
Shadigian, E., & Bauer, S.T. (2005). Pregnancy-associated death: a qualitative systematic review of homicide and suicide. *Obstetrical & Gynecological Surgery*, *60*, 183–190.
Smith, M. V., Rosenheck, R. A., Cavaleri, M. A., Howell, H. B., Poschman, K., & Yonkers, K. A. (2004). Screening for and detection of depression, panic disorder, and PTSD in public-sector obstetric clinics. *Psychiatric Services*, *55*, 407–414.
Spinelli, M. G., & Endicott, J. (2003). Controlled clinical trial of interpersonal psychotherapy versus parenting education program for depressed pregnant women. *American Journal of Psychiatry*, *160*, 555–562.
Spitzer, R. L., Williams, J. B. W., Kroenke, K., Hornvak, R., & McMurry, J. (2000). Validity and utility of the PRIME-MD Patient Health Questionnaire in assessment of 3000 obstetric-gynecologic patients: The PRIME-MD Patient Health Questionnaire Obstetrics-Gynecology Study. *American Journal of Obstretrics and Gynecology*, *83*, 759–769.
Stamp, G. E., Williams, A. S., & Crowther, C. A. (1995). Evaluation of antenatal and postnatal support to overcome postnatal depression: a randomized controlled trial. *Birth*, *22*, 138–143.
Surkan, P. J., Peterson, K. E., Hughes, M. D., & Gottlieb, B. R. (2006). The role of social networks and support in postpartum women's depression: a multiethnic urban sample. *Maternal and Child Health Journal*, *10*, 375–383.
Teixeira, J. M., Fisk, N. M., & Glover, V. (1999). Association between maternal anxiety in pregnancy and increased uterine artery resistance index: cohort based study. *BMJ*, *318*, 153–157.
Viguera A. (2005, June). Poster presentation, Sixth International Conference on Bipolar Disorders. Pittsburgh, PA.
Werrett, J., & Clifford, C. (2006). Validation of the Punjabi version of the Edinburgh postnatal depression scale (EPDS). *International Journal of Nursing Studies*, *43*, 227–236.
Whiffen, V. E. (1988). Screening for postpartum depression: A methodological note. *Journal of Clinical Psychology*, *44*, 367–371.
Wickberg, B., & Hwang, C. P. (1996). Counseling of postnatal depression: A controlled study on a population based Swedish sample. *Journal of Affective Disorders*, *39*, 209–216.
Williams, J. W., Jr., Pignone, M., Ramirez, G., & Perez Stellato, C. (2002). Identifying depression in primary care: a literature synthesis of case-finding instruments. *General Hospital Psychiatry*, *24*, 225–237.
Wisner, K. L., Hanusa, B. H., & Perel, J. M. (2006). Postpartum depression: A randomized trial of sertraline versus nortriptyline. *Journal of Clinical Psychopharmacology*, *26*, 353–360.
Wolman, W., Chalmers, B., Hofmeyr, J., & Nikodem, V. C. (1993). Postpartum depression and companionship in the clinical birth environment: a randomized, controlled study. *American Journal of Obstetrics & Gynecology*, *168*, 1388–1393.
Zax, M., Sameroff, A. J., & Babigian, H. M. (1977). Birth outcome in the offspring of mentally disordered women. *American Journal of Orthopsychiatry*, *47*, 218–230.
Zuckerman, B., Bauchner, H., Parker, S., & Cabral, H. (1990). Maternal depressive symptoms during pregnancy, and newborn irritability. *Journal of Developmental & Behavioral Pediatrics*, *11*, 190–194.

Chapter 14
Supplemental Nutrition Programs During Pregnancy and the Early Postnatal Period

Noel Chávez

Introduction

Pregnancy has been recognized as a period of special nutritional needs across cultures and for many generations. Mothers who have adequate food and good quality nutritional intake generally have better birth outcomes. In cultures where food is scarce, pregnant women may be given extra servings of high quality protein items such as eggs or milk or other scarce but nutritious foods when there isn't sufficient supply for everyone to consume them. However, formal programs to improve women's nutrition during pregnancy as a strategy to improve birth outcomes are relatively recent. This is particularly the case in industrialized countries such as the United States where chronic undernutrition is not endemic. This chapter examines the evidence base for programs that seek to improve birth outcomes through improving pregnant women's nutritional intake and nutrition education. It begins with a brief overview of the role of nutrition in pregnancy, followed by a discussion of the foundation for WIC, the major U.S. prenatal nutrition intervention. A critical evaluation is then presented of the evidence for WIC and selected other programs in reducing low birth weight (LBW) and in reducing racial/ethnic disparities in birth outcomes. The chapter concludes with a discussion of the relevance of the findings for practitioners along with suggestions for future research.

Nutrition in Pregnancy Overview

Good nutrition is essential for all women, but particularly during pregnancy. King (2003) states that "adequate availability of nutrients during gestation is probably the single most important environmental factor influencing pregnancy outcome" (p. 1732S). While a comprehensive review of maternal nutrition and physiology cannot be provided here, there are several excellent reviews that thoroughly cover maternal nutrition and its role in birth outcomes (e.g. King, 2003; Rush, 2001a; 2001b). In brief, though, the fetus receives all of the nutriture for growth and development through the placenta. The fetus has no source of energy substrates, protein, vitamins or minerals other than the mother. Many factors besides maternal food and nutrient intake influence fetal nutrition and growth, including maternal absorption, metabolism, endocrine function and nutrient/energy transfer to the placenta (Fall et al., 2003). Adjustments in nutrient use and metabolism occur during pregnancy, but may be altered if the woman does not have good nutritional status. Differential partitioning of

N. Chávez (✉)
Maternal Child Health Program, Community Health Sciences Division, University of Illinois at Chicago School of Public Health, 1603 W. Taylor MC923, Chicago, IL 60612, USA
e-mail: nchavez@uic.edu

A. Handler et al. (eds.), *Reducing Racial/Ethnic Disparities in Reproductive and Perinatal Outcomes: The Evidence from Population-Based Interventions*, DOI 10.1007/978-1-4419-1499-6_14,

nutrients between mothers and offspring are also highly dependent on the pre-pregnancy nutritional status of the mother; it appears that fetal needs take precedence in severely depleted women, while nutrients are preferentially deposited in maternal tissue in marginally depleted women (King, 2003). There are several possible consequences when women have less than adequate energy and nutrient intakes. A mother's failure to gain sufficient weight during the pregnancy can retard fetal growth or development and may result in or contribute to poor outcomes such as LBW, preterm birth or infant mortality. Fall et al. (2003) note that with an inadequate nutrient supply, the fetus adapts through growth down-regulation and essential tissue development prioritization. Poor nutrition during pregnancy also can contribute to problems with lactation or depleted nutritional status postpartum. Maternal nutritional status is of particular importance for pregnancies occurring at a young maternal age (within 2 years of menarche), and for women with short interpregnancy intervals, namely less than 18 months (King, 2003). King notes that the "risk of LBW or preterm birth among women with early or closely spaced pregnancies in the United States is at least 50% greater than that of adult women with a interpregnancy interval of 18–23 months" (2003, p. 1733S).

The adverse effects of inadequate and/or inappropriate maternal nutrition include LBW, a greater risk of preterm birth with its accompanying potential complications, and infants who are poorly developed for their gestational age (small for gestational age, SGA) or who show signs of IUGR (intrauterine growth restriction). LBW infants can be classified as those who are born preterm (before 37 weeks of gestation) or who are SGA, an indicator of growth restriction. Most research defines LBW as less than 2,500g, and very LBW as less than 1,500g. While nearly all infants born before 37 weeks gestation are LBW, a birth weight of <2,500g can be observed in infants greater than 37 weeks gestation; these infants are usually described as IUGR or SGA. SGA infants are usually identified by comparing their birth weight to national growth norms, and typically reflect those full term infants below the 10th percentile (Kotch, 2005). In looking at LBW it is important to note that we are examining the continuum of infant weight at the point of birth. While examination of the effect of maternal nutrition on the IUGR or SGA outcomes would be preferred, these measures are limited in that they require an accurate gestational age, which many studies are unable to obtain; taking the infant weight at birth is a standard practice and less subject to error than these other measures.

The role of nutrition in healthy birth outcomes is supported by a large volume of observational data and these serve as the "theoretical basis" for many maternal nutrition interventions. Few studies actually present a specific theoretical basis for maternal nutrition interventions, and few conceptual frameworks are found in the literature. One conceptual framework was proposed by Sweeney et al. (1985) and is presented in Fig.14.1; it shows the physiologic relationships between maternal nutrition and birth outcome. This framework provides a detailed, although hypothetical, view of how

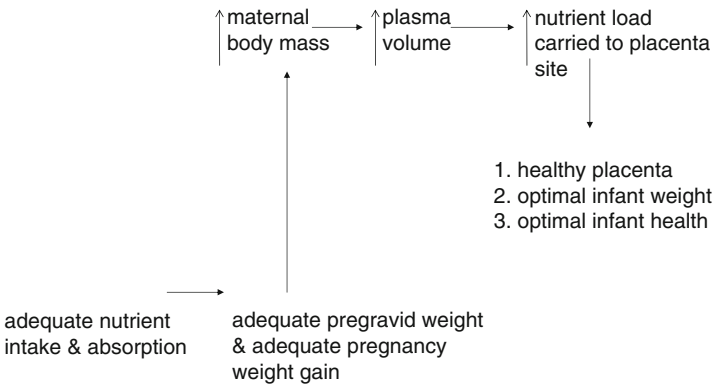


Fig. 14.1 Hypothesized relationships of effects of nutrition on infant outcome (Sweeney, et al., 1985)

maternal physiological function is affected by maternal nutrition. It is important to note that the relationship between maternal nutrition and birth outcome may be nonlinear and there are many other factors (e.g., malaria, smoking, primiparity, pregnancy-induced hypertension, congenital anomalies, and/or other genetic factors) which can influence birth outcomes such as birth weight and other outcomes (Brown & Khan, 1997; Kramer, 2003; Susser, 1991).

While the evidence is strong for an effect of maternal nutrition on infant birth weight overall, birth weight itself is not a highly sensitive indicator of nutrition effects. In fact, as noted by Brown & Khan (1997), birth weight alone is unlikely to differentiate between constitutionally small infants and those whose growth and development are compromised *in utero*. Echoing the "less straightforward" relationships noted by Susser & Stein (1982), Brown and Khan suggest that the relationships between a mother's nutritional exposure and outcomes may be non-linear, and recent research has shown these relationships to be far more complex than reported previously. Thus, given that no single indicator of maternal nutrition can capture the multiple effects of various nutrients and other nutritional exposures on the outcomes of pregnancy, the successful process of fetal growth and development may be better captured by indicators such as proportionate growth, inflexible physiologic responses and chronic disease risk, rather than birth weight (Brown & Khan, 1997). Kramer (2003) noted similarly that most of the research on adverse pregnancy outcomes and nutrition was based on proxy outcomes of mortality and severe morbidity, with LBW being the most common. Despite the problems with birth weight as a maternal nutrition indicator, a consensus reached by the Institute of Medicine (IOM) was that infant birth weight, even with its limitations, is probably the best *general* indicator of the role of maternal nutrition in pregnancy outcome (1985). Thus, birth weight was selected as the primary nutrition outcome for this chapter. Another consideration for selecting birth weight as the pregnancy outcome indicator was that it was the outcome reported most consistently across the prenatal nutrition supplementation literature.

Prevention of LBW through adequate nutrition and nutrition education is a key objective in the 2010 Healthy People document [U.S. Department of Health and Human Services (DHHS), 2000]. From a population perspective, LBW is more frequently observed in low income women, with significant disparities between racial and ethnic groups in developed countries such as the United States. From a nutritional perspective, the best scenario for a healthy baby is to have a well nourished mother who is at an appropriate weight for height with good nutrient stores at conception, who has an appropriate amount and patterning of pregnancy weight gain and who has an energy and nutrient intake during pregnancy that is adequate to supply both the mother and the fetus.

International Studies of Maternal Nutrition and Birth Outcomes

Some of the strongest evidence of the role of nutrition in birth outcomes comes from less developed countries or during famine in industrialized countries. IUGR seems to be the more frequent problem with levels as high as 40% in some developing countries (de Onis, Villar, & Gülmezoglu, 1998). There is a higher risk of poor maternal nutrition in less developed countries, and a large number of studies have demonstrated that poor maternal nutrition, both in terms of total energy intake and intakes of specific nutrients, is associated with higher risk of LBW infants (Rush, 2001a; Susser, 1991). A recent Cochrane Review (Kramer & Kakuma, 2006) reported that among five trials of dietary advice to pregnant women and 13 trials providing balanced energy/protein supplements there were "modest increases in maternal weight gain and in mean birth weight, and a substantial reduction in risk of small-for-gestational-age (SGA) birth" (p. 1). Trials providing supplemental food produced higher birth weight increases than those with nutrition advice alone. While famine, starvation and lack of food reduce fertility and birth weight, there are also adverse effects from less severe nutrition insults (Rush, 2001a).

A frequently cited study was conducted by Ceesay et al. (1997); their research with Gambian women found that supplementation, after adjusting for sex, parity and Parkin score (to assess gestational age), resulted in both increased birth weight (136g, $p<0.001$) and head circumference (3.1mm, $p<0.001$) in the infants of supplemented mothers. Ceesay et al., noted that the supplementation reduced the incidence of "small for their dates" infants but did not decrease prematurity. A recent trial in Nepal (Osrin et al., 2005) showed that micronutrient supplementation reduced the proportion of LBW infants by 25% (mean difference 77g) in the intervention group, although no difference was noted in gestational duration.

While results of single micronutrient studies aren't always conclusive, there seems to be a relationship between maternal iron deficiency and LBW (Rasmussen, 2001). Rush (2001a) reports that there is a 51 g advantage in birth weight for male infants of poor Colombian women who are supplemented during the third trimester. Some studies have supplemented diets with protein. Rush (2001a) notes a 36–83 g birth weight increase in an Indian study of protein-energy supplementation at two levels. There are other relatively small studies of maternal supplementation, weight gain and birth outcome that have been conducted in Central and South America, Africa and Asia. While some of these have design problems, Rush (2001a) concludes that the evidence for supplementation of low and moderate protein density foods during pregnancy to raise birth weights to a modest degree is reasonable and consistent, although studies with larger participant numbers and the strongest designs only observed increments in birth weight in the range of 20–50 g, considerably smaller than had been expected. In describing supplementation programs from both less developed and industrialized countries, Rush (2001b) noted that the birth outcomes were relatively similar, and urged that further interventions and research focus on the subsets of women most likely to benefit from supplementation. The cumulative evidence from both international and more local studies of the role of maternal nutrition has formed the basis of nutritional interventions in industrialized countries such as the United States and Canada.

History of Prenatal Nutrition Programs in the U.S. and Canada

In industrialized countries, the role of maternal nutrition is different from that in developing countries. While problems with food supply and integrity do not typically occur throughout the year, there are some groups at greater risk of having inadequate food intake or food insecurity. For example, lower income women and ethnic/racial minority group status have long been associated with poorer birth outcomes, although there is not regular nutritional status surveillance of U.S. women to document an association of these poor outcomes with maternal undernutrition.

Funding for improving the nutrition of pregnant women (and children) in the U.S. began with the Sheppard-Towner Act in 1921, creating the first federal-state partnership for maternal and child health services and providing states with grants-in-aid for such services (Egan, 1994; Rosen, 1985). Several states used these funds to support nutrition services, including prenatal nutrition. Additional authorization and funding for prenatal nutrition through the Children's Bureau emerged from Title V of the Social Security Act (1935) which required public health nutritionists as staff in state health departments (Dodds & Kaufman, 1991).

The first national U.S. program, the Special Supplemental Program for Women, Infants and Children (WIC), was *authorized* as a pilot program by Congress through the U.S. Department of Agriculture (USDA) in 1972 (the first clinic did not open until 1974). The program emerged from the realization that women, infants, and children of low income and those living in poverty were predisposed to poor nutrition and poor health status, due to inadequate nutrition and health care (Owen & Owen, 1997). Much of the impetus for WIC came from the 1969 White House Conference on Food, Nutrition and Health which acknowledged fairly widespread hunger among

vulnerable populations, including pregnant women, and that hunger contributed to poor birth outcomes (White House Conference, 1969). According to Egan (1994), establishing the WIC program within USDA was the product of a 10-year effort by public health nutritionists to improve food assistance to high risk groups, such as pregnant Women, Infants and Children (WIC). One influence on developing a food assistance program came from an existing supplemental food intervention in Montreal.

The Higgins Nutrition Intervention Program (HNIP) for high risk pregnant women was developed in 1963 by Agnes Higgins, a dietitian at the Montreal Diet Dispensary. The Higgins program included assessment of individual nutritional risk followed by tailored nutrition counseling and provision of supplemental foods to address nutritional risk factors. Initial steps with this intervention began as early as 1948, but were formalized by Ms Higgins in the 1960s. Her group tracked birth outcomes and various reports and papers were published as a result of her work; these reports showed that the combination of counseling and supplemental foods improved infant birth weight in the infants of women at nutritional risk (Higgins, 1976; Higgins, Crampton, & Moxley, 1972). Studies based on this program will be presented and discussed as part of the evidence assessment for this chapter.

Agnes Higgins also conducted training courses on the Higgins Method for health professionals, including those from the U.S. These three week intensive courses, sponsored by the March of Dimes, provided training to 135 U.S. Health professionals (Maternal Nutrition and Agnes Higgins Analysis, n.d.). The work of Ms. Higgins and the Montreal Diet Dispensary were influential as USDA began planning the Special Supplemental Nutrition Program for WIC in the early 1970s (M. Lavan, March of Dimes National Office, personal communication, 2007). Risk assessment, provision of supplemental foods, nutrition education and referral to other health and social services applied on a population level became the hallmark of prenatal nutrition intervention in the U.S. beginning in 1972 (The WIC program was authorized as a pilot program by Congress in 1972 with an amendment to the 1966 Child Nutrition Act; the first clinic opened and certified participants in 1974, and WIC was permanently funded in 1975 as a health and nutrition program in P.L. 94–105) (Oliveira, Racine, Olmsted, & Ghelfi, 2002). USDA has been involved with food programs since the early 1930s when there were many people without adequate food, but the impetus behind the programs during the Depression was to provide financial assistance to farmers. Eventually, surplus commodity foods (excess agricultural production) were distributed to needy populations and continue to be an important aspect of USDA food programs (School Lunch and Breakfast Programs, Summer Food Service Program, Child and Adult Care Food Program and the Nutrition Service Incentive Program). While there was a form of food stamps used during the Depression, the framework for the current Food Stamp Program comes from the 1964 Food Stamp Act, with national eligibility standards established in 1971. Participants receive a monthly allotment of "stamps" (currently electronic benefit transfer cards) which can be used in retail food stores to purchase foods; merchants are subsequently reimbursed in the amount redeemed. Eligibility is tied to family income and assets. Historically nutrition has not been a component of the Food Stamp Program although there is currently a nutrition education program within the overall Food Stamp Program (Note that in September, 2008, the name for the Food Stamp Program was changed to SNAP, Supplemental Nutritional Assistance Program.) (USDA, 2007a, 2009a).

The Current U.S. Prenatal Nutrition Program – WIC

The WIC Program has grown considerably since certifying the first WIC participants in 1974. WIC is administered through the 90 WIC state agencies which operate through 2000 local agencies providing services for low income pregnant, postpartum and lactating women, infants, and children

up to age five (USDA, 2005). In 2006, the most recent year for which participant and program characteristic information are available, the federal budget for WIC was $5.01 billion and the program served 8.8 million participants, 11% of whom were pregnant women, 7% postpartum women and 7% breastfeeding women (USDA, 2007b). That same year, the WIC income threshold was $30,710 for a family of three, the average annualized family income of WIC participants was $15,577, and about 67.4% of WIC participants had incomes at or below 100% of the federal poverty level, compared with 12.6% of the general population (USDA, 2007b). WIC participants receive nutritional risk assessments, supplemental foods, nutrition education/counseling, and referrals to other health and social services.

Unlike the Food Stamp Program, an entitlement program with a general nutrition education component but without nutritional risk requirements that is also sponsored by USDA, WIC is a grant program that requires its participants to meet nutritional risk criteria as well as have a gross household income of less than 185% of the federal poverty level. Congress does not set aside sufficient funds each year to allow every eligible individual to participate in the WIC program. Instead, it annually authorizes a specific amount of WIC funds to USDA which are then distributed as grants to the state agencies. Once the authorized funds are used and allocated caseloads are full, eligible women, infants, and children are put onto waiting list by nutrition risk priority (USDA, 2005). Because of the program focus on healthy birth outcomes, pregnant women have a high priority when funding is limited; currently funding allows most eligible pregnant women seeking services to be served (IOM, 2003). It should be noted that women at nutritional risk can participate simultaneously in both the Food Stamp and WIC programs. Given that there is little USDA research on dietary intake or nutritional status changes in pregnant women, it should be noted that for pregnant participants in both Food Stamps and WIC, joint participation would be a significant confounder in an analysis of either program's effects. Income eligibility for Food Stamps automatically meets the income criterion for WIC; similarly, pregnant Medicaid recipients can be "adjunct income-qualified" for the WIC program. Nevertheless, both Food Stamp and Medicaid recipients must still meet the nutritional risk criteria.

For pregnant women, WIC nutritional risk criteria include low hemoglobin or hematocrit, underweight or overweight, adolescence, history of preterm delivery, LBW, a neonatal loss, and poor nutritional intake. While some of these risk parameters are fairly easily and objectively measured, nutritional intake is generally self-reported via a 24-hour recall or food frequency questionnaire, and thus subject to self-report bias. After the initial nutrition assessment screening, eligible women receive regular nutrition education and counseling from nutrition professionals, along with vouchers, coupons, or an EBT (electronic benefits transfer) card, all of which can be used for purchase of specific foods in a neighborhood grocery store or in some cases a WIC food distribution center. Foods in the WIC package are selected to assure intake of particular nutrients, protein, iron, vitamin C, vitamin A, and folic acid. Protein, iron, calcium, vitamin C, and vitamin A were nutrients "known to be lacking in the diets of populations at nutritional risk" and were the focus of supplementation early in the program; folic acid was added later as evidence emerged of its role in healthy pregnancy outcome (Oliveira et al., 2002, p. 8). The current food package for pregnant women consists of milk, iron- and folic acid-fortified dry cereal, dry beans, cheese, eggs, tuna, peanut butter, carrots and fruit or vegetable juice high in vitamin C [Food Research and Action Center (FRAC), 2005]. Women are also referred to other health or social services as needed.

With this brief background on nutrition program during pregnancy and infant birth weight, along with a description of the WIC program and its participants in hand, the methods used for an examination of the evidence for the effect of prenatal nutrition interventions on LBW is described in what follows.

Methods

The following search terms were used in the PubMed, Web of Science and CINHAL databases to identify and locate relevant peer reviewed articles and reports published between 1985 and the 2008: "maternal nutrition," "prenatal nutrition," "WIC," "supplemental nutrition," "LBW", "birth outcomes," "nutrition interventions," "Montreal Diet Dispensary," and "Higgins Method." In addition to these terms, Cochrane Reviews (Cochrane Database of Systematic Reviews) covering nutrition and pregnancy were examined. Citations in comprehensive reviews and meta-analyses were also reviewed to identify additional relevant articles. The following websites were searched for relevant reports and program evaluations: Food and Nutrition Service, USDA, Food Research and Action Center, National WIC Association, and the Center for Budget and Policy Priorities. Once located, articles were reviewed by the author; those classified as meeting the selection criteria of a program providing a prenatal food supplement and/or nutrition education and counseling were reviewed and scored using the Downs and Black criteria (1998). Publications included in the analysis were those reporting on or using data from a program/intervention that included supplemental foods and/or nutrition education that were found in a peer reviewed article or agency report published between 1985 and February, 2008. While other health and social services might be included in the WIC program, their effects were not reviewed for this chapter. The primary outcome examined in each study was LBW. Aside from some of the meta-analyses which usually incorporated international data, programs analyzed were from the United States and Canada. The overall quality and strengths of the examined studies are presented in Table 14.1. Specific individual study data are in Table 14.2 and findings of meta-analyses are in Table 14.3. They are organized by study source in presenting a summary of the evidence for nutrition interventions to improve infant birth weight.

Overall Assessment of Study Quality and Strength

Table 14.1 shows that the general study quality and strength of the selected articles was "fair," with a number of scores in the "poor" range. Scores were lower than might have been expected due to a variety of factors. There have been a very limited number of randomized clinical trials of nutrition interventions in the U.S.; most of the studies presented are observational. There were also a large number of secondary data analyses based on WIC program and/or Medicaid data. While useful for the review, these analyses rarely provided all of the information needed to yield a high quality score. In some cases it is possible that the scoring criteria were indeed met by a study, but the needed explanations/descriptions were not actually stated in the articles. In addition, many of the studies with primary data were relatively small, and these were more likely to have research design problems, and thus lower scores. Only three studies reported whether there was sufficient statistical power in examining the results. Scores in the other categories were generally several points below the maximum possible for the category, reflecting the relatively low strength of the studies to assess program effectiveness. In a related analysis of interventions to reduce maternal morbidity and mortality and preterm birth, Villar et al. (2003) note that a key element in a systematic review is a "critical evaluation of all primary studies *answering the same question*" (emphasis added) (p. 1609S). A review of the study objectives for the articles selected and discussed in this chapter, showed at least six different primary objectives or study purposes; in the majority of studies at least one of the objectives was to examine the relationship between prenatal WIC participation and birth outcomes. Along with differing study objectives, there were also differences in the study outcomes that were measured and reported, such as: mean LBW, LBW rate, and IUGR. A related problem was that

Table 14.1 Quality rating of studies associated with prenatal supplemental nutrition interventions

Author	Reporting (14)	External validity (4)	Internal validity-bias (9)	Internal validity-confounding (7)	Power (2)	Total quality score ≤14 = poor, 15–19 = fair, ≥20 = good	Suitability of study to assess effectiveness: greatest, moderate, least	Brief study descriptors
Health status outcome: infant birth weight, including low birth weight, very low birth weight								
WIC studies of Medicaid participants								
Schramm (1985)	7	1	3	3	0	13	Moderate	MO Medicaid/WIC, cost data
Stockbauer (1986)	8	1	6	3	0	18	Moderate	MO WIC
Stockbauer (1987)	8	1	4	3	0	16	Moderate	MO WIC
Devaney et al. (1990, 1991)	7	1	4	3	0	15	Moderate	Medicaid-WIC analysis in five states
Devaney (1992)	6	1	5	3	0	14	Moderate	VLBW, five states, Medicaid-WIC
Buescher and Horton, (2000)	10	2	5	4	0	21	Moderate	NC-Medicaid-WIC and medical costs, 1988 data
Ahluwalia et al. (1998)	9	1	4	3	0	17	Moderate	MI SGA births, 1992 data
Buescher et al. (2000)	9	1	5	4	0	19	Moderate	NC Medicaid-WIC and medical costs, 1997 data
Joyce et al. (2005, 2004)	9	3	4	4	0	20	Moderate	NYC '88–'01 WIC-Medicaid – same analysis, but one published, other a working paper
Bitler & Currie (2005)	8	3	5	4	0	20	Moderate	PRAMS-Medicaid data, 19 states
WIC participants – non-Medicaid								
Collins et al. (1985)	5	2	2	2	2	14	Least	Small Alabama study WIC and non-WIC
Metcoff et al. (1985)	10	3	2	5	0	20	Greatest	Randomized, OK
Caan et al. (1987)	8	3	5	4	0	20	Moderate	CA WIC, interpregnancy interval

Rush et al. (1988a)	8	1	3	3	0	15	Moderate	National WIC evaluation-historical study
Rush et al. (1988b)	9	2	4	4	1	20	Moderate	National WIC evaluation-longitudinal study
Gordon and Nelson (1995)	8	3	4	4	0	19	Moderate	1988 National Maternal and Infant Health Study
Frisbie et al. (1997)	8	1	3	4	0	16	Moderate	National Maternal and Infant Health Study
Sloan et al. (2001)	9	1	4	2	0	16	Moderate	National WIC Evaluation data; protein intake and BW
Kowaleski-Jones & Duncan (2002)	9	0	5	4	1	19	Moderate	National Longitudinal Survey of Youth
Reichman & Teitler (2003)	8	1	4	2	0	15	Moderate	NJ HealthStart
Lazariu-Bauer et al. (2004)	8	2	2	4	0	16	Moderate	NY state, early vs. late WIC participation
Joyce et al. (2008)	9	2	4	3	0	18	Moderate	Pregnancy Nutrition Surveillance System data, nine states
Small studies of supplemental nutrition								
Wigda & Lewis (1999)	9	3	3	5	0	20	Least	Small in-home intervention, low income, no WIC
Briley et al. (2002)	6	0	3	3	0	12	Least	In-home, very small, randomized
Long et al. (2002)	7	0	2	2	0	11	Least	Adolescent, small program, nutrition education, not supplemental foods, but intervention on WIC also
Canadian nutrition supplementation studies								
Higgins et al. (1989)	9	2	5	2	0	18	Greatest	Canadian-Higgins, within-mother, two births
Dubois et al. (1997)	9	0	4	4	0	17	Greatest	Canadian-Higgins, teens, cohort study
Desjardins & Hardwick (1999)	10	2	5	5	0	22	Moderate	Healthiest Babies Possible Ontario

Table 14.2 Major outcomes associated with studies of prenatal supplemental nutrition interventions

Author	Study design	Study type	Intervention description what, how and where	Populations Studied (ages, race/ethnicity) and sample size
Health status outcome: infant birth weight *Studies of WIC and Medicaid participants*				
Schramm (1985)	Retrospective	Published article (secondary data analysis)	WIC participation within Medicaid participants in MO; cost analysis (cost information not presented here)	1980 Medicaid births; 7,628 records linked for infant, mom Medicaid, birth certificate and WIC participation
Stockbauer (1986)	Retrospective, quasi-experimental	Published article (secondary data)	WIC prenatal participants in MO for 1980 (10/79–6/81); linked participant information for mom and infant matched with vital records. Also looked at WIC participation as measured by time and coupons redeemed	Able to match 6,732 births (and account for losses), deducted fetal deaths, twins, etc. for group varying in size between 5,574 and 6,657 depending on comparison group
Stockbauer (1987)	Retrospective, quasi-experimental	Published article (secondary data)	MO, linked WIC and vital stats records, 1982; WIC was 13% of all births	9,411 WIC linked data; includes 9,307 live births, 104 fetal deaths, 137 WIC infant deaths. Control population matched on maternal age, education, marital status, gravidity, plurality, child race; 99.1% match with WIC
Devaney et al. (1990, 1991)	Retrospective Medicaid WIC vs. non-WIC, quasi-experimental	Report (secondary data)	All Medicaid births: 1987: FL, MN, NC, SC; Jan–June 1988: TX	~105,000 births; WIC participation varied from 48% of all Medicaid births in TX to 73% in SC; white, black, Hispanic
Devaney (1992)	Retrospective Medicaid WIC vs. non-WIC, quasi-experimental	Report (secondary data)	All Medicaid births: 1987: FL, MN, NC, SC; Jan–June 1988: TX	Same data set as Devaney '90–'91, but only VLBW; looked at WIC participants before 32 weeks and before 30 weeks

Address disparities (yes/no)	Key findings related to intervention effectiveness	Caveats/biases	Findings support the intervention? yes/no for which populations?
Yes, white vs. non-white	WIC births 6 g heavier than non-WIC; LBW rate 2% less. LBW rates lower for WIC than non-WIC for almost all variables tested. Additional findings: Fewer WIC infants admitted to NICUs. Found 40–45% of newborn cost differential explained by WIC/non-WIC birth weight distribution. As WIC costs increased Medicaid costs decreased	Controlled for a number of variables in the analysis, but some confounders unknown	Yes
Compared white vs. non-white	WIC participants had lower percent LBW than non-WIC for all 3 methods, and was significant for standardization method (9.6%) – primarily reflects non-white WIC differences Additional findings: Mean gestational age greater in WIC. Non-white WIC had lower percent of SGA infants (7.8 vs. 9.2) and lower % births before 37 weeks	Acknowledges selection bias because no a priori control/ comparison group. Used three methods to construct post hoc comparison groups	Yes, but effects are small
Yes	Greatest difference between WIC and non-WIC were pre-pregnancy weight, smoking during pregnancy, urban/rural. WIC participation associated with 16% decrease in LBW (22% for blacks vs. 10% for whites) and 27% decrease in <1,500 g; effects greater for blacks for all outcomes	Some potential confounders not available for adjustment. Accuracy issues with birth certificate data. Potential selection bias into WIC	Yes
Yes	Increased BW in WIC in all states; greater BW increase for WIC vs. non-WIC with preterm births <37 weeks Additional findings: Medicaid cost savings for WIC infants; lower preterm birth incidence and longer gestational age in WIC participants	Potential selection bias into WIC; differential WIC participation across states not accounted for; greater differences in state level Medicaid policy, social and demographic factors	Yes
Yes	Lower VLBW in WIC than non-WIC Medicaid, regardless of level of WIC participation; non-WIC moms 2–3X as likely to have VLBW as participants; regardless of WIC participation, VLBW varied by race with blacks higher rates and Hispanics lower; significant differences across states	Biases similar as in Devaney '90–'91	Yes

(continued)

Table 14.2 (continued)

Author	Study design	Study type	Intervention description what, how and where	Populations Studied (ages, race/ethnicity) and sample size
Buescher et al. (1993)	Retrospective, Medicaid women, WIC vs. non-WIC, quasi-experimental	Published article (secondary data)	NC, 1988; WIC/Medicaid	All Medicaid births, black; white; <18 vs. 18+ years; 23,343
Aluwalia et al. (1998)	Retrospective, Medicaid women, WIC vs. non-WIC, quasi-experimental	Published article (secondary data)	WIC/Medicaid MI, 1992	All ages; white, black, Native American, other; 53,782
Buescher et al. (2000)	Retrospective, Medicaid WIC vs. non-WIC, quasi-experimental	Report (secondary data)	NC, 1997, WIC/Medicaid	All Medicaid births; white or minority; 43,276; <18 vs. 18+ years
Joyce et al. (2004, 2005)	Retrospective, Medicaid, WIC vs. non-WIC, quasi-experimental	Published article (secondary data)	NYC, 1988–2001; all Medicaid births, WIC vs. non-WIC; examined variables over time-3 periods: '88–'92,' 93–'97,'98–'01	Singleton and separate twin analysis; first births; 56,002 '88–92; 75,998 '93–'97; 64,656 '09–'01. Mom's age 10–50 year
Bitler and Currie (2005)	Retrospective, Medicaid women with or without WIC, quasi-experimental	Published article (secondary data)	WIC/Medicaid; PRAMS data for 19 states	All ages; white, black, Hispanic; 60,731
WIC studies of non-Medicaid participants				
Collins et al. (1985)	Prospective, quasi-experimental	Published article (primary data with secondary from WIC, hospital and vital records)	WIC women at nutritional risk and low income women not at nutritional risk in same AL Appalachia health department prenatal clinics	Rural women; 341 WIC and 178 non-WIC "controls." 38% white and 62% black; ages 10–30+

Address disparities (yes/no)	Key findings related to intervention effectiveness	Caveats/biases	Findings support the intervention? yes/no for which populations?
Yes	WIC enrollment associated with significantly decreased rates of LBW and VLBW ($p<0.001$); stronger association within black Medicaid population WIC participation fourth most important predictor for LBW; consistent decrease in LBW, VLBW and newborn medical costs with increased WIC participation	Controlled for length. Since one WIC visit counted as participation, results may underestimate WIC benefits	Yes
Yes; adjusted for ethnic, but no separate analysis	Higher BW with longer WIC participation; after adjustment, OR findings held but attenuated	Limited to MI women on Medicaid	Yes
Yes	WIC participants had significantly lower LBW and VLBW rates; WIC participation associated with additional 61g in BW after controlling for risk factors	Gestational age bias may be an issue; controlled for measurable differences in WIC vs. non-WIC, but may still be unmeasurable differences	Yes
PR, other Hispanic, Asian, black, white, other	Unadjusted difference between WIC and non-WIC is 44 g over 14 years of study, but after adjustment, this falls to 25.5g. With those entering PNC earlier, adjusted mean differences decreased to 9.8 g. Only modest association between WIC and BW adjusted for gestational age. Little association between WIC and fetal growth. Modest effect of WIC for black participants (LBW rate for participants is 2.4% less than for non-WIC participants in '88–'92)	WIC participation is self-report. Possible selection bias	Yes, but the authors suggest that WIC only has a minimal effect on poor birth outcomes
Yes, WIC effect analyzed in varying levels of being "disadvantaged"	WIC participants less likely to have low birth weight infant. Additional findings: WIC associated with earlier prenatal care, shorter hospital stay, fewer NICU admits, greater pregnancy weight gain, less likely to breastfeed	Accounted for a large number of observable characteristics, but still some likely unobservable. However, positive selection unlikely since improvements were greatest for most disadvantaged	Yes
Yes	Mean BW same for each group, but more LBW infants in WIC (6.4 vs. 3.8%). BW here below national average (at that time), and below that for AL. Additional findings: Believe that WIC program helps to offset nutritional risks	Limited analyses-primarily descriptive. Two groups were not comparable given importance of nutritional risk	Yes

(continued)

Table 14.2 (continued)

Author	Study design	Study type	Intervention description what, how and where	Populations Studied (ages, race/ethnicity) and sample size
Metcoff et al. (1985)	Prospective, randomized trial with wait list controls	Published article (primary data)	OK. WIC, WIC-eligible on wait list	683 at one hospital, 172 WIC and 238 control babies with complete data
Caan et al. (1987)	Prospective, WIC post-partum >5 months vs. <2 months	Published article (mostly primary, but some secondary)	Women on WIC during first and second pregnancy where interpregnancy interval <27 months; compared limited (<2 months) vs. extended >5 months) post partum WIC	642 women selected from 48 California WIC sites
Rush et al. (1988a) III. Historical Study (National WIC Evaluation)	Retrospective, ecological level analysis, time series	Published article (secondary data)	WIC '72–'81; 888 counties, 14 states and DC	Only black and white singletons used for subgroups stratified by maternal education and race; multiples and "other" races included for other analyses
Rush et al. (1988b) National WIC Evaluation V. Longitudinal Study	Prospective, quasi-experimental	Published article (primary data)	WIC, non-WIC participants, data collected 1983–1984; national data from 174 WIC clinics in 58 contiguous US areas for participants and 55 prenatal clinics without WIC	White, Hispanic, black and other women 5,205 WIC and 1,358 non-WIC
Gordon and Nelson (1995)	Retrospective, WIC, income eligible non-WIC and higher income non-participants; quasi-experimental	Report (secondary data)	National Maternal and Infant Health Survey, 1988	Weighted, nationally representative sample, white and black, 3,868 WIC, 2,303 income eligible non-WIC, 3,783 higher income. Includes both Medicaid and non-Medicaid

Address disparities (yes/no)	Key findings related to intervention effectiveness	Caveats/biases	Findings support the intervention? yes/no for which populations?
74% white, 21% black, 1% NA, 4% other	Positive increase in BW with WIC: 3,254 g (WIC) vs. 3,163 g (non-WIC), 91 g difference, $p=0.0395$, adjusted for sex, gestational age, prenatal visits, pregnancy interval, smoking and previous LBW infant; when adjusted for entry weight of mom, effect of WIC on BW not significant, but WIC had positive effect on infant birthweight of smokers ($p=0.017$)	Adjusted for various confounders; selection bias minimized in that all women qualified for WIC	Yes
Measured, yes	After adjusting for race, parity, BMI, BW of first infant, weight gain, infant sex, smoking, extended WIC associated with significantly greater BW (adjusted for gestational age); OR (limited vs. extended WIC) of LBW was 2.5 (95% CI 0.89–7.00); except for Native Americans, effects about same across race/ethnic groups; BW increases greatest for underweight women at start of second pregnancy	Some potential bias due to selection of WIC agencies	Modestly yes, but stronger for women underweight at start of 2nd pregnancy
Yes	WIC impact on BW was 22.7 g/participant served ($p<0.01$ before and $p=0.07$ after accounting for autocorrelation); effects for blacks and whites differed when stratified by education. Equivocal results of WIC on rate of LBW, with significance only among less educated blacks (−1.49%; $p<0.05$)	Adjusted for secular trend during years analyzed; adjusted for autocorrelation. Results probably underestimate magnitude and significance of WIC impact	Yes
Yes	No significant effect of WIC participation on mean gestational age. No significant relationship between WIC and mean BW or LBW frequency, with or without adjustment for gestational age. However, when analysis was conducted by intention to treat, some effects (not significant) were seen Additional findings: There was a significant relationship between WIC participation and infant head circumference	Outcomes adjusted for maternal ethnicity, gestation duration, other maternal and family characteristics; 25% of control women were in WIC at follow up interview; control group more socially advantaged than WIC group	Yes
Yes	WIC participation associated with 67.9 g increase in birth weight ($p=0.01$); LBW (WIC 7.9% vs. no WIC 10.8%) and VLBW (WIC 1.2% vs. no WIC 2.2%) significant; effect of WIC on birth weight does not differ between blacks and whites Additional findings: WIC participation associated with longer gestational age (approximately ½ week longer); preterm birth (WIC 10.6% vs. no WIC 14.2%) (Results from the basic models. Probabilities predicted using weighted logit models.)	Various issues with the comparison groups, especially since there were notable differences between WIC and income-eligible non-WIC	Yes

(continued)

Table 14.2 (continued)

Author	Study design	Study type	Intervention description what, how and where	Populations Studied (ages, race/ethnicity) and sample size
Frisbie et al. (1997)	Retrospective WIC vs. non-WIC – IUGR including LBW	Published article (secondary data)	National Maternal and Infant Health Survey, 1988; compared WIC vs. non-WIC; Oversample by race (white, AA, MA) and BW	Births over 500 g, 22–50 weeks gestation included; 8,424 total, but does not give numbers for ethnic groups
Sloan et al. (2001)	Retrospective, cohort	Published article (secondary data analysis)	Used 1988 National WIC evaluation data; WIC and non-WIC controls – analysis used birth outcome and dietary data	Women with singleton births, without major fetal growth/survival problems; $n=2,187$ with protein intakes >25 g/day and <150g/day; multiethnic
Kowaleski-Jones and Duncan (2002)	Retrospective WIC cohort ('90–'96 births)	Published article (secondary data, National Longitudinal Survey of Youth)	WIC, non-WIC, national sample; WIC measured self-report Y/N for previous year; BW self-report by mom	(a) WIC, 1984 children born between '90 and '96, moms were age 25–38 (b) siblings: 969 pairs, both born '90–'96, 349 non-WIC and 94 WIC (c) discordant siblings: mom on WIC for one and not both pregnancies, $n=71$
Reichman and Teitler (2003)	Retrospective	Published article (secondary data)	NJ HealthStart program	90,117 enrolled between 1988 and 1996; singleton births; 81% WIC participation; 29% teens; 69% AA or Latino; 30% foreign born
Lazariu-Bauer et al. (2004)	Two stage design: (a) compared WIC groups, then (b) propensity score analysis to further control for selection bias	Published article (secondary, WIC and birth records)	NY state, 1995 births, all WIC	All WIC, only singleton births, 77,601 matched files WIC and birth records; looked at 4 PNC levels

Address disparities (yes/no)	Key findings related to intervention effectiveness	Caveats/biases	Findings support the intervention? yes/no for which populations?
Yes	Likelihood of IUGR significantly lower for women who received WIC and who gained 41 lbs during pregnancy Additional findings: Odds of other adverse outcomes decreased if mom WIC participant	Various adjustments, but still issues with selection bias. Does not specifically address limitations	Yes, for women who gained ≥41 lbs during pregnancy
Yes	Compared to women in the intermediate protein stratum, adjusted mean BW of infants born to mothers in low protein group was 77 g ($p=0.02$) less; mean BW of infants born to mothers in high protein group was 71 g ($p=0.009$) less; this suggests a significant quadratic relationship between maternal protein intake and BW	Adjusted for covariates, including gestational age at delivery and maternal energy intake; 24-h recalls may have attenuated results	Yes
Over-sampled for black, Hispanic and economically disadvantaged whites; some analyses by ethnic group	In all models presented (OLS, fixed effects with and without adjustment for multiple characteristics), prenatal WIC participation had positive impact on BW. In the fully adjusted OLS models, the impact of WIC was significant (0.055 unit difference)	Significant missing data. Sibling sample had small n, but had 90% power to detect 7% impact	Yes
Yes	All four nutrition services associated w/ increase in BW; WIC effects stronger in restricted sample of women with inadequate nutrition; in adjusted model, WIC participation associated with 48 g increase in average BW ($p<0.001$) and a 12% decrease in odds of LBW (OR: 0.88, $p<0.001$)	Controlled for multiple psychological and social variables. Potential adverse selection bias of higher risk women into program	Yes, study compared women at nutritional risk vs. participation but no nutritional risk
Adjusts separately by race/ethnic for selection bias factors and 4 PNC groups	Positive effect of long term prenatal WIC participation on BW for all groups. Full term: infants born to mothers with early WIC participation were a mean of 69.7 g (95% CI 59.7–79.8) heavier than infants born to late WIC participants. *Preterm*: infants born to mothers with early WIC participation were on average 129 g (95% CI 85.9–175.8) heavier than infants born to late WIC participants; *Full term ethnic*: Mean WIC effects larger for black and Hispanic infants (79g (95% CI 60–96), and 75.8 g (95% CI 30–61), respectively) than for white infants (43 g (95% CI 30–61)	Adjusted for both selection and gestation bias; may not have included all the covariates; small preterm sample	Yes

(continued)

Table 14.2 (continued)

Author	Study design	Study type	Intervention description what, how and where	Populations Studied (ages, race/ethnicity) and sample size
Joyce et al. (2008)	Retrospective	Published article (secondary data)	Pregnant and postpartum WIC participants from 9 states submitting data to PNSS, 1995–2004	2,039,102 pregnant women for "intervention" and 742,596 postpartum women as comparison group
Small studies of supplemental nutrition				
Wigda and Lewis (1999)	Prospective, randomized trial	Published article (primary data)	Pregnant women received regular home visits from nutritionist vs. no intervention visits; educational curriculum	40 women completed intervention; 26 controls; mean age 22 vs. 22.8 year; 7 minority women (AA, Asian, Native American) in intervention, and 11 in control group
Briley et al. (2002)	Randomized trial-2 groups	Published article (primary data)	In-home visits for women on WIC in Nebraska	All AA women in WIC, of 27 recruited, 10 intervention and 10 control completed study; 70% each group <21 years
Long et al. (2002)	Quasi-experimental	Published article (primary data)	New Hampshire, Great Beginnings Curriculum, WIC, published 2002	All <20 years (teens) $E = 136$ curriculum +WIC. C-1 = 65 WIC only. C-2 = 50 non-pregnant high school students+curriculum. C-3 = 50 non-pregnant high school students with no curriculum
Canadian nutrition supplementation studies				
Higgins et al. (1989)	Retrospective, within-mother evaluation of HNIP vs. non-HNIP	Published article (secondary data)	Montreal, 1963–1979; Royal Victoria Hospital; Higgins Nutrition Intervention Program (HNIP)	552 sibling pairs-pregnancy for 2nd child was HNIP intervention, all >28 weeks and >999 g; mom age 28 at 2nd pregnancy, parity 4
Dubois et al. (1997)	Retrospective cohort; Higgins program teens vs. randomly selected controls	Published article (secondary data-medical records)	Higgins method: supplemental foods, counseling (HNIP); 1981–1991; Montreal; 15 hospitals	Adolescents; mean delivery age of 17.7 years both groups; 1,203 intervention and 1,203 controls

Address disparities (yes/no)	Key findings related to intervention effectiveness	Caveats/biases	Findings support the intervention? yes/no for which populations?
Analyzed by race/ethnicity: AA, Hispanic	Prenatal WIC participation associated with adjusted mean difference of 63 g in birth weight; 2.7% decrease in LBW and 0.9% decrease in VLBW for entire sample. Analysis by trimester shows a complex picture; women enrolling in third trimester had higher mean birth weight and lower LBW and VLBW rate than postpartum comparison women or women who enrolled in WIC 1st trimester. Prenatal participants averaged 40 g greater birth weight than comparison postpartum women, with lower SGA and LBW	Complex analyses that accounted for many confounders; includes measures of WIC enrollment by trimester	Yes
Yes, but not analyzed	Delivery weight, pre-pregnancy BMI, nutritionist visits and change in energy intake, accounted for 65% of variation in BW	No ethnic analysis; small study	Yes
Entire sample was AA	BW significantly greater in intervention group (3.54 vs. 3.06 kg, $p<0.05$)	Very small sample; confounders not addressed	Yes
Controlled for race/ethnicity-but not discussed separately	E with higher BW than C-1 (3,328.6 vs. 3,206.18 g, $p < 0.05$); LBW rate 9% E and 8% C-1 ($p<0.05$)	Possible selection bias not addressed; no information about when in pregnancy intervention began	Yes, but relatively small, some biases
	BW increased 275 g in undernourished, 319g in stress group, 210 g in multiple conditions group with underweight and 167g in multiple conditions without underweight. Increases all significant. Matched OR shows lower risk for LBW and IUGR for intervention group, especially among undernourished	Minimal selection bias. Changes in smoking patterns may have increased program impact	Yes
Some non-white	Rates of LBW, VLBW, preterm, lower in intervention group; BW of intervention infants 55 g higher than control, and 39% lower rate of LBW, 56% lower VLBW than controls; preterm rate 41% lower	Adjusted for several confounders	Yes

(continued)

Table 14.2 (continued)

Author	Study design	Study type	Intervention description what, how and where	Populations Studied (ages, race/ethnicity) and sample size
Desjardins and Hardwick (1999)	Retrospective, one group	Published article (secondary data)	Toronto Healthiest Babies Possible program, Toronto, ON, 1987–1996. Supplemental milk, home counseling visits with RD/RN. To determine dose of visits with greatest impact on BW	High nutritional risk, low-income women who completed program, sample of 1,883 singleton births; included adolescents (21%)

some studies reported on reduction of LBW risk as a result of program participation while others addressed the greater likelihood of LBW in the unsupplemented women. These differences in units of analysis and reporting made comparisons across studies more difficult.

Few studies reported or referred to a theory or conceptual framework for the intervention or the analysis, resulting in lower quality scores across the board. Overall, studies analyzing secondary data did not note adverse events or negative outcomes due to the intervention, which also lowered the study scores. Likewise, in terms of external validity criteria, studies using secondary data analyses typically did not state if the participants were "representative of the population," although most did statistically control for differences between participants and non-participants. Another external validity criterion that may have led to lower scores for studies using secondary data analyses was the participant eligibility criteria; in the Medicaid-WIC studies, all pregnant women enrolled in Medicaid (e.g., low income) were included in the analysis, with the only selection criterion being whether a woman met the nutritional risk requirements for WIC, however, it was not known whether non-participants met the WIC risk criteria. There was also no information on whether participants were "lost to follow-up," an important criterion for internal validity. Thus, while there are a number of problems with the included studies, they are the key published studies reflecting U.S. programs since 1985, and as such, represent the best available evidence for a review of the effects of supplemental nutrition interventions on birth weight.

Evidence Assessment in the Selected Studies of Supplemental Nutrition Programs on Low Birth Weight

In general, the studies evaluated found that supplemental foods, particularly to at risk women, led to modest increases in mean birth weight and reduced rates of LBW (Table 14.2). These findings are also supported by the meta-analyses that included women from both industrialized and less developed countries (Table 14.3). The majority of maternal nutrition and birth outcomes research in the U.S. involves WIC, the predominant food assistance program targeted toward pregnant women. There are several ways to characterize these studies which help in understanding the state of the evidence for the role of nutrition program in improving birth outcomes. Much of the extant research

Address disparities (yes/no)	Key findings related to intervention effectiveness	Caveats/biases	Findings support the intervention? yes/no for which populations?
No note of race/ethnic analysis. Clients described as multi-cultural	Number of visits independently associated with decreased LBW, controlling for gestational age and smoking. Clients with low (<9) visits have 70% greater odds of having a LBW infant than women receiving 9 visits or more, controlling for risk factors. Probability of LBW decreased by ~10% for each home counseling visit	Supplemental milk combined with home visits from RD/RN	Yes

is relatively dated, having been published 10 or more years ago. Some studies are secondary analyses of program data, in which WIC participants and non-participants are compared; these include studies of the Medicaid program, as well as large national surveys, such as the National Maternal and Infant Health Survey or the National Longitudinal Survey of Youth. One strength of these secondary analysis studies is their relatively large sample sizes, and in several cases, their inclusion of all the Medicaid participant women in a particular state(s).

There are several reports which address nutrition and birth outcomes in specific states (Ahluwalia, Hogan, Grummer-Strawn, & Colville, 1998; Buescher & Horton, 2000; Buescher, Larson, Nelson, & Lenihan, 1993; Devaney, 1992; Devaney, Bilheimer, & Schore, 1990, 1991; Lazariu-Bauer, Stratton, Pruzek, & Woelfel, 2004; Reichman & Teitler, 2003; Stockbauer, 1986, 1987). In the time period examined (1985–2008), there has been one national evaluation of the WIC program (Rush et al., 1988a, 1988b). There was also one moderately sized RCT of the WIC program and birth outcomes (Metcoff et al., 1985). A few other studies examined smaller interventions, usually home visits to pregnant women and/or nutrition education, in conjunction with supplemental foods. Finally, there are a few published reports of maternal nutrition interventions and birth outcomes from Canada. Articles presented in Tables 14.1 and 14.2 have been grouped by whether they examined WIC participants and non-participants within the Medicaid program (Ahluwalia et al., 1998; Bitler & Currie, 2005; Buescher & Horton, 2000; Buescher et al., 1993; Devaney et al., 1990, 1991, 1992; Joyce, Gibson, & Colman, 2005; Schramm, 1985; Stockbauer, 1986, 1987); or only WIC participants (Caan et al., 1985, 1987; Frisbie, Biegler, de Turk, Forbes, & Pullum, 1997; Gordon & Nelson, 1995; Joyce et al., 2008; Kowaleski-Jones & Duncan, 2002; Lazariu-Bauer et al., 2004; Metcoff et al., 1985; Reichman and Teitler, 2003; Rush et al. 1988a), as well as small studies of supplemental nutrition (Briley et al., 2002; Long, Martin, & Janson-Sand, 1995; Wigda and Lewis, 1999). Canadian studies are also grouped (Desjardins & Hardwick, 1999; Dubois et al., 1997; Higgins et al., 1989). The meta-analyses (U.S. General Accounting Office (GAO), 1992; Kramer, 1993; Avruch & Cackley, 1995; de Onis et al., 1998; Merialdi et al., 2003) data are in Table 14.3[1].

[1]A 2009 Cochrane Review (Kramer & Kakuma) draws approximately the same conclusions of a modest effect on birth weight as the other studies in the meta-analyses presented here. For balanced protein/energy supplementation interventions, Kramer and Kakuma rate the studies as of 'variable quality' with several methodological problems.

Table 14.3 Meta-analyses of nutrition supplementation interventions during pregnancy and low birth weight (and related outcomes)

Source	Number of studies/*N*/(nutritional interventions)	Findings	Contextual factors	Disparities/comments
GAO (1992)	17 studies of WIC and non-WIC pregnant women between 1971 and 1988; included Medicaid; primary analysis is for cost savings	Women receiving WIC had 25% fewer LBW infants; greater effect on VLBW-decreased by 44%	9 states were represented, some more than once; sample size varied from 79 to 31,732. Most studies were quasi-experimental, with statistical controls	Not addressed
Kramer (1993)	Twelve trials of balanced energy and protein, balanced isoenergetic protein supplements	Balanced increases energy and protein supplementation produce modest increases in results in maternal weight gain and fetal growth. Limited effect of supplements on gestational age	Included international studies, used data from Oxford Database of Perinatal Trials up to 7/1/92	Not addressed
Avruch and Cackley (1995)	13 studies, WIC vs. non-WIC, Medicaid	WIC participation associated with a 25% reduction in LBW rates (a 16% reduction in non-Medicaid studies and a 28% reduction in Medicaid studies). In non-Medicaid studies, WIC participation associated with a 16% reduction in LBW rates. In VLBW studies, a LBW rate decrease of 44% was associated with WIC participation	Update of 1992 GAO report; studies using random assignment or statistically assigned controls; primary interest is cost savings. Some states represented multiple times and others not at all. Limitations of using Medicaid as a population; not all studies controlled for gestation length. Unclear what elements of WIC contribute to better BW	Not addressed specifically
de Onis et al. (1998)	Systematic review of 7 RCTs of balanced protein/energy supplementation during pregnancy on SGA and LBW; conducted between 1973 and 1988	Across studies, authors reported an average OR of 0.77 (0.58–1.01) for LBW. Attributed these effects to small energy intake increases in the trial women. Supplementation associated with maternal weight gain and mean birth weight increases and SGA decreases of borderline significance	Two of 7 trials in industrialized countries and others were international. Participants generally from "underprivileged" groups. Methodological quality described as variable, and greater in double blinded trials	Not addressed
Kramer & Kakuma (2009)	13 trials with 4,665 women; balanced energy/protein supplementation	Across studies in a random effects model, associations of birth weight increases and supplementation (WMD 37.62 g, 95% CI: –0.21–75.45). Also decreased SGA births (RR = 0.68, 95% CI 0.56–0.84)	Only 1 trial included U.S. women	Not addressed

WIC Only Studies

Two studies which assessed only WIC participants are frequently cited as providing evidence for WIC program effects. The National WIC Evaluation study was published in 1988 and based on data collected between 1983 and 1984; this is the only national evaluation of the program since its inception over thirty years ago. There were various study components to the evaluation, but only those from the Historical Study (Rush et al., 1988a) and Longitudinal Study (Rush et al., 1988b) are included here. The longitudinal study used a prospective, quasi-experimental design to examine birth outcomes with women from 174 WIC clinics throughout the U.S. and women from 55 prenatal clinics without WIC services; the sample size was 5,205 WIC and 1,358 non-WIC pregnant women. For most study birth outcomes (mean birth weight, very low birth weight frequency), whether unadjusted or adjusted for gestation, there were no significant differences between the WIC and the non-WIC births. Using intention to treat analysis, although not significant, the LBW rate was 1.1% lower in the WIC group and 2.6% lower in the control/WIC crossover group (some control women joined WIC late in pregnancy) than the 6.8% rate for the control women. Likewise, there were no significant relationships between WIC participants and gestational age. On the other hand, there were significant differences in infant head circumference associated with WIC participation. There was a significant association between perception of program quality as reported by WIC clinic directors and fetal growth rates when pregnancy duration was controlled; reported LBW rates were 1.9% lower for each standard unit increase in program quality score ($p<0.001$). Generally, there were issues of power in this study because there were fewer controls than expected, and there was confounding due to crossover, as approximately 25% of the controls were later enrolled in the WIC program. Additionally, the control women were more socially advantaged than women in the WIC program; however, this indicates that more disadvantaged women are likely to participate in WIC, meeting a national program goal (Rush et al., 1988b).

As part of the National WIC Evaluation, retrospective program data from the program's federally authorized inception in 1972 (although the first clinic and certified participants began in 1974) to 1981 were analyzed in what Rush et al. (1988a) described as an "ecological model analysis." Data were included from 888 U.S. counties located in 14 states and the District of Columbia. Using multiple regression techniques, it was reported that the data suggested important programmatic effects although WIC's impact on birth weight was not entirely consistent. Among program participants, birth weights were about 250 g higher among white women and 100 g higher among better educated women. The mean birth weight rose over the study period in both whites and blacks, and was 5–6 times higher in well educated white and black women. The authors ultimately concluded that there were "considerable WIC program effects on mean birth weight".

The second most frequently cited study of WIC effects on birth outcomes was an RCT published in 1985 by Metcoff et al. The research, conducted in one Oklahoma hospital, included 172 WIC participants and 238 WIC-eligible control participants (identified from a total of 683 singleton births). Given that the overall demand for WIC participation exceeded the resources allotted to the clinic by the state, the randomization of mothers was deemed ethical. The study found a slightly higher mean birth weight among the WIC women compared to the eligible controls (3,254. vs. 3,163 g, $p=0.0395$), but this difference was no longer significant when the mother's weight at program entry was controlled in the analysis. WIC also had a positive effect on infant birth weights among mothers who smoked ($p=0.017$).

Other studies which examined only WIC participants and non-participants were generally secondary analyses of data from other studies and usually reported positive outcomes, although there were differences in how results were reported. The Frisbie et al. article (1997) and USDA report (Gordon & Nelson, 1995) examined birth outcomes in the 1988 National Maternal and Infant Health Survey participants. Infants of mothers who participated in WIC had decreased IUGR

compared to the infants of non-participants (OR: 0.7, CI = 0.6, 0.9). The USDA study reported significant differences in the LBW rate between WIC participants and income-eligible non-participants of from 1–2%. Kowaleski-Jones and Duncan (2002) examined WIC participation and birth outcomes in the National Longitudinal Survey of Youth. In an adjusted OLS model, they found a significant difference in birth weight between WIC and non-WIC participants of 0.55, although data showed a selection bias with higher WIC participation women at higher risk of LBW. Other studies used state or program level data and generally reported a positive WIC effect on birth weight (New York, Lazariu-Bauer et al., 2004; California, Caan et al., 1987; New Jersey, Reichman & Teitler, 2003; Missouri, Schramm, 1985; Missouri, Stockbauer, 1986, 1987).

A recent analysis by Joyce et al. (2008) examined WIC participation and birth outcomes using 10 years of data from the Pregnancy Nutrition Surveillance System (PNSS) of CDC. Data were included from nine states, and the sample included 2,039,102 pregnant women who participated in WIC during their pregnancy and 742,596 postpartum women who didn't enroll in WIC until after their pregnancies as a comparison group. PNSS provides information on trimester of WIC enrollment (intervention exposure), birth outcomes and some maternal behaviors. The econometric regression models controlled for many covariates and confounders to try to account for biases in past studies. In comparing WIC prenatal participants with those who entered WIC in the postpartum period, they report an adjusted mean birth weight difference of 63 g, and lower rates of LBW and VLBW (2.7 and 0.9%, respectively). Unexpectedly, they found that women who enrolled during their third trimester of pregnancy were more likely to have a higher mean birth weight, and lower LBW rates than women who enrolled during their first trimester or the postpartum comparison group, although they don't provide much explanation for these counterintuitive findings. They reported higher birth weight gains among black and Hispanic women than among non-Hispanic whites, although the mean birth weight for black infants remained lower by about 200 g than that for non-Hispanic whites or Hispanics. According to Joyce et al., early enrollment in WIC appears to be associated with some changes in maternal behavior and health, although they describe the effect of these changes on birth outcome as "modest."

Studies Evaluating WIC Participation Effects among Medicaid Recipients

A number of studies used Medicaid participants to examine birth outcome differences between WIC participants and WIC non-participants who were by their Medicaid participation income-eligible for WIC (nutritional risk status, required to qualify for WIC, was not known in the non-WIC participants). Bitler and Currie (2005) compared birth outcomes among Medicaid participants recorded in the Pregnancy Risk Assessment Monitoring System (PRAMS) data from 19 states and found the odds ratio and (standard error) for the association between WIC participation and LBW to be 0.73 (0.03), significant at the $p < 0.001$ level. Larger differences were found for more disadvantaged women (i.e., those who were single, high school drop out and/or teen mothers). Devaney et al. (1990, 1991), and Devaney (1992) examined all Medicaid births in 1987 for five states (North Carolina, South Carolina, Florida, Texas, and Minnesota) and statistically controlled for differences between WIC participants and non-participants. While there were differences between the states, WIC participation decreased the LBW rate, and resulted in average increases in birth weight ranging from 51 to 117g across the five states. Buescher et al. (1993) examined birth outcomes among Medicaid participants. In 1988, non-WIC Medicaid participants in North Carolina were 1.45 times as likely to have a LBW infant (CI = 1.32, 1.59), and 2.15 times as likely to have a VLBW infant (CI = 1.77, 2.61) compared to WIC Medicaid participants. A more recent analysis of 1997 North Carolina births reported similar findings [LBW OR: 1.36 (1.26, 1.45); VLBW OR: 1.9 (1.61, 2.25) (Buescher & Horton, 2000)].

Joyce et al. (2005) examined all Medicaid births in New York City between 1988 and 2001. For singleton births, the WIC participant LBW effects were greater in 1988–1992 than for the 1993–1997 period when Medicaid benefits were expanded. They reported their findings as "modestly supportive" of a WIC participation benefit. Ahluwalia et al. (1998) examined Medicaid births in Michigan for 1992. The effects of WIC on SGA were reported by WIC exposure time, with greater program effects in women who participated longer (21–37 weeks, OR: 0.82, CI = 0.72, 0.93; 12–20 weeks, OR: 0.78, CI = 0.68, 0.89; and <12 weeks, OR: 0.72, CI = 0.6, 0.86).

Canadian Studies

As noted in an earlier section, some of the initial research into the role of food and nutrition during pregnancy on birth outcomes in industrialized populations was conducted in Canada, and three of these studies are included here. Higgins et al. (1989) reported a within-mother analysis of women who had two births in a Montreal hospital; for the first birth, a woman did not participate in the nutrition program (non-intervention group) but for the second birth each was a HNIP participant (intervention group). Participants received supplemental foods along with nutrition and prenatal services. Matched pairs odds ratios (OR) show a lower risk for LBW and IUGR among program participants, particularly in women who were undernourished at the beginning of pregnancy. Controlling for parity and infant sex between the paired siblings, birth weights averaged 190 g ($p < 0.001$) more in the intervention infant group than in their matched non-intervention siblings. There was also a gradient of birth weight effects with varying degrees of risk among the mothers, with greater increases in the infants of women with greater risk. Another analysis used Higgins Program recipients ($n = 1213$) who were retrospectively matched with controls on various characteristics which showed that the rate of LBW was lower in the intervention infant group (although not significantly so) than the controls (5.7 vs. 6.8%), however, the mean birth weight of infants in the intervention group was 40 g greater, a difference which was significant (Rush, 1981). These differences were even greater in intervention women without prior pregnancies, and for women weighing less than 140 pounds prior to the pregnancy. An adolescent intervention using the Higgins method in 1,203 pairs reported that the adjusted rates of LBW (OR: 0.61, CI = 0.45, 0.82), and VLBW (OR: 0.41, CI = 0.23, 0.82) were lower in the intervention group than the controls (Dubois et al., 1997).

Very Low Birth Weight Studies

Several studies separately reported the related birth outcome of very low birth weight (VLBW) as an outcome. An extension of the Devaney analysis of prenatal WIC participation on infants of Medicaid participants in five states (1992) was published separately examining infants with VLBW. In four of the five states (FL, MN, NC & SC), women's WIC participation during pregnancy was associated with a reduced VLBW probability, with the rates about half those noted in the non-WIC Medicaid participants. She noted that the number of VLBW births prevented in women who began WIC participation by 30 weeks gestation ranged from 191 births in Florida to 352 births in North Carolina. The Buescher et al. (1993) study of North Carolina Medicaid births also reported significantly lower VLBW rates for WIC participants, and noted that reductions were greater in the black WIC participants. After controlling for sociodemographic, prenatal care and medical variables, they found that the non-WIC participants were 2.15 times as likely to have VLBW infants (CI = 1.77, 2.61). In the multivariate analysis, WIC participation was the second greatest predictor of VLBW, following medical risk. Their analysis of data for 1997 (compared to the 1988 data above) confirmed

the findings of the earlier study (Buescher & Horton, 2000), with a VLBW OR of 1.9 (CI = 1.61, 2.25). VLBW findings in a similar range were reported by Stockbauer (1987), where there was a 27% decrease in VLBW. The Dubois et al. study (1997) noted a 56% lower rate of VLBW in the intervention participants than in the controls. Some of these findings are summarized in the GAO meta-analysis (1992) which commented that WIC had a greater effect on VLBW (44% reduction) compared to LBW (25% reduction).

Evidence Summary

When examined as a whole, the majority of studies provide support for a positive impact of the WIC program on birth weight. This is expressed by both an increased mean birth weight and a decreased percent of LBW in program participants when compared to non-WIC participants. However, many of the effect sizes are rather small, and few studies reported the intervention effects in the same way. Some (Ahluwalia et al., 1998; Bitler & Currie, 2005; Buescher & Horton, 2000; Buescher et al., 1993; Caan et al., 1987; Devaney, 1990, 1991, 1992; Dubois et al., 1997; Frisbie et al., 1997; Lazariu-Bauer et al., 2004) presented the effects as OR of having a LBW infant, while other studies (Desjardines & Hardwick, 1999; Devaney, et al., 1990, 1991, and Devaney, 1992; Frisbie et al., 1997; Gordon & Nelson, 1995; Joyce et al. 2005, 2008; Kowaleski-Jones & Duncan, 2002; Reichman & Teitler, 2003; Rush et al., 1988a, 1988b; Sloan et al., 2001; Wigda & Lewis, 1999) reported regression coefficients. Another issue in comparing studies is that different LBW-related outcomes were reported; some studies reported on LBW only (often controlling for gestational age) while others reported infants as being SGA or as having IUGR (intrauterine growth restriction). Studies also varied in how they expressed program effect difference; in some cases, differences in *mean birth weight* are reported, in others, differences in the *percent or rate of LBW* infants were reported. Some studies report decreased risk of LBW for participants while others report the degree of greater LBW risk for non-participants. These differences in outcomes make understanding study differences in WIC impact more difficult to discern.

Keeping in mind that the sample sizes were different across these comparisons, the studies which presented OR showed that non-WIC participants were 1.3–3 times as likely to have LBW infants, or conversely, that WIC participants were between 0.73 and 0.87 as likely to have a LBW infant. When the LBW rates were presented, the range of decreased LBW risk among WIC participants was between 1.1 and 6.9%, while the risk of having a LBW infant in non-participants was increased by 6.4–8.9%.

Methodological Issues in Assessing the Available Evidence for Supplemental Nutrition Programs

In analyzing studies of nutrition intervention effectiveness there were several issues that became apparent across the studies and in the general examination of the role of nutrition in pregnancy outcome. Many of the studies reported here are secondary data analyses where selection of the sample and variables were not part of the original study. While appropriate statistical methods were used, they share some limitations which are described below. Related to this issue is the overall age of the data used for the secondary analyses.

One key problem with evaluating the effect of a program such as WIC on LBW (or other birth outcomes) is that nutrition is just one factor affecting LBW, and with a multi-component program such as WIC, the question arises of whether the birth weight differences are attributable to the

supplemental foods, the nutrition education, the referrals for other services, some combination of these or another factor altogether? None of the studies attempted to determine this. This inability to separate out the nutrition supplement and education (and referrals) components reflects a larger problem in examining WIC program impact which hasn't been studied and will be difficult to resolve in future research. In addressing the issue of reducing LBW in high risk populations, such as low income women and/or ethnic/racial minorities, few evaluations/studies are ethically able to randomize the women into WIC (treatment) or non-WIC (control conditions). In only one case, was randomization possible when there were funding limits restricting enrollment of all at risk women, and women on the waiting list were evaluated as controls. However, this study (Metcoff et al., 1985) and the Rush et al. longitudinal study (1988b) were hampered by subsequent enrollment of the control women into WIC when spaces became available which resulted in much smaller control groups than anticipated. In recent years, increased funding has allowed WIC to accommodate all women with nutritional risk seeking participation (IOM, 2003) making control group selections more difficult. In 2006, the USDA Food and Nutrition Service calculated that nationwide 1,214,682 pregnant women were eligible for WIC, and 845,071 were participating, representing a 69.6% coverage rate (USDA, 2006). Notably, although increased funding and successful outreach have led to increased WIC programmatic coverage over time, methods for estimating and calculating eligibility and participation rates have been questioned; the National Research Council report indicates that because there is underestimation of those eligible for WIC, published coverage rates may be an overestimate (Ver Ploeg & Betson, 2001).

Attempting to overcome the lack of randomization and to reduce selection bias, several studies included non-WIC low income pregnant women enrolled in Medicaid as a comparison group. However, while the women are comparable on income, there may be other selection biases that distinguish WIC and non-WIC participants. Some of these biases could affect birth outcomes/study results favorably (e.g., interest in having a healthy baby and seeking early prenatal care) or unfavorably (e.g., substance abuse, smoking). Since the time when most of these studies were completed the Medicaid program has expanded to cover women at incomes up to 185% of poverty and beyond. WIC funding also has increased to include more at risk women in the program. These changes were made to reduce poor birth outcomes, and while the program expansion highlighted the necessity of more thorough evaluations, it also made them more difficult. In addition, some have argued that too many lower risk women are now being included in the program (Besharov & Germanis, 2001), perhaps diluting program effects. There is other research however showing that more disadvantaged women are more likely to enroll in WIC (Gordon & Nelson, 1995; Joyce et al., 2005). The definition of "disadvantaged" varies by study, although generally low income women, adolescents, and ethnic and underrepresented minorities are included.

It is important to recall when evaluating the WIC studies that maternal nutritional risk is a prerequisite for WIC enrollment, with inadequate nutrient or food intake and anemia being the most common prenatal risk factors. The majority of the studies did not examine nutritional risk beyond whether or not a woman was qualified for WIC, although a few studies did provide information with regard to differential effects of supplementation based on the mother's level of nutritional risk. Reichman & Teitler (2003) noted that the positive effects of WIC grew stronger when the sample was restricted to the 19,307 women classified as having "inadequate nutrition," although this term was never defined in the paper. In Reichman and Teitler's fully controlled model of these high nutrition risk women, there was an average 48 g increase in average birth weight ($p<0.001$) along with and a 12% reduction LBW odds (OR:0.88, $p<0.001$) among the WIC participants. The Higgins program studies (Dubois et al., 1997; Higgins et al., 1989) generally classified their participants by nutritional risk (no risk, underweight, undernourished, stress or multiple conditions) and do report differential supplementation effects. The smallest effects on birth weight were seen in the underweight group (61 g), and the largest in the undernourished group (146 g).

WIC program funding must be annually authorized by Congress; thus, there is a political aspect to the program which has influenced its implementation and evaluation over time. To date, there has been only one national evaluation of the program, which in itself was controversial (GAO, 1989; Rush et al., 1988a, 1988b). It seems unusual that a federal program such as WIC has not been regularly evaluated in its 30 year existence, but this is the case. One result is that none of the research reported here captures the changes that have occurred as a result of either the WIC or Medicaid expansions since the mid-1980s evaluation. For example, there is much greater program coverage now than when the program was last evaluated (Fox, Hamilton, & Lin, 2004), and currently most eligible women are able to receive WIC benefits. Although Medicaid expansions for pregnant women provided higher income eligibility criteria often exceeding the WIC cutoff of 185% of the FPL, and conferred automatic income eligibility for the WIC program, nutritional risk criteria still needed to be met in order to qualify for the program (Fox et al., 2004). Another change in the WIC program which had the potential to affect evaluation studies was the standardization of nutritional risk criteria across state WIC agencies in 1998 (IOM, 1996; Ver Ploeg & Betson, 2001).

A complicating issue in efforts to examine the impact of supplemental foods is the difficulty of obtaining accurate measures of dietary intake, particularly in population studies. Of the larger studies reviewed here, only the National WIC Evaluation presented data on the role of actual dietary intake on birth outcome. Even in studies where food coupon redemption was assessed and used as an indicator of program "effect," there is no guarantee that the women actually consumed the supplemental foods, since they were likely available to the entire family. The need for better assessment of dietary intake in relation to birth outcome was also noted by Owen & Owen (1997) and IOM (2002a).

Information on duration of program exposure is difficult to obtain in secondary data analyses (the majority of studies), since these tend to use birth certificate or Medicaid data which generally only indicate whether or not the mother participated in WIC during the pregnancy rather than how long she participated. Also, even participation from early in the pregnancy does not necessarily mean that women consumed the supplemental foods or put into practice the nutrition education they received. Ahluwalia et al. (1998) found that the women with high exposure to WIC (enrolled before 12 weeks pregnancy) were half as likely to have an infant who was SGA than were women with no exposure to WIC, and there was about a 100 g mean birth weight difference between the two groups. But the issue of program exposure can be confounded by the fact that pregnancies with longer exposure are also often those with longer gestations, making it difficult to tease out the cause of any identified effects. Several of the larger studies did adjust for gestational age or length of gestation in their analyses (Caan et al., 1987; Higgins et al., 1989; Joyce, Gibson, & Colman, 2004, 2005; Lazariu-Bauer et al., 2004; Metcoff et al., 1985; Rush et al., 1988b; Stockbauer, 1986).

Most studies presented here had good sample sizes (often due to the secondary analysis of large data sets), but very few addressed power issues in reporting their results. Another issue is that not all states are represented in the studies, and thus the findings may not be generalizable across program participants/non-participants throughout the country. Related to this is the fact that several states have been studied more than once, and these findings may "tip the balance" of program effects one way or the other.

Summary of the Evidence and the Role of Nutrition Interventions in Reducing Racial/Ethnic Disparities in Low Birth Weight

Racial and ethnic differences in birth outcomes in the United States are well recognized (DHHS, 2000), although to date few solutions to resolve these differences have been fully successful on a population basis. What is generally accepted at this point is that the causal factors for disparities are

multifaceted, and likely are influenced by more variables than can be addressed during a pregnancy (Lu & Halfon, 2003). It should be noted, though, that at the time most of the data for the studies discussed in this chapter were collected, this more comprehensive approach to understanding disparities was not routinely considered.

Many of the studies reviewed here did report differential findings for African-American/black women compared to white women, with a few studies examining outcomes in Hispanic women (Bitler & Currie, 2005; Devaney et al., 1990, 1991; Frisbie et al., 1997; Joyce, 2005; Kowaleski-Jones and Duncan, 2002; and Lazariu-Butler et al., 2004). As with other reports, infants born to African-American women were more likely to be LBW whether or not their mothers were WIC participants, although outcomes were generally improved in African-American WIC participants (Buescher et al., 1993; Devaney et al., 1990,1991; Frisbie et al., 1997; Gordon & Nelson, 1995; Joyce et al., 2005; Kowaleski-Jones & Duncan, 2002; Lazariu-Bauer et al., 2004; Rush et al., 1988a; Sloan et al., 2001). In some of the studies, the improvements noted among African-Americans were the ones that reached statistical significance. For example, in the Stockbauer study (1986), mean birth weight increases in non-white infants were three times those of white women (48 vs. 16 g) even though these differences are relatively small. Stockbauer also reported that the lower LBW rate for the overall sample was due primarily to decreases in LBW among non-white participants. A subsequent analysis also found greater decreases in LBW rates among the non-white infants (Stockbauer, 1987). Strong racial differences were reported in the WIC-Medicaid cost studies conducted in 5 states by Devaney et al. (1990, 1991) and Devaney (1992). Joyce et al. (2005) also found racial differences in New York City, particularly in the 1988–1992 period, which the authors note may be explained by greater crack-cocaine exposure among the U.S. born black non-WIC participants.

Other research explored the potential role of protein and birth weight disparities among African-American women. Low protein intake was initially thought to be related to poor birth outcomes, but the Harlem trial with African-American women found potential adverse effects with high protein supplements and birth outcomes (Rush et al., 1980). A role for protein nutriture in birth outcomes was explored by Sloan et al., (2001) in an analysis of the energy distribution of the National WIC Evaluation participants. They reported that higher protein intakes (and poorer birth outcomes) were found in African-Americans, those on welfare and Food Stamps, and thinner women. Specifically, they found a significant quadratic relationship between maternal protein intake and birthweight (Sloan et al., 2001). These poorer outcomes (mean birth weight decrease of 71 g) were particularly noted in women whose protein intakes were greater than 85 g/day. [The current Dietary Reference Intakes recommend an average protein intake of 71 g/day for pregnant women (IOM, 2002b), but many Americans regularly consume more than 100 g of protein daily.] Thus, for a group of women at risk for poorer birth outcomes, such as African-American women, consumption of a high protein diet could contribute to the existing birth weight disparities. Protein intake emerges as another potential factor to monitor in interventions to reduce LBW among African-American pregnant women. In a Cochrane review of energy and protein intake in pregnancy, Kramer and Kakuma (2006) notes that as a result of these studies, most subsequent trials have used balanced protein/energy supplements and found "modest increases in maternal weight gain and fetal growth." Thus, encouraging a balance between energy and protein intakes among African-American women could contribute to better birth outcomes in high risk women.

In other studies, race and ethnicity were controlled for in the analysis, but LBW rates for these groups were not reported (Ahluwalia et al., 1998; Long et al., 2002; Metcoff et al., 1985; Reichman & Teitler, 2003; Rush et al., 1988b; Sloan et al., 2001). There were several studies, however, that noted greater WIC participation effects in the infants of more disadvantaged women (Joyce et al., 2005; Kowaleski-Jones & Duncan, 2002; Reichman & Teitler, 2003). For example, Kowaleski-Jones & Duncan (2002) stated that there was a negative selection into WIC of mothers most likely to have a LBW infant.

Relevance and Implications of the Evidence for Practitioners

As a program, WIC has been broadly and successfully implemented across the US during the past 30 years, and there have been recent program improvements as noted earlier. However, as this chapter shows, there are no current outcome data on the WIC program at the national level, and commissioning a second national evaluation of the WIC program to demonstrate its effectiveness would be expensive; the cost for the first evaluation, a six year project twenty years ago, was $5.9 million (GAO, 1989). Barriers to successful implementation, as well as systematic and comprehensive evaluation of the WIC program, include its short term year-to-year funding, as well as legislative funding restrictions, which limit the ability of state and local agencies to conduct evaluation research on program effectiveness. While the admirable low administrative costs for WIC have been key to its political and popular success, requirements for keeping these costs low (at 10% of the budget) preclude the ability of state and local agencies to conduct essential program monitoring and evaluation activities that would further our understanding of program effectiveness and enhance our efforts to promote better birth outcomes for program participants. Commissioning a second national evaluation of the WIC program to demonstrate its effectiveness would be expensive; the cost for the first evaluation, a six year project twenty years ago, was $5.9 million (GAO, 1989).

Even though the strength of the evidence for the role of supplemental nutrition interventions is only moderate, and the birth weight improvements due to supplementation are relatively small, it is difficult to argue against providing food and nutrition education to at-risk pregnant women. The existing outcome data do show that women at greater nutritional risk generally deliver higher birth weight infants when there is a nutrition intervention that improves the quality and quantity of food consumed, so it is likely that the masking of individual differences within these studies may lead to underestimates of program effects (Kent & Hayward, 2007). Along with the need to conduct evaluations to better link birth weight increases to dietary supplementation and demonstrate program effectiveness, it is important for practitioners (i.e., program managers, policy makers, clinicians) to identify and intervene with those women determined to be at nutritional risk, regardless of income and eligibility for programs such as WIC. International LBW studies support nutritional risk identification, although risk identification across the U.S. population isn't the norm (Rush et al., 2001a). Groups most often at risk include underweight, undernourished women, low income women, smokers, and adolescents.

Another group of women not specifically examined in the research presented here, but who merit assessment (and are currently considered at nutritional risk group within the WIC program), are women who are overweight and obese. These women may not eat a balanced diet or gain sufficient weight due to their weight concerns or they may gain excessively, both of which can contribute to poorer pregnancy outcomes (Siega-Riz & Laraia, 2006). Because inadequate pregnancy weight gain and poor birth outcomes were the thrust of the IOM (1990) pregnancy weight gain guidelines, there was little focus on weight gain among overweight and obese women, and clinicians have little guidance. The lack of evidence based guidance for overweight and obese women may lead clinicians to not make appropriate assessments or weight gain recommendations for these women. This, along with women's struggle to not gain too much and contribute to their weight problem, may lead women to consume inadequate or inappropriate foods during pregnancy. As there are adverse outcomes associated with overweight and obesity during pregnancy (Siega-Riz & Laraia, 2006), there are multiple reasons to assess, intervene and conduct research on this group.

The best way to identify nutritionally at-risk women is through individual assessment which includes anthropometric, dietary intake, laboratory, social and environmental measures. Ideally, this would occur at a preconceptional visit, although assessment in early pregnancy is also valuable. The importance of preconceptional care is increasingly recognized (Atrash, Johnson, Adams,

Cordero, & Howse, 2006); actually receiving preconceptional care may be more difficult, especially for low-income women. This can be attributed to factors such as low preventive services use by many young women, their likelihood of not having an intended pregnancy, and poor health insurance coverage. Whether assessment occurs as part of preconceptional care or early prenatal care, referral of nutritionally at-risk women to programs such as WIC, or the Food Stamp Nutrition Education Program, the Expanded Food and Nutrition Education Program and/or the Commodity Supplemental Food Program (all offered by USDA) is also an essential step in reducing nutritional risk. The relatively large numbers of working poor women and families, whose incomes are above the income cutoff for WIC, but who are often uninsured and have poorer health than the general population, are a group that often falls between the cracks of health and social services. Outreach, assessment and intervention with women in this group seem essential to continue to improve birth outcomes in the U.S.

Nutritional risk assessment and surveillance should also be expanded beyond that currently in place for the WIC program; the PNSS was initiated in 1979 by CDC and collects prenatal and postpartum data on nutritional and prenatal risk factors primarily on women and their infants in the WIC program (Perry, Yip, & Zyrkowski, 1995). Expansion of this surveillance and monitoring to all pregnant women would provide information that ultimately could contribute to better pregnancy outcomes.

Practitioners (program managers, clinicians, prenatal program personnel) have a role to play in collecting and aggregating data which can be used to better document the role of prenatal nutrition supplementation and counseling in improved birth outcomes. Important information to collect includes the specific nutritional risk factors, nutrition services offered and received (i.e., counseling, referrals, supplemental food prescriptions, duration of program participation), behavior changes resulting from the intervention (i.e., changes in food intake and behavior patterns), maternal history with regard to her own family nutrition and food intake patterns, food access and security, and chronic disease risk. Collection and aggregation of this information across a large number of women in addition to birth outcome data could provide evidence of supplementation effectiveness over time. Atrash et al. (2006) speaks to the need for a national preconceptional care policy as important in improving birth outcomes; a related policy to incorporate nutritional risk surveillance and monitoring during well woman and preconceptional care complements that proposed for overall preconceptional care. Without a national policy, there is less potential for state or local efforts to encourage private providers to assess, collect and submit nutrition risk data which could be used locally or regionally to target programs to reach a broader scope of high nutritional risk women. Over time, this surveillance and monitoring of high nutritional risk women across providers (not just public health programs) could be used in conjunction with existing perinatal data collection systems to compile evidence on supplementation program effects in the larger at risk population not currently served by WIC. While not perfect for evaluation purposes, preconceptional nutrition risk surveillance data would better identify at risk women for intervention. When these data are combined with birth certificate data, this would be a better measure of program effectiveness than we have currently.

The research presented here also indicates that birth outcomes are more favorable when women participate in WIC from early in their pregnancies through delivery. Early referral, and subsequent enrollment, of pregnant women to WIC by practitioners (clinicians, health and social program managers) would increase womens' exposure to the program and thus potentially have a positive impact on birth outcomes such as LBW. To successfully engage high-risk women, greater outreach services are needed across a spectrum of providers and community leaders. This need is reflected in the recent FNS/OANE estimates indicating *eligible* pregnant women exceed *enrolled* women by nearly 30% (USDA, 2006). The role for outreach extends across practitioners from program planners and policy makers to primary care providers and community outreach workers.

Next Steps for Research

Various noted scholars and researchers have pointed out the problems with existing evidence on nutrition and birth outcomes and have emphasized the need for additional research in this area (Brown & Khan, 1997; Joyce et al., 2004, 2005; Kramer, 2003; Rush, 2001a, 2001b; Villar et al., 2003). For example, Rush (2001a, 2001b) asserts that the research standard in the field of nutrition supplementation during pregnancy needs to improve, and Kramer (2003) suggests the use of biological markers or examination of gene-nutrient interactions as potentially important in epidemiological studies. Some suggestions for research emerging from this review are proposed for consideration.

Joyce et al. (2004, 2005) notes that the lack of rigorous studies and definitive conclusions makes it difficult to challenge popular programs. However, a key research need is for the public health community, policy makers, and funders to gather the political and financial wherewithal to evaluate the WIC program. Continuation of the WIC program is dependent on annual Congressional funding, and competition is high across many social and health programs for federal funding to help vulnerable populations. Policy makers and legislators increasingly will demand evidence of program effectiveness for continued funding. Given the time since the last evaluation and the moderate level of accumulated evidence of program success, a comprehensive program evaluation seems essential for continued program viability in the current political and funding environment.

And, while randomized controlled trials (RCT) are optimal and considered the standard in demonstrating intervention effectiveness, there are various reasons why an optimal evaluation may not include a RCT. Rush (2001a) notes that many trials are relatively small and are conducted under ideal conditions which aren't feasible in population based programs such as WIC. Victora, Habicht, and Bryce (2004) point out various problems with conducting RCTs as a means to establish effectiveness evidence in public health. For example, the causal pathways for many interventions are often long, and/or they are often affected by various characteristics of the environment, population or the health system. Both of these factors are possible influences in studies of the effect of prenatal nutrition supplementation effects on birth weight. These authors suggest that *plausibility evaluations*, designed "to make causal statements using observational designs with a comparison group … are not just a "second best" alternative to RCTs, they are indeed the only feasible alternative" (Victora et al., 2004). There are also ethical concerns in conducting a RCT of WIC program participants related to birth outcomes. These various factors demonstrate that a different approach for evaluating the effectiveness of the WIC program is needed.

An ideal comprehensive study of WIC effectiveness would include identification of a suitable comparison group, study birth outcomes by program components (supplemental food, nutrition education, referrals to other services), account for the length of gestation and the length and "dose" of the nutrition intervention, and incorporate biological and dietary measures. Some other considerations of factors to consider in designing a study follow.

In describing supplementation programs from both less developed and industrialized countries, Rush (2001b) noted that the birth outcomes were relatively similar, and urged that further interventions and research focus on the subsets of women most likely to benefit from supplementation. Since knowledge and methodology surrounding assessment of various micronutrients has improved, examining specific micronutrient effects on LBW and other birth outcomes is another way to refine and expand research in this area. Merialdi et al. (2003) propose "developing hypotheses related to the pathophysiological role of calcium, zinc, and magnesium deficiencies and their effects on fetal growth" (p. 1630S). King (2003) suggests that studies need to examine "the overall change in maternal nutritional status across full reproductive cycle" rather than just the short span of one pregnancy. As Villar et al. (2003) note "When planning and implementing program interventions, one should be aware that a few months of nutrient supplementation during pregnancy may not be sufficient to offset decades of low nutrient intake and its adaptive clinical mechanisms" (p. 1622S).

This points to another key body of research that hasn't formally been linked to supplemental nutrition interventions during pregnancy but which should be evaluated, namely the growing body of research showing the role of the early environment, including nutrition, and its long term effect on subsequent development and health of the fetus (Godfrey & Barker, 2001). This chapter has covered the shorter term factors leading to LBW, but there is also need to consider the pioneering work of Barker and others showing longer term effects of maternal nutritional deprivation during or prior to pregnancy on the later health of the child. These outcomes include greater propensity for hypertension, heart disease, obesity and diabetes in the children of women with poor prenatal nutrition. There is also the issue that currently many disadvantaged women enter pregnancy having grown up with some of these "deficit effects" and thus the cycle of inadequate nutrition and subsequent chronic diseases continues.

Another research need is to attempt to link birth outcomes such as LBW to actual food consumption before and during pregnancy. Related to this is a need for health and social service providers to routinely assess food insecurity. Since the 1980s when most of the studies reviewed here were conducted, the ability to document and measure inadequate food access and intake has improved, and its biological, social and emotional effects for individuals and families have been documented (Hamelin, Beaudry, & Habicht, 2002; Olson, 1999).

Food insecurity and hunger are consistent problems among the poor and near poor in the United States (Nord, Andrews, & Carlson, 2006), with rates consistently higher in minority populations. Effects of food insecurity on prenatal outcome haven't been routinely assessed; however, inclusion of these measures in an assessment of the family and social environment should be a part of future studies of WIC program outcome (as well as preconceptional assessment of women planning to become pregnant). As noted earlier by Laraia et al. (2006) food insecurity was a problem for up to 25% of WIC participants. Although most of these women were WIC and/or Food Stamp Program participants at the time of the survey, they still reported problems with obtaining adequate food; therefore, emergency food sources may also be needed by this group. However, food security is not usually screened for during routine prenatal care, and thus an opportunity to improve the nutritional status of these women may be missed.

Yet another issue that has arisen since the research studies examined here were conducted is the greater number of women entering pregnancy overweight and obese; this nutritional risk factor can have adverse effects on birth outcomes (Siega-Riz & Laraia, 2006). Women's food intake may be inappropriate and/or inadequate, leading to inadequate fetal nutrition. In addition, low income and minority women are more likely to be overweight or obese and there is also some evidence that food insecurity is also found in overweight and obese women (Adams, Grummer-Strawn, & Chávez, 2003; Townsend, Peterson, Love, Acterberg, & Murphy, 2001). While none of the studies reported here have addressed the significance of this issue and its potential effects on birth outcomes, future evaluation of evidence for WIC effectiveness needs to include examination of birth outcomes among overweight and obese women. A final research area of importance is the development and testing of theoretically- based conceptual models which describe the role of nutrition supplementation in reducing LBW. This very brief list of prenatal nutrition research suggestions presents multiple opportunities for a coordinated research approach to strengthen the available evidence for population based interventions to improve maternal nutrition as one step toward healthier birth outcomes.

Conclusion

Adequate maternal nutrition is well recognized as an important component of a successful pregnancy and positive birth outcomes; as such, prenatal nutrition supplementation has long been supported as a way to reduce LBW for women at nutritional risk. As the lead public nutrition supplementation program in the United States is WIC, a review of the research exploring the relationship between

WIC program participation (as well as prenatal nutrition supplementation in general) and LBW has been examined in this chapter. The overall conclusion is that there is a modest improvement in birth weight with nutrition supplementation, particularly in women at greater nutritional risk and in women who participate in WIC from early pregnancy. Across the studies, women not participating in WIC were more likely to deliver LBW infants than were participants. The WIC data are supported by other studies, particularly those from the Canadian Higgins Program. It is important to keep in mind however, that this evidence is tempered by the relatively small effect sizes in the studies, the age and quality of the program data, as well as other methodological concerns. Many of these issues were detailed in the chapter and recommendations to overcome some of the limitations have been presented. However, the data suggest that to the extent that racial/ethnic disparities in adverse pregnancy outcomes are attributed to low birthweight independent of preterm delivery, encouraging increased participation in WIC may play a future role in reducing racial/ethnic disparities.

The WIC program is clearly perceived as a significant resource that contributes to the overall health of pregnant women and their children. However, as discussed in this chapter, the last national review of the program was published in 1988. In today's climate of accountability and evidence-based decision making, it is essential to have relevant and timely information. It is due time for a comprehensive evaluation of WIC to fully document program need and outcomes, and to inform the allocation of shrinking resources at the state and national levels. While the costs of conducting such a large-scale and inclusive evaluation might be considered prohibitive, smaller, well-coordinated evaluation studies conducted at the local or regional levels are likely feasible, and would provide valuable information on the impact and value of WIC for policy-makers

Beyond the need for relevant and timely information to promote effective quality programs, changes within the WIC program itself as well as the context within which it is implemented, seem to demand a thorough program evaluation. Importantly, there are notable changes in the demographic and health profiles of the participant population due to: (a) Medicaid program expansions that include pregnant women with various levels of risk that likely brought some women into WIC who previously were not aware of the program, (b) changes in the racial and ethnic profile of the U.S. population (e.g., more Latinos and Asians), and, (c) increasing numbers of women with chronic conditions such as overweight or obesity, diabetes, and hypertension, as well as those who report food insecurity. These groups, while commonly at greater nutritional risk than the general U.S. population, have yet to be explicitly included in an assessment of the impact of supplemental food and nutrition programs on birth weight in the United States. Including adequate numbers of women from these groups to allow for stratified analysis in a well-designed evaluation of WIC will provide a more accurate and complete picture of program effects and potential.

Likewise, changes to the basic food package by USDA based on the recommendations of their commissioned report by the IOM will be implemented nationwide in August, 2009. These changes allow the WIC food package to meet the 2005 Dietary Guidelines and address current nutritional risk profiles. Specifically, the changes will reduce the fat, saturated fat, and sugar content of the prenatal food package. They will also increase access to fruits and vegetables and expand the selection of alternatives in the whole grain, dairy, and meat alternative/legume food groups (IOM, 2006). Impact evaluation of this change in the WIC food package as well as ongoing monitoring of population nutritional profiles seems essential to quality WIC programming.

Finally, the recent revisions of the IOM's 1990 guidelines for weight gain during pregnancy demonstrate the importance of monitoring and responding to changes in the social and economic environment related to nutrition in order to promote and attain optimal reproductive health outcomes. The revisions address the current dynamics of increasing BMI and gestational weight gain among all population subgroups in the United States, as well as the issue of women becoming pregnant at older ages. They also reflect our best understanding of the evidence for relationships between weight gain patterns and maternal and child health outcomes (IOM, 2009). Implementation of these guidelines may have implications for WIC programming in terms of an increase in the

numbers of eligible participants as well as the appropriateness of the WIC package for the dietary needs of specific groups of women.

Over the last decade, many state and local agencies have improved their data system capacities and developed new analytic strategies that allow the generation of baseline information reflecting WIC's programmatic and demographic changes, and provide a foundation for more robust evaluations. In addition, enhanced perinatal surveillance through PRAMS, the revised U.S. certificate of live birth, and technological advancements in data linkage, allow for more timely, accurate and complete information than had been available for the earlier studies reviewed in this chapter. National studies conducted with greater methodological rigor than many of those reviewed in this chapter, will require increased coordination and support from USDA and CDC surveillance systems. Such federal leadership has become increasingly important as demand for WIC services grow and program funding increasingly requires performance-based and client-centered outcome measures.

Given that the number of women at nutritional risk (and many at high-risk) is on the rise in the United States, it is critical to at least maintain, but preferably expand, funding levels for the WIC program as part of a larger national approach to improving pregnancy outcomes. Targeting the program to this ever-growing group of women, coupled with adequate data collection and enhanced analyses are key strategies aimed at developing a strong evidence-base for the role of supplemental foods in not only reducing adverse pregnancy outcomes overall but for reducing ethnic and racial disparities in such outcomes.

References

Adams, E. J., Grummer-Strawn, L., & Chávez, G. (2003). Food insecurity is associated with increased risk of obesity in California women. *Journal of Nutrition*, *133*(4), 1070–1074.

Ahluwalia, I. B., Hogan, V. K., Grummer-Strawn, L., & Colville, W. R. (1998). The effect of WIC participation on small-for-gestational-age births: Michigan, 1992. *American Journal of Public Health*, *88*(9), 1374–1377.

Atrash, H. K., Johnson, K., Adams, M., Cordero, J. F., & Howse, J. (2006). Preconception care for improving perinatal outcomes: the time to act. *Maternal and Child Health Journal*, *10*, S3–S11.

Avruch, S., & Cackley, A. P. (1995). Savings achieved by giving WIC benefits to women prenatally. *Public Health Reports*, *110*(1), 27–34.

Besharov, D. J., & Germanis, P. (2001). *Rethinking WIC: An evaluation of the Women, Infants, and Children Program*. Washington, D.C.: The American Enterprise Institute Press.

Bitler, M. P., & Currie, J. (2005). The effects of WIC on pregnancy and birth outcomes. *Journal of Policy Analysis and Management*, *24*(1), 73–91.

Briley, C., Flanagan, N. L., & Lewis, N. (2002). In-home prenatal nutrition intervention increased dietary iron intakes and reduced low birthweight in low-income African-American women. *Journal of the American Dietetic Association*, *102*(7), 984–987.

Brown, J. E., & Khan, E. S. B. (1997). Maternal nutrition and the outcome of pregnancy, a renaissance in research. *Clinics in Perinatology*, *24*(2), 433–449.

Buescher, P. A., & Horton, S. J. (2000). *Prenatal WIC participation in relation to low birth weight and Medicaid infant costs in North Carolina – a 1997 Update*, Raleigh, N.C.: Center for Health Informatics and Statistics, North Carolina Department of Health and Human Services.

Buescher, P. A., Larson, L. C., Nelson, M. D., & Lenihan, A. J. (1993). Prenatal WIC participation can reduce low birthweight and newborn medical costs: A cost-benefit analysis of WIC participation in North Carolina. *Journal of the American Dietetic Association*, *93*, 163–166.

Caan, B., Horgen, D. M., Margen, S., King, J. C., & Jewell, N. P. (1987). Benefits associated with WIC supplemental feeding during the interpregnancy interval. *American Journal of Clinical Nutrition*, *45*, 29–41.

Ceesay, S. M., Prentice, A., Cole, T., Foord, F., Weaver, L. T., Poskitt, E. M., & Whitehead, R. G. (1997). Effects on birth weight and perinatal mortality of maternal dietary supplements in rural Gambia: 5 year randomised trial. *British Medical Journal*, *315*, 786–790.

Centers for Disease Control and Prevention. (2008). *CDC's Pediatric and Pregnancy Nutrition Surveillance System*. Retrieved January, 2008 from http://www.cdc.gov/PEDNSS/pnss_tables/index.htm

Collins, T. R., Demellier, S. T., Leeper, J. D., & Milo, T. (1985). Supplemental food program: Effects on health and pregnancy outcome. *Southern Medical Journal*, *78*(5), 551–5.

de Onis, M., Villar, J., & Gülmezoglu M. (1998). Nutritional interventions to prevent intrauterine growth retardation: Evidence from randomized controlled trials. *European Journal of Clinical Nutrition*, *52*:S1, S83–S93.

Desjardins, E., & Hardwick, D. (1999). How many visits by health professionals are needed to make a difference in low birthweight? A dose-response study of the Toronto Healthiest Babies Possible Program. *Canadian Journal of Public Health*, *90*(5), 224–228.

Devaney, B. (1992). *Very low birthweight among Medicaid newborns in five states: The effects of prenatal WIC participation*, Alexandria, VA: U.S. Department of Agriculture, Food and Nutrition Service, Office of Analysis and Evaluation.

Devaney, B., Bilheimer, L., & Schore, J. (1990). *The savings in Medicaid costs for newborns and their mothers from prenatal participation in the WIC program*, Vol. 1. Alexandria, VA: U.S. Department of Agriculture, Food and Nutrition Service, Office of Analysis and Evaluation.

Devaney, B., Bilheimer, L., & Schore, J. (1991). *The savings in Medicaid costs for newborns and their mothers from prenatal participation in the WIC program*, Vol. 2. Alexandria, VA: U.S. Department of Agriculture, Food and Nutrition Service, Office of Analysis and Evaluation.

Dodds, J., & Kaufman, M. (1991). *Personnel in public health nutrition for the 1990s*, Washington, D.C.: The Public Health Foundation.

Downs, S. H., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality of both randomized and non-randomized studies of health care interventions. *Journal of Epidemiology and Community Health*, *52*, 377–387.

Dubois, S., Coulombe, C., Pencharz, P., Pinsonneault, O., & Duquette, M. P. (1997). Ability of the Higgins Nutrition Intervention Program to improve adolescent pregnancy outcome. *Journal of the American Dietetic Association*, *97*, 871–878.

Egan, M. C. (1994). Public health nutrition: A historical perspective. *Journal of American the Dietetic Association*, *94*(3), 298–304.

Fall, C. H. D., Yajnik, C. S., Rao, S., Davies, A. A., Brown, N., & Farrant H. J. W. (2003). Micronutrients and fetal growth. *Journal of Nutrition*, *133*, 1747S-1756S.

Fox, M. K., Hamilton, W., & Lin, B-H. (2004). Effects of food assistance and nutrition programs on nutrition and health, volume 4, executive summary of the literature review. *Food assistance and nutrition research report number 19–4*. Washington D.C.: Food and Rural Economics Division, Economic Research Service, U.S. Department of Agriculture.

Frisbie, W. P., Biegler, M., de Turk, P., Forbes, D., & Pullum, S. G. (1997). Racial and ethnic differences in determinants of intrauterine growth retardation and other compromised birth outcomes. *American Journal of Public Health*, *87*(12), 1977–83.

Godfrey, K. M., & Barker, D. J. P. (2001). Fetal programming and adult health. *Public Health Nutrition*, *4*(2b), 611–624.

Gordon, A., & Nelson, L. (1995). *Characteristics and outcomes of WIC participants and nonparticipants: Analysis of the 1988 National Maternal Health Survey*. Alexandria, VA: USDA.

Hamelin A. -M., Beaudry M., & Habicht J. -P. (2002). Characterization of household food insecurity in Québec: Food and feelings, *Social Science & Medicine*. *54*, 119–132.

Higgins, A. C. (1976). Nutritional status and the outcome of pregnancy. *Journal of the Canadian Dietetic Association*, *37*, 17–35.

Higgins, A. C., Crampton, E. W., & Moxley, J. E. (1972). Nutrition and the outcome of pregnancy. *Endocrinology.* Proceedings of the Fourth International Congress of Endocrinology, Washington, Excerpta Medica, Amsterdam.

Higgins, A. C., Moxley, J. E., & Pencharz, P. B. (1989). Impact of the Higgins Nutrition Intervention Program on birth weight: A within-mother analysis. *Journal of the American Dietetic Association*, *89*(9), 1097–1103.

Institute of Medicine. (2003). *Estimating eligibility and participation for the WIC program*. Final Report Committee on National Statistics, Division of Behavioral and Social Sciences and Education (M. Ver Ploeg & D. Betson, Eds.). Washington, D.C.: National Academy Press.

Institute of Medicine, Committee to Study the Prevention of Low Birthweight (1985). *Preventing low birthweight*. Washington, D.C.: National Academy Press.

Institute of Medicine, Food and Nutrition Board. (1990). *Nutrition during pregnancy: Part 1 weight gain*, Subcommittee on Nutritional Status and Weight Gain during Pregnancy. Washington, D.C.: National Academy Press.

Institute of Medicine, Food and Nutrition Board. (1996). *WIC nutrition risk criteria: A scientific assessment*. Committee on Scientific Evaluation of WIC Nutrition Risk Criteria. Washington, D.C.: National Academy Press.

Institute of Medicine, National Research Council. (2001). *Estimating eligibility and participation for the WIC program: Phase I report*. Committee on National Statistics, Division of Behavioral and Social Science and Education. M. Ver Ploeg & D.M. Benson (Eds.), Washington, D.C.: National Academy Press.

Institute of Medicine, Food and Nutrition Board. (2002a). *Dietary risk assessment in the WIC program*. Committee on Dietary Risk Assessment in the WIC Program, Washington, D.C.: National Academy Press.
Institute of Medicine, Food and Nutrition Board. (2002b). *Dietary reference intakes for energy and the macronutrients, carbohydrates, fiber, fat, and fatty acids*, Washington, D.C.: National Academy Press.
Institute of Medicine, Food and Nutrition Board. (2006). *WIC food packages, time for a change*, Committee to Review the WIC Food Packages, Washington, D.C.: National Academy Press.
Institute of Medicine and National Research Council. (2009), *Weight Gain During Pregnancy: Reexamining the Guidelines*, Washington, D.C.: The National Academic Press.
Joyce, T., Gibson, D., & Colman, S. (2004). *The changing association between prenatal participation in WIC and birth outcomes in New York City*. National Bureau of Economic Research Working Paper Series, paper 10796, http://www.nber.org/papers/w10796 September, 2004.
Joyce, T., Gibson, D., & Colman, S. (2005). The changing association between prenatal participation in WIC and birth outcomes in New York City. *Journal of Policy Analysis and Management*, 24 (4):661–685.
Joyce, T., Racine, A., & Yunzal-Butler, C. (2008). Reassessing the WIC effect: Evidence from the Pregnancy Nutrition Surveillance System. *Journal of Policy Analysis and Management*, *27*(2), 277–303.
Kent, D., & Hayward, R. (2007). When averages hide individual differences in clinical trials. *American Scientist*, *95*, 60–68.
King, J. C. (2003). The risk of maternal nutritional depletion and poor outcomes increases in early or closely spaced pregnancies. *Journal of Nutrition*, *133*, 1732S–1736S.
Kotch, J. (2005). *Maternal and child health* (2nd ed.). Sudbury, MA: Jones & Bartlett.
Kowaleski-Jones, L., & Duncan, G.J. (2002). Effects of participation in the WIC program on birthweight: Evidence from the National Longitudinal Survey of Youth. Special Supplemental Nutrition Program for Women, Infants, and Children. *American Journal of Public Health*, *92*(5), 799–804.
Kramer, M. S. (1993). Effects of energy and protein intakes on pregnancy outcome: An overview of the research evidence from controlled clinical trials. *American Journal of Clinical Nutrition*, *58*, 627–635.
Kramer, M. S. (2003). The epidemiology of adverse pregnancy outcomes: An overview. *Journal of Nutrition*, *133*,1592S–1596S.
Kramer, M. S., & Kakuma, R. (2006). Energy and protein intake in pregnancy. Cochrane Database of Systematic Reviews 2003, Issue 4. Art. No.: CD000032. DOI: 10.1002/14651858.CD000032.
Kramer, M. S., & Kakuma, R. (2009). Energy and protein intake in pregnancy. Cochrane Database of Systematic Reviews 2003, Issue 4. Art. No.: CD000032. DOI: 10.1002/14651858.CD000032.
Laraia, B. A., Siega-Riz, A. M., Gunderson, C., & Dole, N. (2006). Psychosocial factors and socioeconomic indicators are associated with household food insecurity among pregnant women. *Journal of Nutrition*, *136*, 177–182.
Lazariu-Bauer, V., Stratton, H., Pruzek, R., & Woelfel, M. L. (2004). A comparative analysis of effects of early vs. late prenatal WIC participation on birth weight: NYS, 1995. *Maternal and Child Health Journal*, *8*(2), 77–86.
Long, V.A., Martin, T., & Janson-Sand, C. (2002). The great beginnings program: Impact of a nutrition curriculum on nutrition knowledge, diet quality, and birth outcomes in pregnant and parenting teens. *Journal of the American Dietetic Association*, *102*(3), S86–S89.
Lu, M. C., & Halfon, N. (2003). Racial and ethnic disparities in birth outcomes: A life-course perspective. *Maternal and Child Health Journal*, *7*(1), 13–30.
Maternal Nutrition and Agnes Higgins Analysis. (n.d.). Retrieved July 11, 2006 from http://library.med.utah.edu/nmw/Tutorial4/higgins.html.
Merialdi, M., Carroli, G., Villar, J., Abalos, E., & Gülmezoglu, A. M. (2003). Nutritional interventions during pregnancy for the prevention or treatment of impaired fetal growth: An overview of randomized controlled trials. *Journal of Nutrition*, *133*, 1626S–1631S.
Metcoff, J., Costiloe, P., Crosby, W. M., Dutta, S., Sandstead, H. H., Milne, D., et al. (1985). Effect of food supplementation (WIC) during pregnancy on birth weight. *American Journal of Clinical Nutrition*, *41*(5), 933–947.
Nord, M., Andrews, M., & Carlson, S. (2006). Household Food Security in the United States, 2005, *USDA Economic Research Service Rport*, Alexandria, VA: USDA.
Oliveira, V., Racine, E., Olmsted, J., & Ghelfi, L.M. (2002). The WIC program: Background, trends, and issues. *Food assistance and nutrition research report, number 27*. Washington, D.C: Food and Rural Economics Division, Economic Research Service, U.S. Department of Agriculture.
Olson C.M. (1999). Nutrition and health outcomes associated with food insecurity and hunger. *Journal of Nutrition*, *129*, 521–524.
Osrin, D., Vaidya, A., Shrestha, Y., Baniya, R.B., Mahandhar, D.S., Adhikari, R. K., Rilteau, S., Tomkins, A., et al. (2005). Effects of antenatal multiple micronutrient supplementation on birthweight and gestational duration in Nepal: A double-blind randomised controlled trial. *The Lancet*, *365*, 955–962.
Owen, A. L., & Owen, G. (1997). Twenty years of WIC: A review of some effects of the program. *Journal of the American Dietetic Association*, *97*, 777–782.

Perry, G. S., Yip, R., & Zyrkowski, C. (1995). Nutritional risk factors among low-income pregnant US women: The Centers for Disease Control and Prevention (CDC) Pregnancy Nutrition Surveillance System, 1979 through 1993. *Seminars in Perinatology*, *19*(3), 211–221.

Rasmussen, K. M. (2001). Is there a causal relationship between iron deficiency anemia and weight at birth, length of gestation and perinatal mortality? *Journal of Nutrition*, *131*, 590S–603S.

Reichman, N. E., & Teitler, J. O. (2003). Effects of psychosocial risk factors and prenatal interventions on birth weight: Evidence from New Jersey's HealthStart Program. *Perspectives on Sex and Reproductive Health*, *35*(3), 130–137.

Revision in the WIC Food Packages, Interim Rule; Special Supplemental Nutrition Program for Women, Infants, and Children (WIC), 72 Federal Register 234 (5 February 2008).

Rosen, G. (1985). *A history of public health*, Baltimore: The Johns Hopkins University Press.

Rush, D. (1981). Nutritional services during pregnancy and birthweight: A retrospective matched pair analysis, *Canadian Medical Association Journal*, *125*, 567–576.

Rush, D. (2001a). Maternal nutrition and perinatal survival. *Journal of Health, Population and Nutrition*, *19*(3), S217–S264.

Rush, D. (2001b). Maternal nutrition and perinatal survival. *Nutrition Reviews*, *59*(10), 315–326.

Rush, D., Alvir, J., Kenny, D., Johnson, S., & Horvitz, D. (1988a). Historical study of pregnancy outcomes. *American Journal of Clinical Nutrition*, *48*, 412–428.

Rush, D., Sloan, N.L., Leighton, J., Alvir, J., Horvitz, D. G., Seaver, W. B., et al. (1988b). Longitudinal study of pregnant women, the National WIC evaluation: evaluation of the Special Supplemental Food Program for Women, Infants, and Children. *American Journal of Clinical Nutrition*, *48*, 439–483.

Rush, D., Stein, Z., & Susser, M. (1980). A randomized controlled trial of prenatal nutritional supplementation in New York City. *Pediatrics*, *65*, 683–697.

Schramm, W. F. (1985). WIC prenatal participation and its relationship to newborn Medicaid costs in Missouri: A cost/benefit analysis. *American Journal of Public Health*, *75*, 851–857.

Siega-Riz, A. M., & Laraia, B. (2006). The implications of maternal overweight and obesity on the course of pregnancy and birth outcomes. *Maternal and Child Health Journal*, *10*, S153-S156.

Sloan, N. L., Lederman, S. A., Leighton, J., Himes, J. H., & Rush, D. (2001). The effect of prenatal dietary protein intake on birth weight. *Nutrition Research*, *21*, 129–139.

Stockbauer, J. W. (1986). Evaluation of the Missouri WIC program: Prenatal components. *Journal of the American Dietetic Association*, *86*(1), 61–67.

Stockbauer, J. W. (1987). WIC prenatal participation and its relation to pregnancy outcomes in Missouri: A second look. *American Journal of Public Health*, *77*(7), 813–818.

Susser, M. (1991). Maternal weight gain, infant birth weight, and diet: Causal sequences. *American Journal of Clinical Nutrition*, *53*, 1384–1396.

Susser, M., & Stein, Z. (1982). Third variable analysis: Application to causal sequences among nutrient intake, maternal weight, birthweight, placental weight and gestation. *Statistics in Medicine*, *1*, 105–120.

Sweeney, C., Smith, H., Foster, J. C., Place, J.C., Specht, J., Kochenour, N. K., et al. (1985). Effects of a nutrition intervention program during pregnancy, maternal data phases 1 and 2. *Journal of Nurse-Midwifery*, *30*(3), 149–158.

The Food Research and Action Center, (2005). *WIC in the states: Thirty-one years of building a healthier america, report summary.* Retrieved August, 2006 from www.frac.org/WIC/2004_Report/Full_Report.pdf.

Townsend, M. S., Peterson, J., Love, B., Achterberg, C., & Murphy, S. P., (2001). Food insecurity is positively related to overweight in women. *Journal of Nutrition*, *131*(6), 1738–1745.

U.S. Department of Agriculture, Food and Nutrition Service. (2005). *WIC at a glance*. Retrieved June 20, 2008 from http://www.fns.usda.gov/wic/aboutwic/wicataglance.htm.

U.S. Department of Agriculture, Food and Nutrition Service Office of Analysis, Nutrition Evaluation (2006). *WIC program coverage: How many eligible individuals participated in the Special Supplemental Nutrition Program for Women, Infants and Children (WIC): 1994–2003*, Alexandria, VA: USDA.

U.S. Department of Agriculture, Food and Nutrition Service. (2007a). *Food distribution programs history and background*. Retrieved June 20, 2008 from http://www.fns.usda.gov/fdd/aboutfd/history.htm.

U.S. Department of Agriculture, Food and Nutrition Service, Office of Analysis, Nutrition Evaluation. (2007b). *WIC participant and program characteristics 2006, WIC-06-PC*, by S. Bartlett, E. Bobronnikov, M. Mendelson, et al. Alexandria, VA: USDA.

U.S. Department of Agriculture, Food and Nutrition Service. (2009a). *A short history of SNAP*. Retrieved March 6, 2009 from http://www.fns.usda.gov/snap/rules/Legislation/about.htm

U.S. Department of Agriculture, Food and Nutrition Service. (2009b). *Legislative history food distribution programs*. Retrieved March 6, 2009 from www.fns.usda.gov/FDD/aboutfd/fd_history.pdf

U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion. (2000). *Healthy people 2010* (2nd ed.) Washington, D.C.: Government Printing Office.

U.S. General Accounting Office. (1989). *Food assistance, the National WIC evaluation: reporting and follow-up issues*, Report to Congressional requesters, GAO/RCED-90-3, Washington, D. C.: Government Printing Office.

U.S. General Accounting Office. (1992). *Early intervention. federal investments like WIC can produce savings*. GAO/HRD-92-18. Washington, D.C.: Government Printing Office.

Ver Ploeg, M., & Betson, D. (Eds.). (2001). *Estimating eligibility and participation for the WIC Program*. Phase I report. Committee on National Statistics, Division of Behavioral and Social Sciences and Education, National Research Council. Washington, D.C.: National Academy Press.

Victora, C. G., Habicht, J. -P ., & Bryce, J. (2004). Evidence-based public health: Moving beyond randomized trials. *American Journal of Public Health*, *94*(3), 400–405.

Villar, J., Merialdi, M., Gülmezoglu, A. M., Abalos, E., Carroli, G., Kulier, R., et al. (2003). Nutritional interventions during pregnancy for the prevention or treatment of maternal morbidity and preterm delivery: An overview of randomized controlled trials. *Journal of Nutrition*, *133*, 1606S–1625.

White House Conference on Food, Nutrition and Health, Final Report. (1969). Washington, D.C.: U.S. Government Printing Office.

Wigda, A. C., & Lewis N. M. (1999). Defined, in-home, prenatal nutrition intervention for low-income women. *Journal of the American Dietetic Association*, *99*(9), 1058–1062.

Chapter 15
Group Prenatal Care and Doula Care for Pregnant Women

Susan C. Vonderheid, Rieko Kishi, Kathleen F. Norr, and Carrie Klima

There is widespread recognition of the need for innovation to improve the quality of care for childbearing women and to reduce the long-standing racial/ethnic disparities in pregnancy outcomes (Carroli, Rooney, & Villar, 2001; Hughes & Runyan, 1995; Institute of Medicine, 2003; Lu & Halfon, 2003; Misra, Grason & Weisman, 2000). This integrated review will examine two promising innovations for childbearing women, group prenatal care and doula care.

Prenatal care remains one of the most important public health interventions for women and children, providing a unique window of opportunity to assess women's health status and to intervene through both effective treatment and the promotion of healthier behaviors [Public Health Service (PHS), 1989]. Group prenatal care is innovative because it fundamentally alters the longstanding model of individual prenatal care for pregnant women.

Whereas prenatal care focuses on pregnancy, doula care traditionally focuses on childbirth. The birth experience is an important event in the lives of women and their families. Yet, many women are not satisfied with their birth experience, often related to the increased medicalization of intrapartum care (see Declercq et al., Chapter 16). Doula care is an enhancement to routine intrapartum care that provides non-medical care including physical, emotional, and social support. More recently, doula care has expanded into the prenatal and postpartum periods.

This chapter reviews both group prenatal care and doula care and their relationship to selected maternal and infant outcomes. This chapter also considers the potential of these two interventions for reducing racial/ethnic disparities in perinatal outcomes.

Group Prenatal Care

Review Method

A search of Medline, CINAHL, PsycINFO, Popline, Web of Science, Cochrane Library, OCLC First Search and Academic Search Premier conducted between June 2006 and August 2007 produced 12 articles about group prenatal care using the following key words: "group prenatal care," "group visits," "group care," "prenatal care," "antenatal care," "Centering Pregnancy," and combinations of these key words. The computer search was limited to research, female participants, and English; no date limits were included. Key outcomes of interest for this review were preterm delivery and/or

S.C. Vonderheid (✉)
846 UIC College of Nursing, Department of Women, Children and Family Health Science,
M/C 802, 845 S Damen Avenue, Chicago, IL 60612-7350, USA
e-mail: vonde@uic.edu

A. Handler et al. (eds.), *Reducing Racial/Ethnic Disparities in Reproductive and Perinatal Outcomes: The Evidence from Population-Based Interventions*, DOI 10.1007/978-1-4419-1499-6_15,

low birth weight. Therefore, only reports of group prenatal care interventions that included outcome data about preterm delivery and/or low birth weight were included in this review. Additionally, a manual search was performed of the references of the studies obtained based on the computer search and generated three additional references. As a result, 15 articles were found. Six out of the 15 articles examined preterm delivery and/or low birth weight (Ford et al., 2002; Grady & Bloom, 2004; Hoyer, Jacobson, Ford, & Walsh, 1994; Ickovics et al., 2003; Ickovics et al., 2007; Rising, 1998) and were included in this review. These six studies are summarized in the Table 15.1. The results of the quality assessment are summarized in Table 15.2. Inter-rater reliability was 1.00.

Description of Group Prenatal Care

Group prenatal care is a model that provides all or mostly all of the components of a usual individual prenatal visit within a group setting. Common components of group prenatal care include provider assessment, self-care training and activities, allotted time for education and support, enhanced health promotion content, and peer support (Ford et al., 2002; Hoyer et al., 1994; Rising, 1998). This model of care brings together a group of 5–12 pregnant women who are expected to give birth around the same time. The first prenatal visit is individual care followed by group visits that begin in the early second trimester (Ickovics et al., 2003, 2007; Rising, 1998). Activities in a group visit generally begin with physical assessments conducted over 30min followed by health education for 60–90min. Women that become medically high-risk might be transferred to individual care for more intensive medical attention but might also continue attending group visits. In early reports of group prenatal care, providers were nurse practitioners and/or nurse-midwives (Fullar, Lum, Sprik, & Cooper, 1988; Rising, 1998). More recent studies of group prenatal care used advanced practice nurses and physicians to provide group care (Ford et al., 2002; Ickovics et al., 2003, 2007).

Given the large amount of time devoted to health education and peer support, it is hypothesized that group prenatal care might improve maternal psychosocial and behavioral outcomes. In turn, improvements in these maternal outcomes might reduce adverse perinatal outcomes, especially among women at risk due to psychosocial factors such as low-income or minority race/ethnicity. Based on the structure of group visits, advantages might include: greater breadth and depth of health promotion content discussed, improved knowledge and self-management skills, greater problem solving ability through sharing of common life experiences, enhanced self-efficacy and empowerment, increased social support, and the potential of participants to develop new social norms for engaging in healthier behaviors. Studies are needed to determine the exact pathways or mechanisms that might contribute to improvements in infant outcomes as a result of participation in group prenatal care.

Early reports of group prenatal care by Fullar and colleagues (1988) and Hoyer et al. (1994) focused on meeting the unique developmental needs of pregnant adolescents. These studies based group prenatal care on developmental theory, social cognitive theory, and peer support. In addition to a group visit lasting 1.5–2 hours, adolescents were paired and conducted physical assessment activities on one another typically reserved for health care professionals including blood pressure, weight, fundal height, fetal heart tones, fetal position and recording results in the record. Verification of assessments was conducted by a health care provider.

CenteringPregnancy®, the most recent and widespread attempt to provide group prenatal care, integrates several conceptual frameworks. Rising (1998) developed CenteringPregnancy® based in part on her experiences with the Childbearing-Childrearing Center at the University of Minnesota, an early attempt at a group model of prenatal care for pregnant females of all ages and their families. CenteringPregnancy® is based upon a midwifery philosophy of care, highlighting women-centered care and the empowerment of women throughout their reproductive lives. Feminism, with its emphasis upon participatory decision-making and equalization of the power inherent in provider–client relationships, is reflected in the symmetrical relationships and access to information found in the group model. Social support theory is

Table 15.1 Major outcomes associated with studies of group prenatal care: birth weight, LBW and preterm births

Author/ study design/ study type	Description of intervention	Sample and setting	Address disparities (Y/N)	Data collection	Outcome	Key findings related to intervention effectiveness	Strengths	Caveats/ biases	Findings support the intervention? Yes/No for which populations?
Hoyer et al. (1994) RCT pilot, Letter to the Editor	I: NP managed peer groups of 6–8 assigned in pairs with similar developmental level Visits 1.5 h: 30 min for assessment (teen pairs performed BP, weights, urine testing, fetal heart tones, fundal heights on one another) and 60 min education. Personal records maintained by teens. Clinic records maintained by provider C: Routine care by MD or NP in comprehensive clinic with nutrition counseling, social services, lab. In waiting room, I C women had access to nutritional snacks and videotapes about teen pregnancy, birth and infant care	HMO clinic in large metropolitan area *N*=65, 14–20 years (median 17.9) AA, low-middle income, 97% unmarried, 7th grade – 2 years college education (median 11th grade completed), 36% employed, 50% planned to finish school, 81% FOB involved Inclusion: ≤20 years, 1st–2nd trimester, low-mod risk, vaginal birth anticipated, singleton No statistically significant differences between study groups	No	Intake (before random assignment) and post delivery Post-test data collection not reported, such as medical record vs. self-report for delivery outcomes	GA, PT labor	GA: Ranged from 34–42 weeks total sample, I mean 39 weeks, C mean 37 weeks (statistical difference not reported) PT labor: <14% – total sample only reported (statistical difference not reported) and suppressed in both groups	Clear descriptions of framework, intervention Random assignment to study group	Intervention fidelity not discussed Potential confounders not controlled for, such as provider type Data collection not reported Data reporting not complete: sample size for study groups and outcomes Small sample, power, intent-to-treat analysis, and attrition not reported Potential for cross-contamination in waiting room	No

(continued)

Table 15.1 (continued)

Author/ study design/ study type	Description of intervention	Sample and setting	Address disparities (Y/N)	Data collection	Outcome	Key findings related to intervention effectiveness	Strengths	Caveats/ biases	Findings support the intervention? Yes/No for which populations?
Rising (1988). One group post-test (with comparison to IC on general clinic data available for ER use, cesarean births and Apgar scores) (published article)	*CP pilot* Traditional IC first visit followed by 10 group visits initiated between 12 and 16 weeks with 8–12 low medical risk women Each group visit was 90 min with traditional prenatal risk assessments conducted within first 20 min followed by group discussion and education. Women weighed themselves, took their blood pressure, determined their GA, and made chart entries. Assessments took place in group space rather than exam room	*N*=111 women in 13 groups with an average no. of 8.75 women per group over 15-months; 3 out of 13 groups were adolescent 111/450 births over 15 months=24.6% of women in clinic Ethnically diverse, primarily Medicaid. Maternal age: 25% <20 years, 64% 20–29 years, 68% primiparous, 70% initiated care 1st trimester and 4% initiated ≥21 weeks	No	Medical records	BW, GA, ER use in 3rd trimester, cesarean births, Apgar scores	PT <37 weeks: 4.5% (no comparison data) LBW: 5.4% (no comparison data) ER use in 3rd trimester: Out of 62 women using ER, 26% (*n*=5) were CP and 74% (*n*=14) were IC Cesarean births: 12.6% CP and 13.5% IC Apgar score <7: <1% (*n*=1) CP 2% (*n*=6) IC. (Statistical analysis to compare CP and IC not reported)	Clear descriptions of framework, intervention	Intervention fidelity not discussed Self-selection Potential confounders not controlled for, such as provider type Comparison data for BW and GA/PT births not reported Statistical analysis to compare CP and IC not reported	No

	Handouts and worksheets completed during beginning of visit facilitated discussion								
Ford et al. (2002). RCT (published article)	Mastery modeling peer-support group intervention I: Received care in groups of 6–8 and learned to perform routine prenatal assessments with a peer partner during visits and educational content provided in groups C: Received individual prenatal care in the same clinics All participants given same hard copy educational material	*Intake* *N*=282 *n*=165 I *n*=117 C BW analysis *N*=255 Sample size for study groups not described Pregnant adolescents from 5 clinics located primarily in Detroit, MI Primarily AA unmarried <21 years of age, ≤high school education, primip- and multiparous	No However, authors stated that "use of this program among urban adolescent mothers may help reduce health disparities" (p. 38)	Self administered forms (Intake), medical records	LBW	Trend showed that experimental group had lower rate of LBW (6.6% vs. 12.5%, *p*=0.18)	Participants randomly assigned to study group Clear descriptions of framework, intervention, study population Statistical analysis controlled for age and parity	Intervention fidelity not discussed Potential confounders not controlled for, such as provider type Power analysis not discussed	Yes, for urban unmarried, AA, <21 years of age

(continued)

Table 15.1 (continued)

Author/ study design/ study type	Description of intervention	Sample and setting	Address disparities (Y/N)	Data collection	Outcome	Key findings related to intervention effectiveness	Strengths	Caveats/ biases	Findings support the intervention? Yes/No for which populations?
Ickovics et al. (2003). Matched cohort: Prospective intervention group matched one-to-one by categories of age, race, or ethnicity, and parity (published article)	I: CP groups of up to 12 women with similar delivery date, met over 10 sessions focused on education and skill building specific to prenatal and postpartum care. CP promoted self care activities and taught women to actively track changes associated with pregnancy C: Received routine IC at same clinics	*N*=458 *n*=229 I *n*=229 C 80% AA 15% Latina 5% white Maternal age mean=21.6, standard deviation=4.2 47% Nulliparous 0.25% had prior term birth Participants from two affiliated hospital based, public clinics in Atlanta, GA and one hospital based public clinic in New Haven, CT Patients who entered PNC < 24 weeks' GA and voluntarily received CP were included. A comparison group of patients who received IC at the same clinics	Yes. "If prenatal care provides a window of opportunity for health promotion and risk reduction, then group prenatal care might be a useful tool in the reduction of adverse prenatal outcomes, reducing racial or ethnic disparities." (p. 1055)	Medical record abstraction	BW, PT births	BW higher for CP vs. IC (3, 228 vs. 3, 159g, *p*=.01 No difference in PT between CP vs. IC (9.2% vs. 9.6%, *p*=0.83) Among term infants, no significant difference in BW (3,313 vs. 3,283g, *p*=0.54) Among PT infants, CP infants were larger than IC infants (2,398 vs. 1,990g, *p*<0.05)	Each clinic implemented CP least 6 months before start of study Clear descriptions of framework, intervention Randomized selection of comparison group conducted by a computer program to reduce potential bias Analysis based on intent-to-treat Power analysis reported	Intervention fidelity not discussed Potential confounders not controlled for, such as provider type Self-selection into study group. Those who selected group PNC might have had better health status or behaviors	Yes, for low medical/ psycho-logical risk, low SES, AA women

						Descriptive analyses indicated that CP infants were less likely than IC infants to be LBW (16 vs. 23 infants); very LBW (3 vs. 6); early PT births (less than 33 weeks; 2 vs. 7 infants)			
Grady & Bloom (2004). Demonstration project comparing CP with retrospective and concurrent IC groups (published article)	In March 2001 the CP program was incorporated as the model of care to be used for all adolescents receiving prenatal care at the Teen Pregnancy Center (TPC) Adolescents who preferred traditional prenatal care were scheduled with a nurse practitioner or nurse midwife within the IC clinic system CP groups consist of 8–12 with an estimated due date within a 6-week period of time	*N*=501 Adolescents attending urban hospital based clinics in St. Louis, MO CenteringPregnancy TPC at Barnes Jewish Hospital (BJH): *n*=124, 93.6% AA, mean age 15.85, attended groups from March 2001 through April 2003, and gave birth at BJH. *2001 Comparison Group* *n*=144, 90.3% AA, age ≤17 yrs (mean age 16.5), who gave birth at BJH in 2001 (excluding those receiving no PNC and the teens in CP) *1998 Comparison Group*	No	Medical record (postpartum progress notes at discharge and on the labor and delivery nursing assessment) Hospital databases for 1998 and 2001 data Only BW, gestational age and delivery type data could be collected on 2001 comparison group	PT births, LBW	CP had lower rates of PT birth (10.5% vs. 25.7% and 23.2%) and LBW (8.9% vs. 22.9 and 8.3%) $p<0.05$	Clear descriptions of framework, intervention 2001 hospital database analysis excluded adolescents receiving no prenatal care	Intervention fidelity not discussed Self-selection into study groups. Those who selected group PNC might have had better health status or behaviors 1998 hospital database analysis included adolescents receiving no prenatal care Potential confounders not controlled for, such as provider type	Yes, for urban AA adolescents ≤17 years of age

(continued)

Table 15.1 (continued)

Author/ study design/ study type	Description of intervention	Sample and setting	Address disparities (Y/N)	Data collection	Outcome	Key findings related to intervention effectiveness	Strengths	Caveats/ biases	Findings support the intervention? Yes/No for which populations?
	Groups begin between 12 and 18 weeks gestation and continue every 2 weeks throughout pregnancy for a total of 12 sessions Sessions focus on assessment, education, and group support directed toward self-care responsibility	*n*=233, 85% AA, age ≤17 years (mean age 16.3), who gave birth at BJH in 1998						Intent-to-treat not used in analysis. Only participants completing CP were included in analysis	
Ickovics et al., (2007), RCT (published article)	3 study groups: CP, CPE with HIV/STD prevention content, and IC CP and CPE group visits had mean of 8 participants	*N*=1, 047 *n*=335 CP *n*=318 CPE *n*=394 IC Maternal age mean 24, 49% aged 14–19 years. ≤HS 76%, 26% dropped out of school, 31% employed, 47% received financial assistance from family/ partner, 22% received public assistance	Yes	Medical record	BW, PT births	Analysis combined CP and CPE into one group and compared to IC	Clear descriptions of framework, intervention. Blocked randomized controlled design, stratified by site, EDC, and study group. Assigned using computerized randomization after eligibility completed.	Intervention fidelity not discussed Provider type not controlled for as a potential confounder	Yes, for low medical risk, low SES AA and Latina women

Participants from large obstetrics clinics in Atlanta, GA and New Haven, CT Obstetric patients who entered PNC <24 weeks' GA, <25 maternal age, not high-medical risk, English or Spanish speaking, and voluntarily received prenatal care in a group setting were included	Group PNC had lower rate of PT birth compared to IC (9.8% vs. 13.8%, p=0.045). Among AAs, group PNC had even lower PT birth rate compared to IC (10% vs. 15.8%, p=0.02) No differences in GA, BW, and rates of LBW and SGA	Analysis controlled for many potential confounders. Analysis based on intent-to-treat and power reported. Loss to follow up described – no study group differences Medical record abstractors were blinded to study group assignment No adverse effects reported	

RCT randomized controlled trial; *AA* African-American; *HS* high school graduate or equivalency degree: *I* intervention study group; *C* control study group; *LBW* low birth weight; *BW* birth weight; *PT* preterm birth; *GA* gestational age; *CP* CenteringPregnancy group visit model of prenatal care; *CPE* CenteringPregnancy group visit model of care enhanced with HIV/STD prevention content; *IC* individual visit prenatal care; *NP* nurse practitioner; *MD* medical doctor; *PNC* prenatal care; *SES* socioeconomic status; *ER* emergency room; *SGA* small for gestational age; *EDC* estimated date of confinement; *FOB* father of the baby

Table 15.2 Quality rating and suitability of studies associated with group prenatal care intervention

Author	Reporting	External validity	Internal validity-bias	Internal validity-confounding	Power	Total quality score ≤14=poor, 15–19=fair, ≥20=good	Suitability of study to assess effectiveness
Hoyer et al. (1994)	3	2	2	3	0	10	Greatest
Rising (1998)	4	0	3	1	0	8	Least
Ford et al. (2002)	9	1	4	3	0	17	Greatest
Ickovics et al. (2003)	10	2	6	3	1	22	Greatest
Grady & Bloom (2004)	7	2	3	1	0	13	Greatest
Ickovics et al. (2007)	12	3	4	6	1	26	Greatest

also fundamental to this model, as it emphasizes the importance of relationships and their effect on health behaviors and other health-related outcomes. Group members form relationships, learn from each other, and normalize the pregnancy experience. Finally, self-efficacy theory is an important foundation of this model as it emphasizes self-confidence and mastery of new skills. Mastery occurs as women are taught to perform self-care activities and try out new skills in the supportive environment of the group.

In CenteringPregnancy®, a group of 8–12 women attend 10 2-h visits together. Centering-Pregnancy®, provides: (1) group discussion with substantially more time for health promotion (15h vs. less than 3h); (2) peer group support; (3) a collaborative patient-provider relationship; and, (4) self-management training and activities such as taking and recording their own blood pressure (Rising, 1998). Centering Pregnancy® is also supported by a national organization Centering Pregnancy and Parenting Association (CPPA) that provides a national training program, consultation, and standardized materials for use by providers and group participants.

Components of Group Prenatal Care Studies

This section describes the group prenatal care studies that examined birth weight and preterm birth outcomes. Studies are described in detail in Table 15.1.

Study Design

The six studies that examined birth weight and/or preterm birth included three randomized controlled trials (RCTs) (Ford et al., 2002; Hoyer et al., 1994; Ickovics et al., 2007), one matched cohort study with a randomly selected control group (Ickovics et al., 2003), one cohort study with a retrospective and concurrent comparison group (Grady & Bloom, 2004), and one descriptive study that used qualitative and quantitative methods but had comparison data for only emergency room visits, cesarean births, and Apgar scores (Rising, 1998). Three out of the six studies were based on data from demonstration or pilot projects (Grady & Bloom, 2004; Hoyer et al. 1994; Rising, 1998).

Quality and Suitability

The small number of studies and lack of large RCTs suggest that the evidence to determine the effectiveness of group prenatal care is limited. Quality rating scores found in Table 15.2 indicate that three studies were rated poor, one was rated fair, and two were rated good. However, examination of the suitability of study designs found that five studies were classified as having the greatest suitability, and one was classified as having the least suitability (see Table 15.2). Among the five studies categorized as having greatest suitability, outcome data were available from

comparison groups. Three studies were RCTs and two other studies used population-based records of births at hospitals to identify comparison groups. In one of the latter studies, the comparison group was randomly selected from these data.

Study Type

All studies were published in peer-reviewed journals as research articles, with one exception. One study was published as a Letter to the Editor (Hoyer et al., 1994). The brevity of the letter format likely contributed to this report missing important information such as data collection method, validity and reliability of measures, and comprehensive reporting of outcome data.

Setting and Sample

Settings include hospital-affiliated clinics and public health clinics. Sample size ranged from 65 to 1, 047 participants with predominantly medically low-risk pregnancies. Samples include adults and adolescents from the Midwest, the Northeast, and the South. Within studies, ethnic disparities were examined only in the most recent study (Ickovics et al., 2007). However, the majority of samples are homogeneous, comprised predominantly of urban, low-income, and African-American women. In 2003, Ickovics et al. identified the need to examine the impact of CenteringPregnancy® among low-income and ethnic minority women and stated "Reducing the risk for low birth weight and preterm delivery is particularly important for black women and Latinas, because of persistent disparities in adverse perinatal outcomes," (p. 1052).

Data Collection and Measures

Methods used to obtain preterm and birth weight data in these six studies included medical records, labor/delivery logs, and hospital databases. The data source was not reported by Hoyer et al. (1994); data might have been collected from the adolescent rather than medical records, given that study participants completed a postpartum survey. In the studies reporting the percents of preterm births, the operational definition for preterm births was less than 37 weeks gestation. Hoyer et al. (1994) only reported gestational age at time of delivery and the mean gestational age for women in group care and individual care; data were not reported to calculate a preterm birth rate. Rising (1998) only reported the number of preterm births and infant birth weight. Therefore, when interpreting the Rising study results, we calculated percents.

Key Findings Related to Effectiveness of Group Prenatal Care

Findings from all group prenatal visit studies that examined birth weight and preterm births are presented together. Studies are not separated by whether they examined group care for adolescents as in earlier reports or whether they examined CenteringPregnancy® as in later reports.

Birth Weight

Five studies of group prenatal care examined birth weight as an outcome. Statistically significant differences in low birth weight were found in only one study. Grady and Bloom (2004) found the low birth weight rate was 2–2.5 times lower for group care compared to individual care. Although not statistically significant, trends in two studies found fewer low birth weight infants for group

care than individual care; 6.6% vs. 12.5%, $p = 0.18$ (Ford, et al., 2002) and 16 vs. 23 infants, respectively (Ickovics et al., 2003). The low birth weight rate among group participants was very low at 5.4% in the Rising (1998) study, but no comparison data were reported. In addition to an analysis of low birth weight infants, one study found that mean birth weight was significantly higher among group care than individual care (3,228g vs. 3,159g) participants (Ickovics et al., 2003). In this same study, Ickovics et al. also found that among preterm infants, birth weight was significantly higher for infants born to participants in group care than for infants born to mothers in individual care (2,398g vs. 1,990g). In the most recent study, a RCT, Ickovics, et al. (2007) found similar birth weights (3,161g vs. 3,112g, $p = 0.24$) and low birth weight rates (11.3% vs. 10.7%, $p = 0.90$) between group and individual care participants, respectively. While this study did not find differences in birth weight, the investigators did find that the higher the number of group visits, the longer the gestation ($r = 0.14$, $p = 0.003$), and the higher the birth weight ($r = 0.13$, $p = 0.003$).

Preterm Births

Five studies reported preterm birth data. Two studies found that compared to individual care, participants in group care had significantly lower preterm births (Grady & Bloom, 2004; Ickovics et al., 2007). In the study with the strongest design and greatest suitability for evaluating the effectiveness of group care, the randomized trial (Ickovics et al., 2007), 9.8% women in group care had preterm births compared to 13.8% of the women in individual care. Further, findings showed the effect of group care to be greater when the analysis only included African-Americans: 10% of African-Americans in group care had preterm births compared to 15.8% of African-Americans in individual care. A third study found similar preterm birth rates between group and individual care, 9.2 and 9.6%, respectively (Ickovics et al., 2003). The preterm birth rate was low at 4.5% in the Rising (1998) study, but no comparison data were reported. Another study reported gestational ages ranging from 34 to 42 weeks for the total sample, with the mean gestational age at 39 weeks for group care and 37 weeks for individual care (Hoyer et al., 1994). Statistical significance was not reported.

Study Strengths and Limitations

Major methodological strengths of the majority of studies in this review include clear descriptions of the conceptual frameworks, the interventions, screening criteria for study inclusion, study populations, and findings related to birth weight and prematurity.

Limitations of many of these studies in this review include lack of random assignment to study group, incomplete description and control of principal confounders such as provider type, failure to describe adverse events or negative outcomes that might be a consequence of group care, failure to discuss loss to follow-up, lack of reporting about the fidelity of the intervention, small samples, and failure to report a power analysis.

The fidelity of group prenatal care interventions is unknown. The only measure reported has been number of visits. While this measure captures exposure to the intervention, it does not capture the process of care or the implementation of the components of prenatal care. This limits the researcher's ability to explore the components of group prenatal care that potentially might make it more effective in reducing disparities in perinatal outcomes. The same is true of the control groups: no study reported monitoring the process or content of care provided in individual visits. Another limitation of group prenatal care studies is the impossibility of blinding study participants and staff to study group assignment. Regarding ethnic disparities, only one study examined outcomes among diverse ethnic groups (Ickovics et al., 2007).

Summary of Group Prenatal Care Evidence

Increases in prenatal care utilization have not reduced racial and ethnic disparities in prenatal care or pregnancy outcomes (Alexander, Kogan, & Nabukera, 2002; Martin et al., 2006). African-Americans continue to have higher rates of inadequate prenatal care, and their infants are still two to three times more likely to die or be low birth weight compared to other racial and ethnic groups (Kochanek, Murphy, Anderson, & Scott, 2004; MacDorman, Martin, Mathews, Hoyert, & Ventura, 2005; Martin, et al., 2006). However, prenatal care remains one of the most important public health interventions for women, providing a unique window of opportunity to improve women's pregnancy and health related behaviors. There is increasing recognition that the prenatal care delivery model in the U.S. is not meeting the needs of women, particularly women of color. A recent study of women who all initiated prenatal care in the first trimester found that African-Americans continued to have poorer perinatal outcomes (Healy et al., 2006).

Data from this review suggest that group prenatal care holds promise as an effective strategy for improving birth weight and preterm birth outcomes in vulnerable populations. However, few studies have been conducted and, while trends indicate favorable outcomes among women receiving group care, additional studies showing consistently statistically significant differences between participants in group and individual care are needed. Findings indicate that group prenatal care appears effective among predominately urban, low-income African-American women with medically low-risk pregnancies. The one study that specifically analyzed racial/ethnic group differences found that the group model of care had a greater impact on the preterm birth rate among African-American women than on the entire sample studied (Ickovics et al., 2007).

Two very similar models of group prenatal care were identified. The earlier model (Hoyer et al., 1994) focused on adolescents and, in addition to group visits, this model paired teens who conducted physical assessments on one another. The more recent model of care, CenteringPregnancy® (Grady & Bloom, 2004; Ickovics et al., 2003; Novick, 2004; Rising, 1998; Rising, Kennedy, & Klima, 2004), designed for both adolescent and adult pregnant women, continues to increase in popularity. CenteringPregnancy® has been implemented in the U.S., Canada, and Australia for women from diverse racial and ethnic backgrounds, income and education levels, and ages (adolescents and adults).

Given the increasing popularity and wide implementation of CenteringPregnancy®, even in the absence of a consistent body of strong evidence, it is urgent to conduct rigorous evaluations to determine the impact of this innovative model on maternal and infant health. The urgency to evaluate group care using RCTs with high quality process evaluations is compounded as the number of sites offering group prenatal care rises. With an increasing number of sites providing group care it becomes increasingly difficult to conduct RCTs since women might be interested in choosing group versus individual care and less willing to be randomized into care. This model of care is likely to gain in popularity over time since evidence suggests group care increases provider satisfaction (Klima, Vonderheid, & Norr, 2006; Rising, 1998) as well as patient satisfaction (Ickovics et al., 2007; C. Klima, personal communication, 2007). Additionally, group visits are likely to reduce health care costs. Findings from a recent finance study suggest a quick return on the initial investment of training staff. Group prenatal care allows clinics to serve up to "twice as many women in the same time and at a much lower cost in terms of labor resources" compared to individual care (Cox, Obichere, Knoll, & Baruwa, 2006, p. 1).

In conclusion, group care appears to address the shortcomings of individual care through enhanced education, skills building, and support of all pregnant women. While group care was not developed to specifically address ethnic/racial disparities, the facilitative group process enhances the ability of caregivers to respond to each groups' unique needs. To date, group prenatal care has been offered to women at socially higher risk of adverse outcomes including African-Americans and adolescents and has shown promise in these groups. Innovative models such as group prenatal care may be especially

beneficial for groups at greatest risk of adverse outcomes and may play a major role in reducing the long-standing racial/ethnic disparities in pregnancy and health related outcomes.

Doula Care Studies

Review Method

A search of Medline, CINAHL, PsycINFO, Popline, Web of Science, Cochrane Library, OCLC First Search and Academic Search Elite from June 2006 to November 2006 produced 83 references (including three meta-analyses) with key words of "doula" or "doulas," or combinations of "labor support" and "continuous." Literature on doula care was limited to published research written in English; however, no date limits were included. Studies that focused on persons other than doulas such as partners or nurse-midwives were also excluded. Conference abstracts and duplicated studies, non-original studies, and non-intervention studies were excluded. So that no studies were included twice, original studies used in the three meta-analyses were also excluded. These excluded original studies included in meta-analyses are listed in Table 15.3. Additionally, a manual search was performed on the references of the studies obtained based on the computer search. As a result of this search process, 15 studies were found. Studies were then examined to identify outcomes evaluated across studies. Mode of delivery (cesarean section) and length of labor were chosen as the major outcome measures in this review because these measures: (1) were commonly examined across studies; (2) indicate important labor processes; and, (3) represent outcomes that can affect both mothers' birth experiences and their recovery in the postpartum period. Nine of the 15 studies examined length of labor and/or cesarean section rates as the outcome measures and were therefore included in the review. These nine studies are summarized in Table 15.4.

Three out of the nine studies included in this review are meta-analyses (Hodnett, Gates, Hofmeyr, & Sakala, 2006; Scott, Berkowitz, & Klaus, 1999; Zhang, Bernasko, Leybovich, Fahs, & Hatch, 1996) (Table 15.5). The first original study was published in 1980 by Sosa, Kennell, Klaus, Robertson, and Urrutia. This study was included in two meta-analyses (Zhang et al., 1996; Scott et al., 1999) and therefore, is not examined separately in this review. The latest study was published in July/August 2006 (Campbell, Lake, Falk, & Backstrand, 2006). The results of the quality

Table 15.3 Doula studies cited in the three meta-analyses

Original study reports	Meta-analyses		
Authors	Zhang et al. (1996)	Scott et al. (1999)	Hodnett et al. (2006)
Sosa et al. (1980)	X	X	
Klaus et al. (1986)	X	X	X
Cogan and Spinnato (1988)			X
Hodnett and Osborn (1989)	X	X	X
Hemminki et al. (1990)		X	X
Hofmeyr et al. (1991)	X	X	X
Kennell et al. (1991)	X	X	X
Breart et al. (1992)		X	X
Kennell and McGrath (1993)		X	
Gagnon et al. (1997)			X
Langer et al. (1998)			X
Madi et al. (1999)			X
Dickinson et al. (2002)			X
Hodnett et al. (2002)			X

assessment for the six original studies not included in the three meta-analyses are summarized in Table 15.6. Initial inter-rater reliability for this quality assessment was 0.90 (range 0.82–1.00). Any differences were discussed until a consensus was reached.

Description of Doula Care

Doula care involves positive physical, emotional, and social support provided by an experienced woman who continuously supports another woman during labor and delivery (Klaus, Kennell, & Klaus, 1993). Content areas of this non-medical support may include massage, encouragement, praise, reassurance, information, and advocacy (Gilliland, 2002; Newton, 2004). Medical tasks such as taking blood pressures and pelvic exams are not considered part of the doula role (Klaus & Kennell, 1997; Meyer, Arnold, & Pascali-Bonaro, 2001). Doulas do not replace medical health care providers, partners, or other family members (Stein, Kennell, & Fulcher, 2003). Doulas usually have expertise based on their personal birth experiences and/or prior support of women in labor. In addition to personal experiences, doulas may be professionally trained (Rosen, 2004).

While doula care is often considered an innovative model in modern societies such as the U.S., it has existed for generations around the world. In fact, the term "doula" means "women's servant" in old Greek, referring to an experienced woman who helps other women (Klaus et al., 1993). In the late 1960s an American anthropologist, Dr. Dana Raphael, introduced doula care in a breastfeeding study (Raphael, 1973). In the late 1970s, Klaus and Kennel's breastfeeding study unintentionally found the positive effects of doula care on women (Gilliland, 2002). The first study to specifically examine the effect of doula support on women in labor found positive impacts on maternal-infant interactions and medical interventions such as lower rates of cesarean sections (Sosa et al., 1980). Since the 1980s, several studies have examined doula care in the U.S. as well as other countries. During this time period, while there has been an increase in awareness about the medicalization of childbirth, there has also been an increase in the rate of cesarean births (Gilliland, 2002) (see Declercq et al., Chapter 16). Obstetrical care increasingly includes induction or augmentation of labor, lack of ability for ambulation related to monitoring, lack of support for oral hydration and nutrition during labor, and epidural analgesia (DeClercq, Sakala, Corry, Applebaum, & Risher, 2006; Leavitt, 1986). Additionally, the medicalization of childbirth has been accompanied by an increasing nursing shortage that began in 1998 (Buerhaus et al., 2007) contributing to less direct care by nurses including support during childbirth (Gale, Fothergill-Bourbonnais, & Chamberlain, 2001). One study found that nurses spent only 27.8% of their total time in contact with women during childbirth and only 12.4% of their total time was spent providing supportive care including physical comfort measures, emotional support, instruction/information, and advocacy (Gale et al., 2001). It is possible that inadequate supportive care from nurses and other professionals has contributed to some women's decisions to seek assistance by doulas in addition to the attendance by family and friends during childbirth.

In 1992, the first doula organization, Doulas of North America International (DONA International), was founded to support doulas by developing training and certification programs for birth doulas and, more recently, postpartum doulas (DONA International, 2005). More than ten years later however, in 2005, only 3% of childbirths were accompanied by doulas in the U.S. (DeClercq et al., 2006). Additionally, since the service fees for doulas are usually not covered by medical insurance, the majority of doula-supported childbirths are among affluent women who can afford to hire a doula privately at an additional cost ranging from $300 to $1, 800 (Campbell et al., 2006). Therefore, historically, doula care has not been an option for low-income women, although their needs for social and emotional support may be more urgent than those of socially and economically advantaged women.

Table 15.4 Major outcomes associated with studies of doula care intervention: length of labor and mode of delivery

Author/study design/study type	Description of intervention	Sample and setting	Address ethnic/racial disparities (Y/N)	Data collection
Gagnon and Waghorn (1999). Secondary analysis of an RCT (published article)	One-to-one presence of a nurse during labor and birth Doulas had at least 30h of training	*N* =100 first mothers stimulated with oxytocin, 37 weeks or more gestation, singleton, in labor with vertex presentation, CD <5cm, no CS plan, no planned induction, no attendance 637-bed university hospital that allowed other support during labor Quebec, Canada	No	Medical records
Gordon et al. (1999). RCT prospective (published article)	Attendance from admission to labor by a trained doula Doulas attended "approved" training in community, were supervised for 2 births, and had a half day of hospital orientation	*N* = 314, *n* = 149 E, *n* = 165 C Primiparous moms with at least one family attendance, no complication, no planned CS, spontaneous labor, CD <5cm, 79% 18–35 years old, 20% >35 years old. About half white, with college degree. More than 86% had childbirth preparation, about 90% had planed breastfeeding. 100% had labor companion (other than doula for those in doula group) 3 HMO hospitals in NC, U.S.	No	Medical records
Trueba et al. (2000). RCT (published article)	Attendance during labor and delivery by doulas, including Lamaze Doula training conducted by a childbirth educator	*N* =100 *n* =50 E *n* =50 C Baby-friendly hospital First mothers, term, CD> 3cm, young (<25), no complications, no formal preparation for birth Mexico City, Mexico	No	Medical records
Shelp (2004). Descriptive, mixed intervention (published article)	Comfort, praise, and reassurance Non-medical support during labor and birth by Somali trained doulas as hospital employees Two day doula training program used a combination of the American and Hofmeyr models. All doulas were Somali women same as study participants	*N* =104 births Somali immigrant women, female circumcisions. Culture of participants resulted in men not being involved during labor but wanted other women with then during labor Limited English skills Minnepolis, MN, U.S.	No	Medical records

Measure	Key findings related to intervention effectiveness	Strengths	Caveats/Biases	Findings support the intervention? Yes/no for which populations?
CS, instrumental delivery, length of labor	Trend of lower *CS* among doula group: 56% reduction in risk of total CS rates RR 0.44 95% CI 0.19–1.01 *Instrumental delivery*: RR 1.39 95% CI 0.71–2.73 *Length of labor* Mean difference –0.7 95% CI 2.7–1.3	Controlled for major potential confounding factor by examining only oxytocin-stimulated women	Length of labor measured from study entry Fidelity of intervention monitored	No
CS, operative delivery, uncomplicated vaginal delivery	*CS* E: 16.8% C: 15.8% $p>0.05$ *Operative delivery* E: 19.2% C: 28.8% $p>0.05$ *Uncomplicated vaginal delivery* E: 67.8% C: 60.0% $p>0.05$	RCT Fidelity of intervention: MD and RN evaluate doula after each birth to identify any problems and then follow-up	Small sample size variances across the 3 sites Needed larger sample to detect differences in uncomplicated delivery Fidelity of intervention: doulas not in-house and initiation of doula care during labor unknown	No
Length of labor, CS	*Length of labor*: Shorter in the intervention group E: 14.51h C: 19.38h $p>0.05$ *CS*: Lower in the intervention group E: 2% C: 24% $p<0.003$	RCT CS: Striking effects in spite of the small sample size Combination of doula care and Lamaze	Small sample No continuity of care (depending on the circum-stances) Hospital environment limited doulas' ability to support	Yes, for Latina, young, healthy primparous, women
CS	*CS*: E: 14.4% C: 27.1% ($p=0.0025$)		Retrospective No randomization Fidelity of intervention not reported Likely that planned CS cases were included	Yes, for Somali immigrant women in U.S.

(continued)

Table 15.4 (continued)

Author/study design/study type	Description of intervention	Sample and setting	Address ethnic/racial disparities (Y/N)	Data collection
Campbell et al.(2006). RCT (published article)	Birth attendance by trained lay doulas who were self-identified support person and female friend or family member Two 2-h classes with certified doula	*N* = 586 *n* = 291 E *n* = 295 C Nulliparous, singleton pregnancy, low-risk at the enrollment in the study, able to identify a doula, no planned CS, no complications 1/2 white, 1/3 AA, 18% Hispanic Predominantly young, low-income Tertiary perinatal center allowing support persons in labor New Jersey, U.S.	No	Medical records
Dundek (2006). Quasi-experimental, no randomization to experimental and control group. retrospective (published article)	Attendance by trained doula Intervention emphasized emotional aspect-Hofmeyr model. All doulas and participants were Somali women.	*N* = 267 *n* = 123 E *n* = 225 C *Primiparous* *N* = 112 *n* = 68 C *n* = 44 E Somali refugees/ immigrants, most had female circumcisions. Culture of participants resulted in men not being involved during labor but wanted other women with then during labor Twin cities, Minnesota, U.S	No	Medical charts, birth log and birth certificates

RCT randomized controlled trial; *E* experimental group/doula support; *C* control group/usual care; *CS* cesarean continuous care by doula.

At the time of writing this chapter a randomized controlled trial completed about 15 years ago (McGrath & Kennell, published meta-analyses. This study was recently published as an article and focused on doula care, initiated shortly accompanied by a male partner during labor and delivery. Results show that women in the doula care group had Kennell (2008). A randomized controlled trial of continuous labor support for middle-class couples: Effect on

Measure	Key findings related to intervention effectiveness	Strengths	Caveats/Biases	Findings support the intervention? Yes/no for which populations?
Length of labor, CS	CS E: 18.9% C: 17.9% ($p = 0.7$) *Length of labor* E: 10.4 ± 4.3h C: 11.7 ± 4.8h ($p = 0.004$) *Length of 2nd stage* E: 58 ± 51min C: 64 ± 57min ($p = 0.2$)	*RCT* Statistical power Consistent doula training	Fidelity of intervention not reported No mention of provider type in labor Comparison group had 1–3 support persons during labor Authors suggest high technology setting over powers benefit of doulas	Yes, for low medical risk, low-income women
CS, vaginal birth (including operative delivery)	CS *All Somali mothers* E: 17.1% C: 26.6% *Primiparous Somali mothers* E: 18.2% C: 27.9% (statistical significance not reported)	Standardized intervention	No random assignment to study group, retrospective comparison data, statistical significance not reported Length of doula training based on Hofmeyr model was not reported. Doulas were not DONA certified at start of intervention, but they did eventually obtain certification Fidelity of intervention not reported	Yes, Somali refugees/ immigrants in U.S.

section rate; *CD* cervix dilation; *MD* medical doctor; *RN* registered nurse; *INTM* intermittent care by doula; *CONT*

2008) was available only as a conference abstract and therefore, did not meet the inclusion criteria of this review and after admission and provided continuously through delivery, to nulliparous middle- and upper-income women significantly lower CS rates (13.7% vs. 25.7%) compared to women in the control/usual care group. McGrath. & cesarean delivery rates. *Birth 35* (2), 92–97.

Table 15.5 Doula care meta-analyses

Source	Number of studies/*N*/ (% receiving doula care)	Findings	Contextual factors	Disparities/comments
Hodnett et al. (2006)	15 selected RCT studies *N*=12,791 % Receiving doula care not reported Length of labor: 9 trials *n*=10,322 Operative vaginal birth: 14 trials *n*=12,757 CS: 15 trials *n*=12,791 Spontaneous vaginal birth: 14 trials *n*=12,757	Calculated summary statistic *Length of labor* *No difference* Weighted mean difference −0.28, 95% CI −0.64–0.08 *Operative vaginal birth* *Favors doula care*: Doula group less likely to have operative vaginal birth RR=0.89, 95% CI 0.83–0.96 CS *Favors doula care*: Doula group less likely to have CS RR=0.90, 95% CI 0.82–0.99 Spontaneous vaginal birth: *Favors doula care*: Doula group more likely to have a spontaneous vaginal birth RR=1.08, 95% CI 1.04–1.13	All women in labor were included *Intervention*: Continuous presence and support. One-to-one intrapartum care by a health care professional (nurse, midwife) trained as a doula or childbirth educator, or a family member, friend or stranger with no special training in labor support *Hospitals in 11 countries* Australia, Belgium, Botswana, Canada, Finland, France, Greece, Guatemala, Mexico, South Africa, and the United States. Widely disparate hospital conditions, regulations, and routines 9 out of 15 trials allowed women to be accompanied by their husbands/partners or other family members during labor. Other six trials did not allow additional support people At least four trials involving at least 1,000 women	No studies present data by race/ethnicity or SES Included interventions by hospital staff: nurses and nurse-midwives Doula type, training, initiation and duration of doula care, and relationship to laboring women varied across trials Hospital policy varied across trials regarding the attendance of support persons during labor Loss to follow-up/attrition of subjects: sample size varies across outcomes
Scott et al. (1999)	11 selected RCT studies % Receiving doula care not reported CS: *n*=2,580 INTM and C	Meta-analysis *INTM vs. C* *Length of labor* *No difference* Weighted mean difference −0.21, 95% CI −0.46–0.04 CS *No difference*: OR 0.91, 95% CI 0.67–1.2 *Forceps* *No difference*: OR 0.72, 95% CI 0.50–1.0 *CONT vs. C* All outcomes favor continuous doula care	Included healthy, low and middle income, nulli/primiparous women at or near term Canada, US, Guatemala, Belgium, France, Greece, Finland, and South Africa Analysis controlled for the duration of doula support – findings indicate that studies that do not require the continuous presence of a doula during labor underestimate or lessen the positive effect size of doula care	No studies present data by race/ethnicity or SES Doula type, training, relationship to laboring women, initiation and duration of doula care varied across trials Hospital policy varied across trials regarding the attendance of support persons during labor Loss to follow-up/attrition of subjects: sample size varies across outcomes

	n=1,824 CONT and C Length of labor: *n*=2,153 INTM and C *n*=791 CONT and C Forceps delivery: *n*=2,391 INTM and C *n*=1,063 CONT and C	Length of labor *Favors doula care*: Shorter for continuous doula care vs. no doula care by a mean of 1h and 38min Weighted mean difference −1.64, 95% CI −2.3 to −0.96, $p<0.001$ CS *Favors doula care*: 51% decrease in odds among continuous doula care group vs. no doula care OR 0.49, 95% CI 0.37–0.65 *Forceps* *Favors doula care*: 57% decreased odds among continuous doula care group vs. no doula care OR 0.43, 95% CI 0.28–0.65		
Zhang et al. (1996)	4 selected RCTs *N*=1,349 *n*=504 (37%) E *n*=845 (63%) C Sample size varies across outcomes	Meta-analysis *Length of labor* *Favors doula care*: 2.8h shorter labor, 95% CI 2.2–3.4 *Forceps* *Favors doula care*: 54% decreased RR 0.46, 95% CI 0.3–0.7 CS *Favors doula care*: 46% decreased RR 0.54, 95% CI 0.4–0.7 *Spontaneous vaginal birth* *Favors doula care*: RR 2.01, 95% CI 1.5–2.7	Included primiparous women, no complications, CD<6cm, young (19–21 years old), low-income, no labor companion in the busy labor unit Urban setting 3 countries: Guatemala, US, South Africa	Presented data by SES. Primiparous, low-income, young urban women in hospitals with limited support personnel No studies present data by race/ethnicity Doula type, training, relationship to laboring women, initiation and duration of doula care varied across trials Hospital policy varied across trials regarding the attendance of support persons during labor Different inclusion criteria for labor

RCT randomized controlled trial; *E* experimental group/doula support; *C* control group/usual care; *CS* cesarean section rate; *CD* cervix dilation; *INTM* intermittent care by doula; *CONT* continuous care by doula; *OR* odds ratio

An updated Cochrane review (Hodnett et al., 2007) was published after studies for this chapter were reviewed. This updated Cochrane review added the RCT by Campbell et al. (2006), also included in this chapter, and reports a change in findings related to length of labor. Based on 10 RCTs (*n*=10, 922), continuous (doula) care is associated with a slightly shorter length of labor (weighted mean difference −0.43h, 95% CI −0.83 to −0.04). Findings related to operative delivery, CS and spontaneous vaginal birth continue to favor doula care. Continuous (doula) care is associated with fewer operative deliveries (15 trials, *n*=13, 357; RR 0.89, 95% CI 0.82–0.96), fewer CS (16 trials, *n*=13, 391; RR 0.91, 95% CI 0.83–0.99) and more spontaneous vaginal births (15 trials, *n*=13, 357; RR 1.07, 95% CI 1.04–1.12).

Table 15.6 Quality rating of studies associated with doula care intervention

Author	Reporting	External validity	Internal validity-bias	Internal validity-confounding	Power	Total quality score ≤14=poor, 15–19=fair, ≥20=good	Suitability of study to assess effectiveness
Gagnon et al. (1999)	7	3	3	3	0	16	Greatest
Gordon et al. (1999)	8	4	4	4	0	24	Greatest
Trueba et al. (2000)	6	3	3	3	0	15	Greatest
Shelp (2004)	4	4	2	2	0	12	Least
Campbell et al. (2006)	9	4	4	4	1	26	Greatest
Dundek (2006)	5	2	1	0	0	8	Least

Two intrapartum doula models commonly found in the literature are the "American" and the Hofmeyr models. Components or types of support in the American model include physical, guidance, emotional, educational, facilitation of communications, and advocacy (Shelp, 2004; Dundek, 2006). The Hofmeyr model emphasizes emotional support including comfort, praise, and reassurance (Shelp, 2004; Dundek, 2006). A third model, the community-based doula model, provides more extended doula care specifically for women of socially disadvantaged populations (Abramson, 2004; Abramson, Breedlove, & Isaacs, 2006). Initiated in 1996 (Chicago Health Connection, 2003, 2004) as a demonstration project for adolescents in Chicago and funded by private foundations, this model trains women from the community to support pregnant women from the prenatal period into the postpartum period. The community-based doula model has focused on care of disadvantaged women and their families, and has been adopted by diverse communities across the U.S. To date, no reports on community-based doula care delivered through a RCT have been published although such a trial has been conducted at the University of Chicago.

Components of Doula Care Studies

This section describes the doula care studies that examined the outcomes of mode of delivery and length of labor. All studies that met this review's inclusion criteria examined doula care and focused on the intrapartum period. These studies are summarized in Table 15.4.

Study Design

Study designs included four RCTs including one secondary analysis of an RCT, three meta-analyses, and two quasi-experimental designs. Original studies included in the meta-analyses were not included as separate studies in this review of the literature.

Quality, Suitability and Type

The number and type of studies suggest the evidence to determine the effectiveness of doula care is strong. However, quality ratings classified two studies as poor, two as fair, and two as good (Table 15.6). Four studies were classified as having the greatest suitability, while two studies was classified as having the least suitability. All studies were published in peer reviewed journals or the Cochrane Database of Systematic Reviews.

Setting and Sample

Six original studies examined doula care in hospitals in Canada (Gagnon & Waghorn, 1999), the United States of America (Campbell et al., 2006; Dundek, 2006; Gordon et al., 1999; Shelp, 2004), and Mexico (Trueba, Contreras, Velazco, Garcia, & Martinez, 2000). The three meta-analyses included studies conducted in numerous countries across several continents (Table 15.5). Zhang et al. (1996) included women in three countries, Scott et al. (1999) included women in eight countries, and Hodnett et al. (2006) included women in 11 countries. Samples across all of the studies in our review represent women primarily in North and Central American, African, and European countries.

Five studies limited participants to primiparous women with no medical complications. The majority of the studies (including one meta-analysis) focused on women with social disadvantages such as young age, low-income, or immigrant status (Campbell et al., 2006; Dundek, 2006; Gagnon & Waghorn, 1999; Zhang et al., 1996). Three studies (including one meta-analysis) reported that women did not have a companion during childbirth other than the doula, due to hospital policy (Zhang et al., 1996) or cultural values (Shelp, 2004; Dundek, 2006). In contrast, one study included women socially advantaged with respect to their income and education who had at least one family member present during childbirth (Gordon et al., 1999). These women were also older, well educated, and about half were white. Little information was reported about the sample characteristics of women in two meta-analyses (Scott, Berkowitz, & Klaus, 1999; Hodnett et al., 2006). One of these meta-analyses (Scott et al., 1999) reported characteristics other than being nulliparous or primiparous for only three out of ten trials. These characteristics varied: one trial included middle-income, white, married women; the second included low-income, Hispanic or African-American women; and the third included low- or middle-income women. The Cochrane meta-analysis (Hodnett et al., 2006) had the broadest inclusion criteria (all pregnant women in labor regardless of their characteristics such as parity, age, medical risk) and likely included a wide range of participant backgrounds. Additionally, this meta-analysis reported that 9 out of 15 trials had hospital policies allowing women to be accompanied by a support person during childbirth.

Data Collection and Measures

Data on mode of delivery and length of labor were obtained from hospital records in each randomized trial.

Modes of Delivery

The modes of delivery examined include spontaneous vaginal, operative vaginal delivery, and cesarean section. Natural vaginal deliveries indicate no major complications for the mothers and fetuses in the labor and delivery process. Operative delivery is indicated when maternal exhaustion, or fetal intolerance during the second stage occurs (Gabbe, Niebyl, & Simpson, 2007). A cesarean section is indicated when maternal progress of labor is inadequate despite adequate uterine forces, there is fetal intolerance in labor, or there is fetal malpresentation (Gabbe et al., 2007). In our review, all original studies excluded pregnancies complicated by malpresentation, multiple gestation, and maternal medical complications. One meta-analysis (Zhang et al., 1996) examined studies with women that had no antenatal complications. The inclusion/exclusion criteria related to medical and obstetrical risk were not clear in two meta-analyses (Hodnett et al., 2006; Scott et al., 1999). Therefore, in our review, cesarean section rates represent the rates of unplanned cesarean sections of women without complications at the onset of labor or at study enrollment.

As a point of reference, in the U.S. in 2005 the cesarean rate for all women was 30.3% (Martin et al., 2007). In 2003, the rate for all low-risk women was 23.5%, for all primiparous women 27.1%, and for low-risk primiparous women 23.6% (Menacker, 2005).

Length of Labor

Guidelines for normal length of labor based on research published in the 1950s are being re-evaluated as support increases for multifactorial labor curves (Cesario, 2004; Greenberg et al., 2006, 2007; Jones & Larson, 2003). Length of labor may be influenced by numerous factors such as women's parity, age, ethnicity, psychosocial well-being, epidural analgesia, induction of labor, and oxytocin augmentation. Specifically, primiparous women have longer labors than multiparous woman. Most women in the reviewed nine studies were primiparous and without complications. One study only included women stimulated with oxytocin (Gagnon & Waghorn, 1999).

Key Findings Related to Doula Care

Modes of Delivery

Cesarean section rate was the most frequently reported outcome among the studies included. Most studies showed that women who received doula care during labor and delivery had lower rates of cesarean sections and operative vaginal deliveries, and higher rates of uncomplicated spontaneous vaginal deliveries. Findings were strongest among socially disadvantaged women compared to socially advantaged women. The six original studies, all examined cesarean section rates. Two studies of Somali immigrant women (Dundek, 2006; Shelp, 2004) and another study conducted at a busy, crowded hospital in Mexico City (Trueba et al., 2000), found reductions (significant in the Shelp and Trueba studies) in the cesarean section rates in doula care groups. However, results of the study of doula care in Mexico City need to be interpreted with caution due to the small sample size: 13 out of 100 participants experienced a cesarean section (n = 1 Doula care, n = 12 control group). One study that included only nulliparous women stimulated with oxytocin showed a 56% reduction in the risk of having a cesarean section (Gagnon & Waghorn, 1999).

Two out of the six original studies did not show a reduction in cesarean sections among women receiving doula care (Campbell et al., 2006; Gordon et al., 1999). In one of these studies, participants were socially advantaged (older and well educated, about 50% white, had at least one labor companion, and approximately 87% attended childbirth preparation class) (Gordon et al., 1999). The lack of difference between study groups might be related to later initiation of doula care and relatively high rates of epidural analgesia (85% in the doula group and 88% in the control group) and oxytocin augmentation (46% in the doula group and 49% in the control group). Doulas were "called to the hospital when needed. By the time doulas arrived to work with patients…labor might have progressed beyond the point of benefit of having a doula" (Gordon et al., 1999, p. 425). In the other study (Campbell et al., 2006) that did not show a difference in cesarean section rates, participants were young, had low incomes, were about 50% white, and nearly half of the women in the control group had at least one labor companion. Cesarean section rates in both groups were lower than the overall hospital rate. Campbell et al. (2006) proposed that "participants might have been enlightened as to the potential benefits of a female companion after the informed consent process and chose to seek out and bring a female companion with them after they were randomized to the control group" (p. 461). Another reason for similar cesarean section rates is the type of doula and brief training session for the doulas. After enrollment in the study, the pregnant participants identified

a female family member or friend to become their doula and the doula then completed only 4h of training.

All three meta-analyses found lower cesarean rates in doula care groups compared to control groups (Table 15.5). At the time of writing this chapter, the Cochrane meta-analysis (Hodnett et al., 2006) was the most current and largest meta-analysis of doula care. It was updated in 2007 and results are provided in a footnote to Table 15.5. Hodnett et al. (2006) included 15 RCTs with data from 11 countries and found women in the doula care (continuous support) group to be 10% less likely to have a cesarean section (15 trials; $N = 12, 791$; RR = 0.90, 95% CI 0.82–0.99), 11% less likely to have an operational delivery (14 trials; $n = 12, 757$; RR = 0.89, 95% CI 0.83–0.96), and 8% more likely to have a spontaneous vaginal birth (14 trials; $n = 12, 757$; RR = 1.08, 95% CI 1.04–1.13). Subgroup analysis found that the effects of doula care (continuous support) on spontaneous vaginal birth and cesarean section rates were stronger: (a) in settings where women were not allowed a labor companion other than the doula; and, (b) when doula care was initiated before active labor. Similarly, the effects of doula care (continuous support) on spontaneous vaginal birth, operative vaginal birth, and caesarean birth were larger when doula care was not provided by a facility staff member. The other two meta-analyses found even larger significant effects: women in the doula group were 51 and 46% less likely to have cesarean sections, and 57 and 54% less likely to have forceps deliveries, respectively (Scott et al., 1999; Zhang et al., 1996). Women in one of these meta-analyses were socially at risk (e.g., young, urban residence, had low-income, and had no female labor companion) (Zhang et al., 1996). Consistent with the Cochrane meta-analysis, the meta-analysis by Zhang et al. (1996) showed women in the doula group to be twice as likely to have a spontaneous vaginal birth compared to women in the usual care group.

Length of Labor

Length of labor was examined in five studies: two original studies (Campbell, et al., 2006; Trueba et al., 2000) and all three meta-analyses (Hodnett et al., 2006; Scott et al., 1999; Zhang et al., 1996). While findings suggest women who received doula care had a shorter length of labor, the difference in length of labor only reached statistical significance in three out of the five studies. Campbell et al.'s original study had a large sample size ($N = 586$, $n = 291$ doula group, $n = 295$ control group) and found a significantly shorter length of labor in the doula care group (an average difference of 1.3h). Campbell and colleagues suggested the impact of doula care might be greatest in the first stage of labor. In the original study by Trueba et al., the length of labor was nearly 5h shorter (a 34% reduction) in the doula care group compared to the control group but, this difference was not statistically significant. Given this study's small sample size ($N = 100$, $n = 50$ in each study group), it is possible that there was insufficient statistical power to detect a difference.

Findings were mixed in the meta-analyses (Table 15.5). The Cochrane meta-analysis (Hodnett et al., 2006) did not find an association between doula care and shorter length of labor (9 trials, $N = 10,322$; weighted mean difference −0.28, 95% CI −0.64–0.08) (see note in Table 15.5 on updated Cochrane Review, Hodnett et al., 2007). In contrast, the other two meta-analyses found a significant reduction in the length of labor among the women who received doula care: the length of labor was shortened by 1 hour and 38min in the doula care group in Scott et al.'s (1999) meta-analysis ($N = 2,153$), and by 2h and 48min in Zhang et al.'s (1996) meta-analysis of primarily socially disadvantaged women ($N = 1,149$).

Inconsistent findings across the doula studies with respect to length of labor might be related to lack of a uniform operational definition of this outcome and variations in doula care related to type and training of doula, initiation and continuity (continuous vs. intermittent) of the doula at the bedside of the laboring woman, and whether or not the doula provider changed when a shift ended. Despite the study limitations, there was a consistent trend of shortened length of labor in the doula care groups.

Study Strengths and Limitations

Strengths of the majority of studies (both individual studies and the meta-analyses) include random assignment (seven out of the nine studies, except for Shelp, 2004 and Dundek, 2006), adequate sample size (seven out of the nine studies, except for Gordon et al., 1999 and Trueba et al., 2000), use of individual subjects as the unit of analysis versus randomization by clinic site, and consistent definitions for modes of delivery (all nine studies). All but two studies used random assignment with a control group to avoid selection bias. Most of the studies had large sample sizes. In particular, the meta-analysis by Hodnett et al. (2006) contained more than 1,000 women for each outcome variable.

Limitations of the studies reviewed include: the difficulty with blinding care providers; variations in doula type, training, timing of the initiation of doula care (before or during active labor), and duration (continuous vs. intermittent) of doula at the side of a woman in labor; lack of monitoring of the implementation of doula care and the experiences of the control groups; lack of a uniform measure for length of labor; and, lack of control for medical interventions in the analysis. Blinding care providers to study group assignment might influence care providers to deliver additional support to women in the control group, thus potentially reducing the differences in outcomes between study groups (Shadish, Cook, & Campbell, 2002). Variations in the delivery of doula care might have resulted in smaller effect sizes of the intervention. The types of doula providers varied greatly, including trained and untrained lay women, student nurse-midwives/childbirth educators, nurses (retired or active) and female relatives and friends. The Cochrane meta-analysis (Hodnett et al., 2006) found doula care provided by hospital staff members to have a smaller effect on outcomes compared to doula care by non-staff members. It is likely that hospital staff do not continually attend women during labor and delivery, given the other responsibilities that take them away from the laboring women.

Doula training also differed substantially across original studies in this review ranging from 4h (Campbell et al., 2006) to two days, (Shelp, 2004) to more than 30h (Gagnon & Waghorn, 1999). Some studies did not provide details about doula training (Dundek, 2006; Gordon et al., 1999; Trueba et al., 2000). In the studies included in the meta-analyses, there was substantial variation in the extent to which doulas were trained. In some studies, the doulas were well-trained and experienced; in others, they had no training (Hodnett et al., 2006; Scott et al., 1999; Zhang et al., 1996). In addition, two different doula models of care were reported on in the studies. For example, Dundek (2006) reported on the Hofmeyr model; Shelp (2004) reported on both the Hofmeyr and American models. The meta-analyses also included studies using both models.

Monitoring of the care provided for women in the doula and control groups was not adequate in the studies included in this analysis. The continuous presence of a doula has been recognized as one of the most essential components of doula support (Hodnett et al., 2006). Whereas most studies compared doula care and usual care, one of the three meta-analyses compared continuous and intermittent doula care (Scott et al., 1999). This meta-analysis found that the continuous presence of a doula during childbirth was associated with shorter lengths of labor, fewer forcep deliveries, and lower cesarean section rates. In contrast, intermittent doula support was not associated with any improvements in childbirth outcomes. While these findings highlight the importance of monitoring the implementation of the doula intervention, only two studies specifically reported monitoring doula care. One of these studies had staff physicians and nurses evaluate doulas after each birth, but the criteria were not described and the evaluation findings were not used in the analysis (Gordon et al., 1999). The other study monitored the consistency of care by reviewing the doulas' record of the care they provided such as stress management and physical comfort. The intervention protocol allowed doulas to have 20min breaks for meals and up to 10min breaks every 4h as well as to change doulas if labors lasted more than 10h from randomization. Despite this protocol, duration of care for each woman in labor was not reported (Gagnon &Waghorn, 1999). The lack of precise measurement of the intervention and subsequent identification of potential quality improvement activities, possibly decreased the effect sizes found in most of the studies (Shadish et al., 2002).

Outcomes of care also might be affected by the lack of a uniform measure of length of labor and medical interventions not measured or controlled for in the study design and/or analysis. Doula care that both begins before active labor and is documented/measured before active labor begins might allow for a potentially larger effect of doula care to be demonstrated. Additionally, length of labor tends to interact with and is influenced by other medical interventions. For example, oxytocin use and use of operative delivery can shorten length of labor. One study controlled for oxytocin use (Gagnon & Waghorn, 1999). Subgroup analysis in the Cochrane meta-analysis found that the positive effects of doula care on spontaneous vaginal birth and cesarean section rates were stronger in settings where epidural analgesia was not routinely available. Similarly, the positive effects of doula care on spontaneous vaginal birth were stronger in settings where electronic fetal monitoring was not routine. As a result, it is not certain whether the shorter length of labor was due to doula care or other factors.

Summary of Doula Care Evidence

The weight of the evidence suggests that doula care during labor and delivery reduces rates of cesarean section and operative vaginal deliveries, and shortens length of labor. Further, these positive effects of doula care were greater for women who were socially disadvantaged such as low-income, young, unmarried, primiparous, giving birth in a highly medicalized hospital setting without a companion, and women with language/cultural barriers. No adverse effects were reported.

The current main approach to doula care is limited to labor attendance during the intrapartum period. However, a more recent doula care model, the community-based doula model tested in the Chicago Doula Project (Abramson et al., 2006), and expanded into the prenatal and postpartum periods may have stronger positive effects on pregnant women and their infants. Likewise, in support of increased potency with a larger dose of doula care, one of the meta-analyses (Scott et al., 1999) reported on here found continuous rather than intermittent doula support to be associated with significant differences in intrapartum outcomes.

No study examined the effect of doula care on racial/ethnic disparities in maternal or infant outcomes. The majority of the studies included women with social disadvantages such as being young, low-income, or of immigrant status (Campbell et al., 2006; Dundek, 2006; Gagnon & Waghorn, 1999; Shelp, 2004; Zhang et al., 1996) and showed that doula care improved one or more birth outcomes. Based on the evidence, improvements in perinatal outcomes due to the presence of doulas are expected to be greater among disadvantaged women, although there is less evidence for the role of doula care in reducing racial/ethnic disparities.

The doula care model is used globally, and RCTs examining doula care have been conducted in at least 11 countries (Hodnett et al., 2006). In the U.S. in 2005, "four out of five of women (81%) who did not receive care from a doula had heard about this type of caregiver and care, including a majority (61%) who said that they had had a clear understanding of this type of caregiver and care" (DeClercq et al., 2006, p. 30). However, only 3% of births in 2005 in the U.S. were attended by doulas (DeClercq et al., 2006). A major barrier to wider use of doula care is the cost of the service in addition to the traditional costs of childbirth. Until the cost-effectiveness of doula care is demonstrated, hospital and community-based birthing centers, and insurance providers are not likely to pay for this service. Currently, as stated earlier, fees for doula care are typically not covered by insurance and only women who can afford this out-of-pocket expense have access to doula care. This expense precludes socially disadvantaged women who are under insured, uninsured, or enrolled in Medicaid from using doulas, and these are the women that stand to gain the greatest benefit. Policy changes supporting reimbursement for doula care would increase the availability of doula services for all women.

Studies are needed to examine the cost-effectiveness of doula care. Studies have shown that doula care can increase normal delivery rates and decrease medical interventions; both of these can contribute

to lower medical costs (Hodnett et al., 2006; Scott et al., 1999; Zhang et al., 1996). While there would be an initial outlay of funds to train doulas if the doula model became more widespread, the long term cost-savings would likely more than offset the initial investment. Cost-savings are expected to be highest among women at greatest risk of adverse pregnancy outcomes related to social factors.

Having professional health care providers act as or become doulas is controversial. On the one hand, health care providers working in hospitals (nurses and nurse-midwives) could be ideal doula care providers, because they have professional knowledge and skills regarding perinatal care and are familiar with the birth environments. However, these professional health care providers are commonly busy and need to prioritize other health care needs in addition to the provision of non-medical support of laboring women (Gagnon & Waghorn, 1999; Gale et al., 2001). In addition, the benefits (such as lower cesarean rates) of continuous doula care compared to intermittent care are thought to be related to the training and type of doula in these studies. Findings from Hodnett and colleagues (2006) suggest that lay personnel might be best suited to implement doula care. In fact, they suggest that doulas should not be members of the hospital staff allowing them to more easily to provide continuous support and remain a patient advocate not affected by hospital factors.

Because the role of doulas and the role of hospital staff such as nurse-midwives and nurses may overlap with respect to support, professional conflicts between professional health care staff and doulas have been a concern. Greater collaboration among professional health care staff and doulas should be addressed. Evidence that the effect of doula care on improved birth outcomes is greater when the doula is not a hospital staff member (Hodnett et al., 2006), and that lay doulas are more likely to provide continuous support during labor, emphasize the need for greater collaboration for the benefit of patients.

Studies are needed to monitor the fidelity of doula interventions and to identify the most effective doula training programs and care components. Findings could be used to maximize the effects of doula care by enhancing doula training and practice. Studies are also needed to examine the effects of doula care on racial/ethnic disparities in the perinatal outcomes that can be affected by intrapartum interventions, particularly if doula care remains an intervention primarily delivered in the intrapartum period.

Finally, intervention studies that examine extended doula support from early pregnancy until late postpartum are needed, especially with a strong design such as an RCT with valid and reliable measurements. As the doula care model expands its focus from the intrapartum period to include early pregnancy until late postpartum, a time frame used in the community-based doula model, evaluations are also expected to be undertaken. Outcomes such as rates of premature births, low-birth weight births, and infant and child outcomes such as child abuse rate and child development, might be clinically important indicators to determine the effects of this extended support.

Conclusion: Review of Findings Related to Group Prenatal Care and Doula Studies

Providers and women have expressed dissatisfaction with the current system of prenatal, intrapartum, and postpartum care. Few alternatives to individual prenatal care and routine hospital care are available. Group prenatal care and doula care (intrapartum and beyond) are two alternatives to routine care for childbearing women and their families. Both of these models show potential for improving pregnancy-related health outcomes among women at greatest social risk. However, additional studies are needed to further test, refine, and explore the components of both interventions and to understand their potential to have a positive effect on reducing racial/ethnic disparities in perinatal outcomes.

References

Abramson, R. (2004). The critical moment and the passage of time: Reflections on community-based doula support. *International Journal of Childbirth Education, 19*(4), 34–35.

Abramson, R., Breedlove, G., & Isaacs, B. (2006). *The community-based doula: Supporting families before, during, and after childbirth*. Washington, DC: Zero to Three.

Alexander, G. R., Kogan, M. D., & Nabukera, S. (2002). Racial differences in prenatal care use in the United States: Are disparities decreasing? *American Journal of Pubic Health, 92*(12), 1970–1975.

Breart, G., Mlika-Cabane, N., Kaminski, M., Alexander, S., Herruzo-Nalda, A., Mandruzzato, P., et al. (1992). Evaluation of different policies for the management of labour. *Early Human Development*, *29*, 309–312.

Buerhaus, P. I., Donelan, K., Ulrich, B. T., Norman, L., DesRoches, C., & Dittus R. (2007) Impact of the nurse shortage on hospital patient care: Comparative perspectives. *Health Affairs. 26*(3):853–862.

Campbell, D. A., Lake, M. F., Falk, M., & Backstrand, J. R. (2006). A randomized control trial of continuous support in labor by a lay doula. *Journal of Obstetric, Gynecologic, & Neonatal Nursing, 35*(4), 456–464.

Carroli, G., Rooney, C., & Villar, J. (2001). How effective is antenatal care in preventing maternal morbidity? An overview of the evidence. *Paediatric and Perinatal Epidemiology, 15*(Suppl. 1), 1–42.

Cesario, S. K. (2004). Reevaluation of Friedman's Labor Curve: A pilot study. *Journal of Obstetric, Gynecologic, & Neonatal Nursing, 33*(6), 713–722.

Chicago Health Connection. (2003). *History*. Retrieved December 16, 2006, from http://www.chicagohealthconnection.org/about/history.php.

Chicago Health Connection. (2004). *Annual report*. Retrieved December 16, 2006, from http://www.chicagohealthconnection.org/publications/chc_annual_report_2004.pdf.

Cogan, R., & Spinnato, J. A. (1988). Social support during premature labor: Effects on labor and the newborn. *Journal of Psychosomatic Obstetrics & Gynecology, 8*, 209–216.

Cox, R., Obichere, T., Knoll, F., & Baruwa, E. M. (2006). *A study to compare the productivity and cost of the Centering Pregnancy model of prenatal care with a traditional prenatal care model*. Final Report to March of Dimes.

DeClercq, E., Sakala, C., Corry, M. P., Applebaum, S., & Risher, P. (2006). *Listening to mothers: Report of the second national survey of women's childbearing experiences*. New York: Childbirth Connection. Maternity Center Association/Harris Interactive.

Dickinson, J. E., Paech, M. J., McDonald, S. J., & Evans, S. F. (2002). The impact of intrapartum analgesia on labour and delivery outcomes in nulliparous women. *Australian and New Zealand Journal of Obstetrics and Gynaecology, 42*(1), 59–66.

DONA International. (2005). *DONA international certification*. Retrieved March 2, 2006, from http://www.dona.org/develop/certification.php.

Dundek, L. H. (2006). Establishment of a Somali doula program at a large metropolitan hospital. *Journal of Perinatal & Neonatal Nursing, 20*, 128–137.

Ford, K., Beglicki, L., Kershaw, T., Schram, C., Hoyer, P. J., & Jacobson, M. L. (2002). Effects of a prenatal care intervention for adolescent mothers on birth weight, repeat pregnancy, and educational outcomes at one year postpartum. *Journal of Perinatal Education, 11*(1), 35–38.

Fullar, S. A., Lum, B., Sprik, M. G., & Cooper, E. M. (1988). A small group can go a long way. *MCN: The American Journal of Maternal Child Nursing, 13*(6), 414–418.

Gabbe, S., Niebyl, J., & Simpson, J. L. (2007). *Obstetrics*. Philadelphia, PA: Churchill Livingstone.

Gagnon, A. J., & Waghorn, K. (1999). One-to-one nurse labor support of nulliparous women stimulated with oxytocin. *JOGNN – Journal of Obstetric, Gynecologic, & Neonatal Nursing, 28*, 371–376.

Gagnon, A. J., Waghorn, K., & Covell, C. (1997). A randomized trial of one-to-one nurse support of women in labor. *Birth, 24*, 71–80.

Gale, J., Fothergill-Bourbonnais, F., & Chamberlain, M. (2001). Measuring nursing support during childbirth. *American Journal of Maternal Child Nursing, 26*(5), 264–271.

Gilliland, A. L. (2002). Beyond holding hands: The modern role of the professional doula. *JOGNN – Journal of Obstetric, Gynecologic, & Neonatal Nursing, 31*, 762–769.

Gordon, N. P., Walton, D., McAdam, E., Derman, J., Gallitero, G., & Garrett, L. (1999). Effects of providing hospital-based doulas in health maintenance organization hospitals. *Obstetrics & Gynecology, 93*(3), 422–426.

Grady, M. A., & Bloom, K. C. (2004). Pregnancy outcomes of adolescents enrolled in a Centering Pregnancy program. *Journal of Midwifery & Women's Health, 49*(5), 412–420.

Greenberg, M. B., Cheng, Y. W., Hopkins, L. M., Stotland, N. E., Bryant, A. S., & Caughey, A. B. (2006). Are there ethnic differences in the length of labor? *American Journal of Obstetrics & Gynecology, 195*(3), 743–748.

Greenberg, M. B., Cheng, Y. W., Sullivan, M., Norton, M. E., Hopkins, L. M., & Caughey, A. B. (2007). Does length of labor vary by maternal age? *American Journal of Obstetrics & Gynecology, 197*(4), 428.e1–7.

Healy, A. J., Malone, F. D., Sullivan, L. M., Porter, T. F., Luthy, D. A., Comstock, C. H., et al. (2006). Early access to prenatal care: Implications for racial disparity in perinatal mortality. *Obstetrics and Gynecology, 107*(3), 625–631.

Hemminki, E., Virta, A. L., Koponen, P., Malin, M., Kojo-Austin, H., & Tuimala, R. (1990). A trial on continuous human support on childbirth outcomes. *Journal of Psychosomatic Obstetrics & Gynecology, 11*, 239–250.

Hodnett, E. D., Gates, S., Hofmeyr, G. J., & Sakala, C. (2007). Continuous support for women during childbirth. *Cochrane Database of Systematic Reviews, 2003*(3), CD003766. DOI: 10.1002/14651858.CD003766.pub2. Retrieved November 11, 2007 from the Cochrane Library.

Hodnett, E. D., Lowe, N. K., Hannah, M. E., Willan, A. R., Stevens, B., Weston, J. A., et al., (2002). Effectiveness of nurses as providers of birth labor support in North American hospitals: A randomized controlled trial. *JAMA, 288*(11), 1373–1381.

Hodnett, E. D., & Osborn, R. W. (1989). Effects of continuous intrapartum professional support on childbirth outcomes. *Research in Nursing & Health, 12*, 289–297.

Hofmeyr, G. J., Nikodem, V. C., Wolman, W. L., Chalmers, B. E., & Kramer, T. (1991). Companionship to modify the clinical birth environment: Effects on progress and perceptions of labour, and breastfeeding. *British Journal of Obstetrics & Gynaecology, 98*, 756–764.

Hoyer, P. J., Jacobson, M., Ford, K., & Walsh, E. (1994). Pregnancy care for the adolescent. *The Nurse Practitioner, 19*(4), 27–28, 31–32.

Hughes, D. C., & Runyan, S. J. (1995). Prenatal care and public policy: Lessons for promoting women's health. *Journal of the American Medical Women's Association, 50*(5), 156–159, 163.

Ickovics, J. R., Kershaw, T. S., Westdahl, C., Magriples, U., Massey, Z., Reynolds, H., et al. (2007). Group prenatal care and perinatal outcomes – A randomized controlled trial. *Obstetrics and Gynecology, 110*(2), 330–339.

Ickovics, J. R., Kershaw, T. S., Westdahl, C., Rising, S.S., Klima, C., Reynolds, H., et al. (2003). Group prenatal care and preterm birth weight: Results from a matched cohort study at public clinics. *Obstetrics and Gynecology, 102*(5), 1051–1057.

Institute of Medicine. (2003). *Priority areas for national action: Transforming health care quality*. (K. Adams & J. M. Corrigan, Eds.). Washington, DC: National Academy Press.

Jones, M., & Larson, E. (2003). Length of normal labor in women of Hispanic origin. *Journal of Midwifery & Women's Health*. *48*(1), 2–9.

Kennell, J., Klaus, M., McGrath, S., Robertson, S., & Hinkley, C. (1991). Continuous emotional support during labor in a US hospital. A randomized controlled trial. *JAMA, 265*, 2197–2201.

Kennell, J. H., & McGrath, S. K. (1993). Labor support by a doula for middle income couples: The effect on cesarean rates. *Pediatric Research, 32*, 12A.

Klaus, M. H., & Kennell, J. H. (1997). The doula: An essential ingredient of childbirth rediscovered. *Acta Paediatrica, 86*, 1034–1036.

Klaus, M. H., Kennell, J. H., & Klaus, P. H. (1993). *Mothering the mother: How a doula can help you have a shorter, easier, and healthier birth*. New York, NY: Harper Collins Publisher.

Klaus, M. H., Kennell, J. H., Robertson, S. S., & Sosa, R. (1986). Effects of social support during parturition on maternal and infant morbidity. *British Medical Journal Clinical Research Ed., 293*, 585–587.

Klima, C., Vonderheid, S. C., & Norr, K. F. (2006). *Current Research on the Centering Model: Centering Pregnancy in Chicago*. Paper presentation at Group Health Care for the 21st century annual multidisciplinary conference, Chapel Hill, NC.

Kochanek, K. D., Murphy, S. L., Anderson, R. N., & Scott, C. (2004). Deaths: Final data for 2002. *National Vital Statistics Report, 53*(5), 1–115.

Langer, A., Campero, L., Garcia, C., & Reynoso, S. (1998). Effects of psychosocial support during labour and childbirth on breastfeeding, medical interventions, and mothers' wellbeing in a Mexican public hospital: A randomized clinical trial. *British Journal of Obstetrics & Gynaecology, 105*, 1056–1063.

Leavitt, J. W. (1986). *Brought to bed: childrearing in America 1750–1950*. New York: Oxford University Press.

Lu, M. C., & Halfon, N. (2003). Racial and ethnic disparities in birth outcomes: A life-course perspective. *Maternal and Child Health Journal, 7*(1), 13–30.

MacDorman, M. F., Martin, J. A., Mathews, T. J., Hoyert, D. L., & Ventura, S. J. (2005). Explaining the 2001–02 infant mortality increase: Data from the linked birth/infant death data set. *National Vital Statistics Report, 53*(12), 1–22.

Madi, B. C., Sandall, J., Bennett, R., & MacLeod, C. (1999). Effects of female relative support in labour: A randomized controlled trial. *Birth, 26*(1), 4–8.

Martin, J. A., Hamilton, B. E., Sutton, P. D., Ventura, S. J., Menacker, F., & Kirmeyer, S. (2006). Births: Final data for 2004. *National Vital Statistics Report, 55*(1), 1–101.

Martin, J. A., Hamilton, B. E., Sutton, P. D., Ventura, S. J., Menacker, F., Kirmeyer, S., et al. (2007). Births: Final Data for 2005. *National Vital Statistics Report, 56*(6), 1–104.

McGrath, S. K., & Kennell, J. H. (2008). A randomized controlled trial of continuous labor support for middle-class couples: Effect on cesarean delivery rates. *Birth, 35*(2), 92–97.

Menacker, F. (2005). Trends in cesarean rates for first births and repeat cesarean rates for low-risk women, United States 1990–2003. *National Vital Statistics Report, 54*(4), 1–8.

Meyer, B. A., Arnold, J. A., & Pascali-Bonaro, D. (2001). Social support by doulas during labor and the early postpartum period. *Hospital Physician, 37*(9), 57–65.

Misra, D. P., Grason, H., & Weisman, C. (2000). An intersection of women's and perinatal health: The role of chronic conditions. *Women's Health Issues, 10*(5), 256–267.

Newton, P. C. (2004). The doula's role during L&D. *RN, 67*(3), 34–38.

Novick, G. (2004). Centering pregnancy and the current state of prenatal care. *Journal of Midwifery & Women's Health, 49*, 405–411.

Public Health Service. (1989). *Expert panel on the content of prenatal care. caring for our future, the content of prenatal care: A report of the Public health service expert panel on the content of prenatal care*. Washington, DC: Public Health Service. Department of Health and Human Services.

Raphael, D. (1973). *Tender gift: Breastfeeding*. Englewood, NJ: Prentice-Hall.

Rising, S. S. (1998). CenteringPregnancy. An interdisciplinary model of empowerment. *Journal of Nurse-Midwifery, 43*(1), 46–54.

Rising, S., Kennedy, H., & Klima, C. (2004). Redesigning prenatal care through CenteringPregnancy. *Journal of Midwifery and Women's Health, 49*, 398–404.

Rosen, P. (2004). Supporting women in labor: Analysis of different types of caregivers. *Journal of Midwifery & Women's Health, 49*, 24–31.

Scott, K. D., Berkowitz, G., & Klaus, M. (1999). A comparison of intermittent and continuous support during labor: A meta-analysis. *American Journal of Obstetrics & Gynecology, 180*, 1054–1059.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.

Shelp, A. (2004). Women helping women: The Somali Doula Initiative. *International Journal of Childbirth Education, 19*, 4–7.

Sosa, R., Kennell, J., Klaus, M., Robertson, S., & Urrutia, J. (1980). The effect of a supportive companion on perinatal problems, length of labor, and mother-infant interaction. *New England Journal of Medicine, 303*, 597–600.

Stein, M. T., Kennell, J. H., & Fulcher, A. (2003). Benefits of a doula present at the birth of a child. *Journal of Developmental & Behavioral Pediatrics, 24*, 195–198.

Trueba, G., Contreras, C., Velazco, M. T., Garcia, L. E., & Martinez, H. B. (2000). Alternative strategy to decrease cesarean section: Support by doulas during labor. *Journal of Perinatal Education, 9*, 8–13.

Zhang, J., Bernasko, J. W., Leybovich, E., Fahs, M., & Hatch, M. C. (1996). Continuous labor support from labor attendant for primiparous women: A meta-analysis. *Obstetrics & Gynecology, 88*, 739–744.

Chapter 16
Contemporary Childbirth in the United States: Interventions and Disparities

Eugene Declercq, Mary Barger, and Judith Weiss

Introduction

Why focus on the process of childbirth in an examination of evidence for interventions to reduce disparities in reproductive and perinatal health? Is it because childbirth is the leading cause of hospitalization among women in the United States with the method of delivery having both short and long term implications for a woman's health? Or, because cesarean section has become the most common major surgery in the United States, where these rates vary widely by age, race, and ethnicity? Or, because rehospitalization of women for morbidities associated with childbirth accounts for more than 40,000 hospital visits annually (Kozak, DeFrances, & Hall, 2006)? Perhaps the best reason for studying the relationship between interventions in childbirth and reproductive and perinatal health is the degree to which these issues characterize the larger challenges of judiciously applying evidence to maximize the benefits and limit the risks associated with health care interventions. Simply put, the contemporary practice of maternity care in many significant areas only tangentially adheres to best practices, despite multiple efforts to encourage the widespread use of evidence based practice.

Maternal and child health research and practice have long emphasized precursors (contraception and prenatal care) and consequences (primarily the health of the baby) of birth, with little attention to the birthing process itself. The study of the birthing process has been largely left to clinicians examining alternative methods of intervention (Bloom et al., 2006) and social scientists who have studied the social and cultural context surrounding birth (Devries, Benoit, van Teijlingen, & Wrede, 2001). The public health community is primarily involved in the reporting and surveillance of vital statistics, particularly birth certificate, maternal, infant and fetal death data. The analysis of these data has notably focused on access (essentially utilization of prenatal care), maternal behaviors (smoking), and outcomes (prematurity, low birth weight, and infant mortality), with little consideration of the relationship between intra-partum interventions and/or method of delivery and maternal or infant outcomes. Not surprisingly much of the emphasis in the analysis of these measures has been on disparities across key subgroups, either in terms of age (historically teens and more recently older mothers) or race/ethnicity. The particular emphasis on disparities in adverse infant outcomes is manifested in the reporting of differences in infant mortality as a core MCHB Block Grant outcome measure (Maternal and Child Health Bureau, 2006).

This lack of interest in birth as a process is symbolized by the CDC's Pregnancy Risk Assessment and Monitoring System (PRAMS), a robust and extremely valuable population-based postpartum

E. Declercq (✉)
Department of Community Health Sciences, Boston University School of Public Health, Crosstown Center, 4th Floor, 801 Massachusetts Avenue, Boston, MA 02118, USA
e-mail: declercq@bu.edu

A. Handler et al. (eds.), *Reducing Racial/Ethnic Disparities in Reproductive and Perinatal Outcomes: The Evidence from Population-Based Interventions*, DOI 10.1007/978-1-4419-1499-6_16,

survey of mothers, carried out in 39 states (as of March 2008). The PRAMS questionnaire includes 35 core questions regarding the prenatal period, 16 questions on the postpartum period, and, several demographic background questions. The only question concerning birth is question 36, "When was your baby born?" There are also dozens of standard questions that states can select to supplement the core questions; of these, four concern birth and all involve method of delivery (vaginal or cesarean). This is not to suggest that questions on substance use during pregnancy (ten core questions on tobacco and alcohol alone), methods of fertility enhancements, and birth control are not valuable, but it does illustrate a lack of interest in the birth process itself in maternal and child health data systems. If the public health community does not systematically develop and report measures related to the birth process, there will be little basis for encouraging improvements and ensuring quality standards in how the process is implemented.

For this chapter, the importance of these concerns emerges in three crucial ways. First, there is little reliable systematic national data for us to analyze concerning the birth process in general, let alone evidence based analysis. Second, the work that is available is predominantly from, and aimed at, clinicians and not based on a public health approach involving community level prevention and programmatic initiatives. Third, the research that has been done on the birth process has rarely focused on disparities, the primary subject of this book. We have chosen to address these problems by focusing on five major interventions used in the birth process: induction, electronic fetal monitoring, epidurals, episiotomy, and cesarean section. We will briefly summarize the existing research on evidence based approaches to promote or prevent certain maternal and/or perinatal outcomes in these areas with an understanding that such research is rarely definitive and, even when it appears to be clear, it may not have a major influence on clinician behavior. We will then examine the distribution of these interventions across different groups, most notably by race/ethnicity. Since parity is such a crucial element in determining birth related behavior, in many cases we will also distinguish the results for primiparous and multiparous mothers.

The Changing Climate for Birth in the US

Rates of medical intervention, most notably cesarean section, have been shifting rapidly in recent years. The cesarean section rate in the US rose 50% between 1996 and 2006 (Hamilton, Martin, & Ventura, 2007; Martin et al., 2007) while epidurals, a rarely used anesthetic before the 1990s has become the dominant form of obstetrical anesthesia. Episiotomies in the meantime declined by 64% among mothers with a vaginal birth from 1980 to 2004 (Kozak et al., 2006), more than one in three mothers has her labor started by induction, and electronic fetal monitoring is now virtually universal. Aside from some attention toward trends in cesarean section rates by the public health community, most notably in the establishment of a much disputed (Sachs, Kobelin, Castro, & Frigoletto, 1999) Healthy People 2010 target of 15.5% primary cesarean rate for lower risk mothers, in general, these changes have occurred with little comment in the scientific literature and with even less note taken of the changes in these rates by subgroups. In fact, the general uptake of medical interventions during the birth process, has not been based on solid evidence of the efficacy of such interventions in improving maternal and infant outcomes. This chapter is an attempt to rectify that neglect.

Data Sources and Methods

The data for this chapter are drawn from several different sources, reflecting the fragmented nature of data systems related to the birth process. Overall national trends and disparities by race/ethnicity

are based on birth certificate data available from the Natality Branch of the National Center for Health Statistics (NCHS) either in their published reports (Martin et al., 2006) or through analysis of the CD-ROM public access file [National Center for Health Statistics (NCHS), 2006] for data on cesareans, electronic fetal monitoring and induction. National hospital discharge data, as noted in the introduction, are the source of our information on episiotomy use (Kozak et al., 2006). Massachusetts natality data are also used because unlike NCHS data they include method of payment for the delivery and thus allow stratification of cesarean data by whether the birth was paid for by private or public insurance (Massachusetts Department of Public Health, 2008).

Another data source utilized is *Listening to Mothers II* (LTMII), a national retrospective survey of 1,573 mothers (200 via telephone interviews, 1,373 from an online national database maintained by Harris Interactive) that explores womens' experiences during childbirth in 2005 (Declercq, Sakala, Corry, & Applebaum, 2006). The LTMII questionnaire which took 30 min to complete, combined closed and open-ended questions that provided data on all five interventions discussed in this chapter, including data on mothers' attitudes regarding these procedures. The results were weighted and the profile of respondents is comparable to the national birthing population for 2005. The LTMII survey is the primary source of data on electronic fetal monitoring (EFM), epidurals and episiotomies. Mothers responding to LTMII reported slightly higher levels of EFM and much higher levels of induction than that reported in national birth certificate data. However, there is concern with both underreporting and weak conceptualization of these items in the national birth certificate data (Lydon-Rochelle, Holt, Easterling, & Martin, 2000). The EFM measure on the birth certificate does not distinguish between a brief episode of monitoring and continuous monitoring, nor does it distinguish between external and internal fetal monitoring. The induction measure provides no information on the method or intensity of the induction. Nonetheless, there are pronounced differences in the rates of reported induction across racial/ethnic groups in the birth certificate data, and there is no reason to expect differential reporting related to these groups.

When presenting results from LTMII broken down by race/ethnicity we present the results of statistical tests. However, in keeping with the practice of the National Centers for Health Statistics, when presenting U.S. natality data by race/ethnicity, no statistical tests are presented because the very large numbers (more than four million births) result in virtually all comparisons being statistically significant, regardless of clinical significance.

Our approach to identifying the evidence relative to outcomes for selected interventions was first to focus on meta-analyses related to those interventions. We also report on issues that may account for differences or similarities in interpretation of the meta-analyses, especially in relation to typical practice in American obstetrics. If there were no published meta-analyses or if a published meta-analysis primarily reported a single, large randomized clinical trial, then we report on specific studies in detail. Details on the meta-analyses conducted for three of the interventions, electronic fetal monitoring, epidural anesthesia, and episiotomies, are presented in Tables 16.5, 16.6, and 16.9. A review of individual studies (outcomes and quality) related to induction at 41 weeks gestation and planned c-section for breech birth are presented in Tables 16.3, 16.12–16.14.

Evidence Base for Birth Related Obstetrical Practice and Interventions

Evidence-Based Practice

The current movement to practice evidenced-based care had its beginnings with Archie Cochrane, M.D., who felt that treatments for patients should be evaluated using randomized controlled trials. In 1979 he awarded the "wooden spoon" to the field of obstetrics, the area of medicine he judged

both to have the most poorly-researched practices and to be the least likely to change its procedures when new high-quality evidence clearly demonstrated specific practices to be harmful or others to be beneficial. Obstetrical research was criticized for lack of comparison populations, few well-conducted randomized trials, and studies with too few subjects to answer the scientific questions being asked.

Cochrane's work inspired others, most notably Iain Chalmers, who founded the National Perinatal Epidemiology Unit (NPEU) at Oxford in 1978. The NPEU did much of the work leading to the 1989 publication of *Effective Care in Pregnancy and Childbirth* (Chalmers, Enkin, & Keirse, 1989), a summary of systematic reviews of effective childbirth practices. Since that time, a number of other perinatal research centers and groups have been conducting randomized clinical trials with populations sufficiently large to test hypotheses adequately, developing prospective cohort studies, and/or performing meta-analyses of existing research. In the United States, the National Institute of Child Health and Human Development's Maternal Fetal Medicine Unit (MFMU) Network has conducted large multi-site studies of interventions intended to prevent pre-eclampsia and preterm birth and examine vaginal birth after cesarean (National Institute of Child Health and Human Development, MFMU Network, n.d.).

The principal approaches used by the NPEU have been applied to all areas of medicine with the 1993 founding of the Cochrane Database of Systematic Reviews, which publishes reviews of the effects of health care interventions and treatments. Findings from quality studies with sufficiently consistent study designs are summarized statistically using meta-analysis techniques. Although other types of evidence may be considered, studies included in these reviews are almost exclusively randomized controlled trials. In particular, this may limit the number of perinatal studies reviewed, because in some cases it may not be feasible or ethical to randomize care involving both a vulnerable developing fetus and/or a pregnant woman.

Almost 20 years ago, the NPEU published a list of obstetric and related practices that should definitely be abandoned, given the evidence already available (Chalmers et al., 1989). The list included several practices that we will examine specifically – routine use of episiotomy, routine continuous fetal heart rate monitoring without availability of fetal scalp blood sampling, routinely inducing labor before 42 weeks gestation or with pre-labor rupture of membranes at term, and routine repeat cesarean section after a previous cesarean. Some of these practices which were routine 20 years ago even though they had been refuted by the best evidence at the time, continue to be very much a part of routine obstetrical practice today. The evidence for each of these practices will be examined in turn.

Induction

Labor induction, the stimulation of contractions prior to the onset of labor, is a medical intervention used for a variety of reasons. Generally, labor induction can be classified as medically indicated or elective. Concern for the health of either the mother or the fetus may prompt a medically indicated induction. For the mother, medical indications include worsening pre-eclampsia or deteriorating maternal condition, chorioamnionitis, diabetes at term, the presence of a fetal demise, or timing issues such as a previous fast labor. For the fetus, induction is medically indicated for reasons including the prevention of stillbirth or neonatal complications such as isoimmunization, fetal growth restriction, low amniotic fluid, post-dates pregnancy, multiple gestation, and acute fetal distress (American College of Obstetricians and Gynecologists, 2006b). In addition, women with premature rupture of membranes at term may be induced, although a strategy of waiting for spontaneous labor onset is an acceptable alternative (Hannah et al., 1996). Another reason cited for induction is suspected fetal macrosomia, although studies have not shown any benefit except for type 1 diabetic women (Irion & Boulvain, 2000).

Elective reasons for induction are numerous, including convenience of provider and/or mother, desire for a specific provider to be present at the birth, transportation and weather issues, and logistical concerns related to the mother, provider, or hospital (Leonhardt, 2006). There are some contraindications to labor induction, including the presence of placenta previa, an active genital herpes lesion, a uterus with a previous vertical incision, or an abnormal fetal lie (such as transverse).

Labor induction can be accomplished using pharmacological and non-pharmacological methods. Oxytocin is the most commonly used drug in obstetrics and is used to induce women at or near term gestation. Mifepristone (RU-486) may also be used to induce labor and is more commonly used in women more remote from term. Non-pharmacological methods of induction include breaking the amniotic sac (amniotomy) and membrane stripping.

What are the risks associated with induction of labor? One is iatrogenic prematurity. Another is uterine hyper-stimulation, which can lead to fetal heart rate decelerations. The decelerations in turn prompt more interventions, such as increased operative delivery. Several studies have found that nulliparous women undergoing elective induction have increased rates of: operative birth (including cesareans), epidural analgesia, and neonatal problems (Maslow & Sweeny, 2000; Seyb, Berka, Socol, & Dooley, 1999). However, two recent meta-analyses did not show an increase in cesareans or poorer perinatal outcomes among women induced at 41 weeks or later (Crowley, 2000; Sanchez-Ramos, Olivier, Delke, & Kaunitz, 2003).

Rates of labor induction in the United States have more than doubled over the last 15 years, regardless of the dataset used (Kirby, 2004; Martin et al., 2007; Zhang, Yancey, & Henderson, 2002) with the most recent estimates based on 2005 birth certificate data at 22% overall (see Table 16.1) and based on 2005 hospital discharge data at 17% (CDC, 2008b). There are considerable and different measurement problems associated with each dataset, but there is little doubt that the rate has been climbing quickly. Although induction rates have increased, there is large geographic variability in rates. In 2005, induction rates varied from 11% in California, to 35% in Utah, and 34% in West Virginia. Within-states, the variation is also substantial: in 2003, the rate among New York counties ranged from 8 to 53% (NCHS, 2006). Likewise, institution-specific studies have shown large differences between academic hospitals and community hospitals. In one study using abstracted medical records, induction rates ranged from 18% at a university hospital to 34% at a community hospital. The vast majority of inductions in the university setting met ACOG criteria for medically indicated (95%), whereas 44 and 57% of the inductions at the two community hospitals were elective (Beebe et al., 2000).

There are several potential reasons for the increase in inductions. Cervical ripening agents began to be more widely used in the early- to mid-1990s. These agents will soften and efface, and occasionally dilate, a cervix in a woman who has an 'unripe' cervix. Prior to the introduction of these agents, providers were more reluctant to use oxytocin alone to induce a woman with an unripe cervix, since the woman was less likely to go into active labor, increasing the risk of a cesarean birth. Also, both improved screening technology and improved care of premature infants may have played a role in increasing the induction rate. Newer imaging technology and other diagnostic tools enable providers to identify more at-risk pregnancies and improved care of premature infants may have led providers to now consider the risks of an earlier birth to be less than the risks of continuing some pregnancies (Rayburn & Zhang, 2002).

Table 16.1 Proportion of mothers receiving induction U.S., natality data, 2005, by Race/Ethnicity

2003 Natality data	White non-Hispanic %	Black non-Hispanic %	Hispanic %	All %
Induction – all	27	20	16	22
Induction – primiparas	30	24	19	26
Induction – multiparas	25	17	13	20

Source: 2005 Natality data. Centers for Disease Control and Prevention (2008b)

Lastly, newer, more methodologically sound studies examining the risks of stillbirths, show a sharp increase in the stillbirth rate after 39 weeks gestation (Rand, Robinson, Economy, & Norwitz, 2000). The human tragedy of a stillbirth, whatever its cause, may create a potential malpractice liability for providers; while the likelihood of a stillbirth increasing after 39 weeks is rare, providers may become reluctant to undertake an expectant management approach for pregnancies at 40 or 41 weeks. On the other hand, the financial costs associated with early induction could be quite substantial. A decision-tree analysis of the economic costs and health outcomes of elective induction demonstrated that elective induction of 100,000 nulliparas at 39 weeks gestation would result in increased health care costs of close to $100 million, along with an excess of 12,000 cesareans, and 133 fewer fetal deaths (Kaufman, Bailit, & Grobman, 2002).

The rapidly increasing rates of induction, especially of elective inductions at less than 41 weeks gestation, partially account for a shift of the gestational age distribution of births, with 39 weeks now being the most common gestational age (Davidoff et al., 2006). U.S. natality data show that 60% of inductions in 2005 occurred at 39 weeks or earlier (Centers for Disease Control and Prevention, 2008b; Davidoff et al.). With recent attention to the increased morbidity of 'near-term' infants born at 34–36 weeks gestation (Shapiro-Mendoza et al., 2006), this trend toward earlier inductions may have important public health implications.

National birth certificate data indicate that induction rates vary by race/ethnicity. Table 16.1 presents induction rates based on 2005 U.S. natality data. Regardless of parity, rates were highest among white non-Hispanic women, lowest among Hispanic women, with black non-Hispanic women falling in between. First time mothers in each group have higher rates than multiparous mothers.

National birth certificate data only report rates of induction resulting in labor (22% overall in 2005). The *Listening to Mothers II* survey included a series of questions concerning induction, allowing a richer perspective on the phenomenon than is available in vital statistics or administrative datasets. Surveyed mothers were asked if they self-induced, that is began labor themselves, or had a medical induction attempted, and whether these efforts began labor. LTMII mothers reported higher rates of labor induction (34%) than are generally found in other data sources. The higher rates may be due in part to mothers reporting a broader array of techniques used for induction (e.g., stripping the membranes) than are reported on birth certificates.

Table 16.2 presents results from LTMII with respect to induction by parity and race/ethnicity. Overall, over one in five (22%) women tried to self induce. White non-Hispanic mothers (25%) were slightly more likely to try to self induce compared to the other two groups, a finding that did not vary by parity. White non-Hispanic mothers were also significantly more likely to have a medical provider use drugs or techniques in an attempt to induce them, with 53% of white non-Hispanic first time mothers reporting an attempted induction, compared to 36% of Hispanic mothers. Not surprisingly, white non-Hispanic mothers were most likely to have their labor begun by their provider (38%) particularly when compared to black non-Hispanics mothers (27%). Contrary to national birth certificate data, Hispanic mothers who had given birth before reported the highest rates of attempted and successful induction. The major techniques used by providers to induce labor (breaking membranes, pitocin, sweeping membranes) did not vary greatly by mother's race/ethnicity (data not shown). The largest recent change in obstetric practice has been in "routinely" inducing women at 41 weeks gestation instead of providing non-invasive testing of fetal well-being, generally the non-stress test, and if results are normal, waiting until women meet the definition for a post-dates pregnancy, greater than 42 weeks gestation (Norwitz, Snegovskikh, & Caughey, 2007).

Table 16.3 presents a summary of two major randomized controlled trials (Hannah et al., 1992; Heimstad et al., 2007) examining the key outcomes associated with induction, reflecting both the wide array of potential consequences and the limitations of existing studies. Neither study showed any benefit of induction at 41+ weeks gestation vs. 42+ weeks gestation in terms of improved neonatal outcomes. The Hannah study (1992) found improved maternal outcomes with earlier induction

Table 16.2 Results from Listening to Mothers II related to induction by Race/Ethnicity

Induction	White non-Hispanic %	Black non-Hispanic %	Hispanic %	All %	p Value
All mothers (n)	(n = 980)	(n = 192)	(n = 329)	(n = 1,561)	
Tried to self-induce	25	17	18	22	(0.006)
Attempted medical induction	44	34	40	41	(0.043)
Induced medically	38	27	34	34	(0.033)
Primiparas (n)	(n = 324)	(n = 62)	(n = 112)	(n = 517)	
Tried to self-induce	25	18	11	21	(0.013)
Attempted medical induction	53	46	36	48	(0.018)
Induced medically	48	38	26	41	(0.003)
Multiparas (n)	(n = 656)	(n = 130)	(n = 217)	(n = 1,044)	
Tried to self-induce	24	16	22	23	(0.038)
Attempted medical induction	39	29	43	38	(0.058)
Induced medically	33	22	38	32	(0.041)

Source: Adapted from Declercq, Sakala, et al. (2006)

(41 vs. 42 weeks) with fewer cesareans performed for fetal distress in labor, but the Heimstad et al. (2007) study did not find any such benefit. Only one of the studies addressed disparities (Hannah et al., 1992). On the whole, the research on induction documents a rapid rise in the use of the intervention, variation in the rates across different racial and ethnic groups, and mixed evidence on the beneficial impact of the intervention, at least in terms of perinatal mortality.

Electronic Fetal Heart Rate Monitoring

The use of electronic fetal heart rate monitoring (EFM) in labor is the most frequent obstetric intervention in the United States, reportedly used in 85% of all births in 2005 (CDC, 2008b). Virtually all the mothers in the LTMII survey (97%) reported being attached to a monitor for at least a portion of their labor and more than three out of four were continuously monitored. The rates did not vary by parity or to any great degree by race/ethnicity (Table 16.4), which is not surprising given such widespread use.

EFM was introduced into obstetrical practice in the late 1960s, with the hope of using it as a screening method to identify labors in which the fetus was at risk for severe asphyxia, thus allowing for interventions before neurological damage occurred. Thus, after EFM was introduced, many believed that the incidence of cerebral palsy would decline dramatically. However, subsequent research has demonstrated that only 10–20% of diagnosed cerebral palsy is due to intrapartum events, while the vast majority of cerebral palsy cases originate during the antenatal period (Greene, 2006). The rate of cerebral palsy today is similar to the rate in the 1960s despite the nearly ubiquitous use of EFM (Winter, Autry, Boyle, & Yeargin-Allsopp, 2002).

EFM was incorporated into the practice of obstetrics without any prior randomized clinical trials (RCT) to demonstrate either safety or efficacy. Since that time many RCTs of EFM have been conducted; Table 16.5 presents the results of three meta-analyses of these RCTs. When the first RCT was done in the mid-1970s, obstetric providers were reluctant to believe the results, which showed that there were no differences in fetal and infant outcomes between the EFM group and the control group; however, the EFM group experienced a 2.5 times higher rate of cesarean section (Thacker,

Table 16.3 Major outcomes associated with studies of induction at 41 weeks gestation

Health status outcome #1 perinatal mortality								
Author (year)	Study design	Study type	Description of intervention: what, how and where	Populations studied (ages included, race and ethnicity) and sample size	Address disparities (yes/no)	Key findings related to intervention effectiveness (OR with CI)	Caveats/biases	Findings support the intervention? Yes/noFor which populations?
Hannah et al. (1992)	RCT	Published paper	Enrolled at 41 weeks; induction within 4 days of randomization vs. non-stress testing 3×/week	Low-risk women at 41 weeks in Canada; Total N = 3,407; Induction N = 1,701; Monitoring N = 1,706	Data on race and SES collected; black women had higher risk of CS	Monitored group had more CS (24.2% vs. 21.2%) OR 1.22 (1.02, 1.45) No difference in perinatal deaths or neonatal morbidity	66% of induction group were actually induced	Yes – less cesareans No for neonatal morbidity or perinatal deaths
Heimstad et al. (2007)	RCT	Published paper	Induction at 41 weeks, 2 days by ultrasound dating vs. monitoring with non-stress testing at every 3 days intervals with induction at 42 weeks 6 days if undelivered	Singleton pregnancies in Norway; N = 508 Induction N = 254; Monitoring N = 254; maternal ages not given; 98% of both groups were Caucasian	No	No difference found in neonatal morbidity scores – study powered to assess this measure No differences in operative deliveries; induced women had shorter second stages and shorter labors	No information on oxytocin regimen to assess generalizability to US practice standards	No

Table 16.4 Rates of electronic fetal monitoring by Race/Ethnicity, 2005

Electronic fetal monitoring	White non-Hispanic % (n = 748)	Black non-Hispanic % (n = 102)	Hispanic % (n = 156)	All % (n = 1,049)	
All mothers					
Used EFM at all	96	100	98	97	(p = 0.135)
EFM continuous all labor	75	85	78	76	(p = 0.261)

Source: Adapted from Declercq, Sakala, et al. (2006)

Table 16.5 Meta-analysis table: electronic fetal monitoring

Source	Number of studies (N)	Findings	Contextual factors	Disparities/ comments
Alfirevic et al. (2006)	12 studies, N = 37,615 women	*Increased risk of*: Caesarean section RR 1.66, 95%CI 1.30, 2.13, heterogeneity among studies; overall analgesia RR 1.09, 95%CI 1.01, 1.18 *No difference in*: epidural analgesia, perinatal mortality or cerebral palsy	4 studies identified were excluded, authors do not say why 3 studies have low-risk patients, 5 high risk, 4 mixed risk	No information on race/ethnicity or SES
Vintzileos et al. (1995)	9 RCT studies, N = 18,561, 50.6% received EFM	*Increased risk of*: caesarean for suspected fetal distress, forceps/ vacuum for FD, surgical intervention *No difference in*: caesarean rate for no FD, overall perinatal mortality rate *Decreased risk of*: perinatal mortality due to fetal hypoxia	2 trials excluded Patients included were 26 weeks gestation with live fetus Included both low and high risk patients and preterm gestation Analysis relied heavily on study excluded by Thacker	No information on race/ethnicity or SES Likelihood of detecting true difference in perinatal mortality small due to small sample sizes and settings
Thacker (1987)	6 RCT studies, N = 3,928, ~50% received EFM 1 RCT, N = 12,964, not included in pooled analyses	*Increased risk of*: Cesarean delivery, total operative deliveries *No difference in*: operative vaginal deliveries	Reports both pooled and individual analyses	No information on race/ethnicity or SES

1987). In the past 30 years, there have been at least 12 adequately designed trials of EFM (as judged by Cochrane criteria) which included over 37,000 women. Results of the RCTs and the subsequent meta-analyses have been remarkably consistent (Alfirevic, Devane, & Gyte, 2006; Thacker; Vintzileos et al., 1995). EFM has not been shown to decrease overall perinatal mortality rates or the rate of admissions to intensive care units. In addition, inter- and intra-observer reliability of interpretation of fetal heart rate tracings from EFM has been poor (Parer & King, 2000), definitions of fetal heart rate patterns have not been standardized, and provider knowledge about the pathophysiology needed to interpret FHR patterns may not be current (Parer & King). There is evidence that EFM infants have half the incidence of neonatal seizures (RR 0.50; 95% CI 0.31, 0.80), but in long-term follow-up there was no difference in rates of cerebral palsy or other neurological sequelae (Alfirevic et al.; Graham, Peterson, Christo, & Fox, 2006; Vintzileos et al.). All meta-analyses found

an increase in operative births associated with EFM, including an increased risk of 50–66% for cesareans and 16–23% for assisted vaginal birth (Alfirevic et al.; Vintzileos et al.).

Although EFM has not been shown to have external validity as a screening method to prevent long-term neurological sequelae, has poor reliability and appears to increase the risk of operative birth, its use continues to be ubiquitous, in part for economic reasons. The practical alternative to EFM is intermittent auscultation, which requires either one-on-one nursing or midwifery care. Given health care costs, it is likely more expensive for hospitals to invest in the nursing staff necessary to implement routine intermittent auscultation for most patients. Thus, providers are concentrating on enhancing the quality of EFM, through improving the reliability of fetal heart rate interpretation with the promulgation of uniform definitions (National Institute of Child Health and Human Development Research Planning Group, 1997) and the standardization of EFM management using agreed-upon algorithms, with some exploration of the use of computerized neural network systems (Parer & King, 2000). There are no known disparities in the application of EFM in the population, and none appear in either the birth certificate or the LTMII data we analyzed. Given EFM's present universal use, it might now be difficult to enroll enough non-EFM subjects to find any potential disparities in EFM application. However, EFM's near-ubiquity remains an issue in itself, which could benefit from the application of public health expertise on the risks and benefits of screening technologies applied universally to large predominantly healthy populations.

Epidurals

Throughout history women have sought methods to deal with the pain of childbirth. To date, epidural analgesia/anesthesia (note: when used to assist women in labor, "analgesia" is the more appropriate term; if used for a cesarean, the larger dose results in an anesthetic effect) is the most effective method of relieving that pain. This section focuses on epidural analgesia use during vaginal births. Epidural analgesia is contraindicated for women who have low platelet levels (which can occur with severe pre-eclampsia or HELLP[1] syndrome), have recently used heparin, have an untreated infection or a skin infection at the site of injection, or have refractory low blood pressure.

Fully 91% of mothers in LTMII who received an epidural reported it was either very (81%) or somewhat (10%) helpful in reducing pain. In one study women rating their pain using a visual analog pain scale had less perceived pain when receiving epidural analgesia, but a meta-analysis of 21 randomized trials of epidural analgesia in labor compared to other methods, failed to demonstrate any difference in women's satisfaction with pain relief (Amin-Somuah & Howell, 2005). Therefore, given that women fail to cite pain relief as one of the most important factors in assessing their satisfaction with their childbirth experience (Hodnett, 2002) the idea that pain relief in labor is the major influence on satisfaction during labor and birth appears to be a misconception.

The relief of pain from epidural analgesia does not come without some risks and costs. Women choosing epidural analgesia must have an intravenous infusion, be confined to bed, have the fetal heart rate continuously monitored, have a bladder catheter, and also need frequent blood pressure checks. Women using epidural analgesia are also more likely to need oxytocin augmentation of their labor. They are at greatly increased (20 times) risk of experiencing at least one episode of hypotension (Amin-Somuah & Howell, 2005), which can lead to fetal heart rate bradycardia, and 1–2% of the time will result in the need for an emergent cesarean (Leighton & Halpern, 2002).

[1]The acronym HELLP was coined in 1982 to describe a syndrome consisting of hemolysis, elevated liver enzyme levels and low platelet count. The syndrome has been considered a variant of preeclampsia, but it can occur on its own or in association with preeclampsia (Padden, 1999).

As shown in several meta-analyses, the risks associated with epidural analgesia include longer labors for women overall by about 1 h, longer second stages of labor in 30–50% of women, and, increased risk of operative vaginal birth (Amin-Somuah & Howell, 2005; Leighton & Halpern, 2002; Lieberman & O'Donoghue, 2002). The latter may be due to the fact that infants of women with epidurals are more often in the occiput posterior position near the time of birth (Lieberman, Davidson, Lee-Parritz, & Shearer, 2005). Table 16.6 presents the results of two meta-analyses that calculated summary statistics. Other systematic reviews of epidural analgesia provide a more qualitative analysis of study methods and results (Lieberman & O'Donoghue).

Mothers in LTMII with a vaginal birth and an epidural reported an average length of labor of 10.4 h compared to 6.7 h for those without an epidural. In addition, LTMII respondents with epidurals had a 3–5 times increased risk of developing a fever greater than 38°C (100.4°F). Although the fever is unlikely to be infectious in origin, it is impossible to distinguish this fever from one of infectious origin. This situation often leads to overtreatment of the mothers with antibiotics and sepsis evaluations, as well as overtreatment of their neonates (Lieberman & O'Donoghue, 2002).

There is a substantial literature addressing whether or not epidural analgesia increases the risk of cesarean delivery. While the clinical consensus appears to suggest it does not, numerous methodological questions have been raised about the quality of the studies that found no relationship, including differences in labor management style among the settings in which trials took place and the typical management style in most American hospitals. A summary of the conflicting views on the relationship between epidurals and cesarean section has been published as part of a special issue of the *American*

Table 16.6 Meta-analysis table: effects of epidural vs. no epidural or no analgesia on labor

Source	Number of studies (N) (% not receiving/ receiving intervention)	Findings	Contextual factors	Disparities/ comments
Amin-Somuah and Howell (2005)	20 RCT studies for CS outcome Epidural group = 3,226 (15% no epidural) Non-epidural use group = 3,308 (23% had epidurals)	**Calculated summary statistic** **No difference** in cesareans, backache, low 5 min Apgar, maternal satisfaction, less naloxone to infants *Favors epidural use*: better pain relief (1 study) *Favors non-epidural use*: shorter 2nd stage, less instrumental births, less urinary retention, less catheterization, less maternal fever, less CS for fetal distress 1.42 [0.99, 2.03]	12 studies from N. America No analysis separating early epidurals (<4 cm) from late epidurals	No studies present data by race/ ethnicity or SES
Leighton and Halpern (2002)	14 RCTs and 2 prospective cohort studies for CS outcome and several maternal and neonatal outcomes. For CS, Epidural group = 2,161. Non-epidural group = 2,136. Cross-over not accounted for.	**Calculated summary statistic for CS only** No difference in cesareans *Favors epidural*: better pain relief, improved satisfaction, fewer 1-min apgar <7, less naloxone to infants *Favors non-epidural use*: shorter 2nd stage, less instrumental births, less oxytocin use, less fever, less hypotension	7 studies from N. America; 4 from the U.K., 3 from Scandanavia Parity: 8 used only nulliparas, 1 only multipara, and 5 had mixed parity 8 of the studies known to exclude induced labors	No studies present data by race/ ethnicity or SES

Journal of Obstetrics and Gynecology on pain relief in labor (Leighton & Halpern, 2002; Lieberman & O'Donoghue, 2002; Thacker & Stroup, 2002). Another review articulates the differences between oxytocin dosing regimens, noting that in the majority of epidural randomized clinical trials, there is a rapid dose increase to high maximum levels, compared to North American usage which is at a slower rate of increase with much lower maximum levels (Kotaska, Klein, & Liston, 2006).

The only national data currently available on the epidural rate in the U.S. comes from the *Listening to Mothers* surveys (epidurals are included on the new version of the birth certificate but less than half the states have adopted this form) which in 2002 (LTMI) reported a rate of 59%, and in 2005 (LTMII), a rate of 71%. Overall, white non-Hispanic mothers were most likely and black non-Hispanics least likely to receive an epidural with their vaginal birth (Table 16.7). However, black non-Hispanic mothers who did receive an epidural were most likely to rate it very helpful, Epidural rates for first time mothers were higher in all groups when compared to experienced mothers, but particularly so (46% points) among black non-Hispanic mothers. Mothers could also choose a category of "no pain medications for labor" and while the rates were generally much higher for experienced mothers, there were no major differences across racial and ethnic groups. Using the pregnancy files from the Medical Expenditure Survey 1996–2000, which is drawn from respondents to the National Health Interview Survey, a study of singleton, live vaginal births found no racial (black vs. white) differences in epidural use; however, it did document that women of Hispanic ethnicity were half as likely as non-Hispanic mothers to obtain an epidural (Atherton, Feeg, & el-Adham, 2004).

Epidural rates have been found to vary by insurance status, with publicly insured women less likely to receive epidural anesthesia (Atherton et al., 2004; Obst, Nauenberg, & Buck, 2001). Medicaid managed care in one state (TN, Tenncare) resulted in a sharp decrease of epidural use among covered women (Johnson, 1995). A New York state study demonstrated that insurance status was the strongest predictor of epidural use (Obst et al.). Epidural analgesia is costly, since it requires an additional physician or nurse anesthetist to attend the woman in labor, as well as enhanced nursing care to monitor and respond to potential adverse effects (Bell et al., 2001; Huang & Macario, 2002).

Table 16.7 Epidural usage and satisfaction and women using no pain-relieving drugs for a *vaginal birth* by Race/Ethnicity among mothers in the Listening to Mother II

Among vaginal births, use of epidurals and no pain-relieving medicine	White non-Hispanic (n = 667) %	Black non-Hispanic (n = 112) %	Hispanic (n = 251) %	All[a] (n = 1,077) %	p Value
Among all w/epidural	74	63	69	71	(0.009)
Among those rating epidural as 'very helpful'	75	90	82	77	(0.007)
Parity					
Epidural use – first-time mothers	81	97	77	82	(0.062)
Epidural use – experienced mothers	72	51	61	67	(0.000)
Payer source for those using epidural					
Medicaid payer	68	54	71	66	(0.073)
Private payer	80	85	67	74	(0.111)
Used no pain med during labor					
First-time mothers	10	6	7	8	(0.647)
Experienced mothers	25	24	30	21	(0.521)

Source: Adapted from Declercq, Sakala, et al. (2006)

[a] All includes self-described Asian and "other" mothers

Episiotomy

Episiotomy, an incision made in the area between the vagina and the anus as the baby's head is emerging, has been a routine part of obstetrical practice in the United States for nearly 90 years. Typical explanations found in obstetric texts over the years have argued that the use of episiotomy protects the pelvic floor from long-term damage resulting in pelvic floor relaxation, as well as preventing more serious kinds of perineal trauma. More than 20 years ago, randomized clinical trials of routine vs. restrictive use of episiotomy demonstrated that women in the routine episiotomy group experienced more anal sphincter and rectal tears, and that there were no differences in pelvic floor relaxation or associated urinary incontinence between the groups (Thacker & Banta, 1983). Several meta-analyses have been published of the RCTs related to episiotomy, the most recent in 2005 (Carroli & Belizan, 2000; Hartmann, Viswanathan, & Palmeri, 2005). All have concluded not only that women undergoing episiotomy are at greater risk for perineal trauma, but also that the use of episiotomy at birth has no effect on the integrity of the pelvic floor. Most recently the American College of Obstetrician-Gynecologists practice bulletin on episiotomy emphasized that there is no evidence to support the liberal use of episiotomy, and that it should be limited either to fetal indications of a potential difficult birth or to maternal indications of potential severe trauma (ACOG, 2006a).

Despite the evidence, episiotomy remains a common obstetric procedure, though one that has declined in recent years. Studies using national hospital discharge data show a decrease in episiotomies from over two million in 1980 to 667,000 in 2004, this translates into 64% of vaginal births accompanied by an episiotomy in 1980 to only 23% in 2004 (Kozak et al., 2006). Given that reported episiotomy rates from hospital discharge data are only 70–84% sensitive compared to medical record abstraction (Lydon-Rochelle et al., 2005; Yasmeen, Romano, Schembri, Keyzer, & Gilbert, 2006), the actual episiotomy rate is likely higher. Intriguingly, the *Listening to Mothers II* survey found a comparable overall rate of 25% in 2005, with first-time mothers reporting an episiotomy rate of 36% while experienced mothers reported a rate of 20% (Table 16.8). The LTMII survey also found disparities in episiotomy rates across racial and ethnic groups, with a low of 18% among black non-Hispanic women, to a high of 27% among white non-Hispanic women. Among primiparous women the disparities are even more pronounced, with a rate of 40% among white non-Hispanic women compared to 25 and 27% for black non-Hispanic and Hispanic women, respectfully. The survey also demonstrates that in general, women are not given a choice regarding the episiotomy procedure, as only 18% of women reported being given such a choice.

Episiotomy rates appear to vary significantly by provider specialty. In LTMII, among first time mothers with vaginal births, the highest reported use of episiotomy was by obstetricians (44%), compared to family practice physicians (25%), and midwives (13%). A study of one institution from 1995 to 2000 showed that 55% of nulliparous women with a term, singleton, vertex birth had an episiotomy, but that private obstetricians had a rate seven times higher than that of academic physicians or residents (Howden, Weber, & Meyn, 2004).

Table 16.8 Episiotomy by Parity and Race/Ethnicity, U.S., 2005

Episiotomy in vaginal births	White non-Hispanic % (n = 665)	Black non-Hispanic % (n = 112)	Hispanic % (n = 251)	All % (n = 1,077)	p Value
All	27	18	21	25	(0.039)
Primips	40	25	27	36	(0.044)
Multips	21	15	18	20	(0.419)
Mother had a choice about episiotomy	18	5	22	18	(0.241)

Source: Adapted from Declercq, Sakala, et al. (2006)

Table 16.9 Meta-analysis table: episiotomy

Source	Number of studies (N) (% receiving episiotomy)	Findings	Contextual factors	Disparities/ comments
Banta and Thacker (1983)	Reviewed approx. 400 studies, no RCT studies	*No difference in*: ease of repair, third degree lacerations, and prevention of fetal brain damage *Associated risk of*: pain, dyspareunia, edema, infection	– 250 studies deemed acceptable, authors do not go into specifics about these studies – Historical context of episiotomy presented	No studies present data by race/ ethnicity or SES
Carroli and Belizan (2000)	6 RCT studies Restrictive group = 2,441 (27%) Routine group = 2,409 (73%)	**Calculated summary statistic** **No difference** in severe perineal/vaginal trauma, dyspareunia, urinary incontinence, severe maternal pain *Favors restricted use*: less posterior trauma, less suturing, fewer healing complications *Favors routine use*: less anterior trauma	– All but one study used mediolateral episiotomy as routine – Settings: Argentine, Saudi Arabia, United Kingdom, Canada – Longer term outcomes (3 months and 3 years) limited to one study	No studies present data by race/ ethnicity or SES
Hartmann et al. (2005)	RCT and prospective cohort studies; 7 RCTs – same as above with 2004 study, N = 146	*Synthetic meta-analysis fair to good evidence of no difference*: severity of laceration, pain, and pain medication use *Favors restricted use*: less dysparuenia **Fair to poor evidence episiotomy does not protect against fecal or urinary incontinence**	As above with new study from Germany 3–4 prospective cohort data add to the robustness of this finding; follow-up 3 months to 4 year post-birth	No studies present data by race/ ethnicity or SES

In an analysis of National Hospital Discharge Data from 1979 to 1997, use of episiotomies was found to vary by hospital size and geographic region (lowest rates in the West), race and payer source (Weber & Meyn, 2002). Adjusting for the previously mentioned variables, as well as age, marital status, and year of birth, black women were 54% less likely than white women, and women with government insurance were 46% less likely than women with private insurance to have an episiotomy. These disparities may be due to larger numbers of black women and poor women living in urban settings and having an increased likelihood of birthing in a teaching hospital where episiotomy rates are lower (Garcia, Miller, Huggins, & Gordon, 2001).

Table 16.9 summarizes the results of a systematic review (Banta & Thacker, 1983) and two meta-analyses of RCTs and prospective cohort studies (Carroli & Belizan, 2000; Hartmann et al., 2005) on the use and outcomes of episiotomies. These analyses find little support for routine use of episiotomy. Although there has been sufficient evidence against routine use of episiotomies since 1986, only recently have major provider groups and policy makers urged adoption of an evidence-based approach to management of the perineum. Nonetheless, rates of episiotomy use remain consistently above 20%, with even higher rates among first time mothers.

Cesarean Birth and Vaginal Births After Cesareans

The research literature on the appropriateness of cesarean section has, in general terms, gone through four phases. Starting in the mid 1970s, concerns arose about increasing cesarean section

rates leading to discussions of both the reasons for the growth and recommendations for reducing cesareans. In September 1980, the National Institute of Child Health and Human Development sponsored a Consensus Development Conference on cesarean birth (National Institutes of Health, 1981). One of the notable recommendations of the conference was support for an increase in the use of vaginal birth after cesarean (VBAC) reflecting a series of studies published in the late 1970s and into the 1980s, citing evidence of the safety of VBACs (Flamm, Newman, Thomas, Fallon, & Yoshida, 1990). After this, VBAC rates began to climb (see Fig. 16.1), reaching almost 30% of all births to mothers with a prior cesarean in the mid 1990s (Martin et al., 2007). Subsequently, research (Lydon-Rochelle, Holt, Easterling, & Martin, 2001; Mozurkewich & Hutton, 2000), sounding boards (Sachs et al., 1999), and editorials (Greene, 2001) questioning the safety of VBACs emerged. In 1997, VBAC rates began to decrease slowly and then more rapidly, reaching the lowest point since 1989, when rates were first recorded (Fig. 16.2). Since the latter part of the 1990s and continuing as this chapter is being written, the very safety of vaginal birth itself is being questioned with an increased an emphasis on the benefits of elective cesarean section (Bump, 2002; O'Boyle, Davis, & Calhoun, 2002), for the integrity of the pelvic floor. As in 1980, NIH sponsored another conference in March 2006, to address these emerging issues through an examination of a concept termed "Cesarean delivery by maternal request" (National Institutes of Health, 2006). The NIH conference explored current research to identify whether cesarean delivery might be a preferred alternative to vaginal birth for low risk women. The conclusions of the expert panel emphasized the lack of knowledge concerning the outcomes of what they termed "planned cesareans," and the need for systematic research to help better resolve the safety of cesareans to mothers without a medical indication.

Trends in Cesarean Birth and VBAC

National cesarean rates have increased in most industrialized countries though few experienced either the decline that occurred in the U.S. from 1989 to 1996 or such a rapid increase since 1996 reaching an all time record for the United States of 31.1% in 2006, and resulting in the U.S. having the third highest rate among industrialized countries in 2005 trailing only Italy and South Korea (Menacker, Declercq, & MacDorman, 2006).

When national cesarean rates in the U.S. are broken down by subgroups (Fig. 16.1), an even more interesting pattern emerges. Black non-Hispanic mothers began the 1990s with a cesarean rate

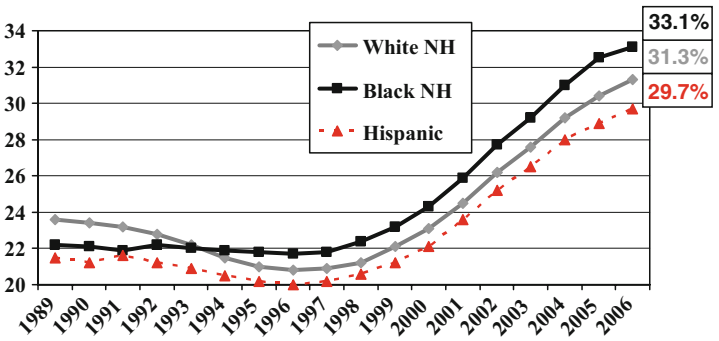


Fig. 16.1 Overall cesarean rates by Race/Ethnicity, U.S., 1989–2006. *Source*: Adapted from Martin et al. (2007) and Hamilton et al. (2007)

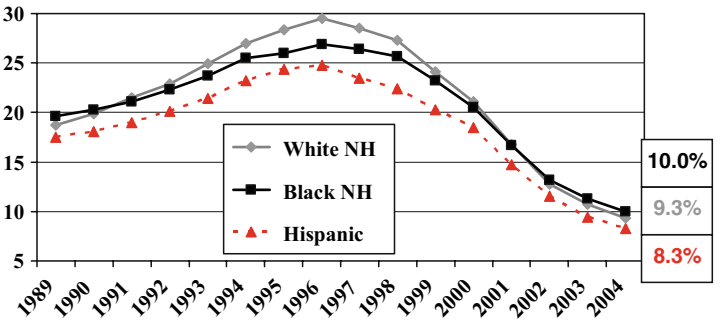


Fig. 16.2 VBAC rates by Race/Ethnicity, U.S. 1989–2004 (NCHS has not reported national VBAC rates for 2005 and 2006). *Source*: Adapted from Martin et al. (2006)

almost 2% points lower than that of white non-Hispanic mothers in the U.S. However, by 2006 their rate was 1.8% points higher than whites. How did this change come about? The decline in cesareans in the early 1990s that was experienced by whites and Hispanics did not occur among black mothers whose cesarean rate remained steady at about 22% for the period from 1989 to 1997. Therefore, by 1994, black non-Hispanic mothers had the highest cesarean rate of any of the three subgroups. When rates began to rise starting in 1997, however, rates for black non-Hispanic mothers rose as rapidly as those in the other groups, so that by 2006 their cesarean rate was 33.1% compared to 31.3 and 29.7% respectively for white non-Hispanic and Hispanic mothers (Hamilton et al., 2007).

The national rate of VBACs in the U.S. shows a similar, but inverse overall pattern (Fig. 16.2), with an increase from 1989 to 1996 and a rapid decline thereafter. However, the subgroup patterns are different from those of overall cesareans. Hispanic mothers, who had distinctly lower cesarean rates, also had lower VBAC rates (and hence higher repeat cesarean rates) throughout the period studied. White non-Hispanic mothers for most of the period had the highest VBAC rates, reaching a total of 30% in 1996, 5% points higher than Hispanic mothers. White non-Hispanic mothers then experienced the most rapid decline after 1996 and by 2006, black non-Hispanic mothers had the highest VBAC rate (10%) among all three groups.

In such analyses there is always the dilemma of how much to emphasize the similarities compared to the differences. It is important to note that overall the same general pattern emerged among all three groups – a decrease in primary cesareans with a corresponding increase in VBAC from 1989 to 1996, followed by a reversal of those trends from 1996 to 2006. When comparing subgroups however, different patterns are seen, particularly for black non-Hispanic and Hispanic mothers. Black non-Hispanic mothers in 2004 had an overall cesarean rate 12% higher than Hispanic mothers, and a primary cesarean rate 24% higher (data not shown).

Table 16.10 presents the data for the three subgroups for *first-time mothers* broken down by age for the periods 1991–1996 and 1996–2005. What is striking about the age breakdown is that when one examines even the 1991 rates by age, black non-Hispanic first-time mothers' cesarean rates are the same as the rates of white non-Hispanic mothers among mothers under 20 only, but are consistently higher in every other age group. What keeps the black non-Hispanic overall rates lower in 1991 is the much larger proportion of births to younger mothers. A similar, though not quite as pronounced pattern is seen comparing white non-Hispanic and Hispanic mothers in 1991. The data from 1991 to 1996 clearly illustrate the substantially smaller declines experienced by black non-Hispanic mothers in all age groups during this period compared to others. However, for the 1996–2005 period, the increases are substantial for white non-Hispanics (41%), black non-Hispanics (43%), and to a lesser extent, Hispanics (33%).

There are interesting findings at both ends of the age spectrum. Table 16.10 illustrates the extraordinarily high primary cesarean rates for older mothers in 2005 with over half of first-time mothers 40 and older having a cesarean, including rates of 62 and 59% for black non-Hispanic and Hispanic mothers, respectively. The largest shifts across the two periods are generally seen among teenage mothers, with non-Hispanic white (41% increase) and black (43% increase) teens' cesarean

Table 16.10 Proportion of cesareans[a] to first time mothers by Age, and Race/Ethnicity, United States, 1991, 1996, 2005; and percent change between 1991–1996, and 1996–2005

Age and race/ethnicity	% Cesareans to first-time mothers			Percent change 1991–1996	Percent change 1996–2005
	1991	1996	2005		
All races[b]					
Total	23.8	21.3	29.8	–11	40
Under 20	16.8	14.7	20.4	–13	39
20–24	22.0	19.2	25.6	–13	33
25–29	26.9	23.3	30.7	–13	32
30–34	31.6	28.4	37.5	–10	32
35–39	38.6	35.0	46.4	–9	33
40–49[c]	46.4	42.6	57.6	–8	35
White, non-Hispanic[d]					
Total	24.4	21.5	28.3	–12	41
Under 20	17.1	14.4	19.0	–16	44
20–24	22.2	19.0	23.2	–14	33
25–29	26.4	22.7	28.1	–14	32
30–34	30.9	27.4	34.6	–11	34
35–39	37.7	33.5	43.3	–11	35
40–49[c]	45.2	40.6	52.2	–10	39
Black, non-Hispanic[d]					
Total	23.2	22.4	29.6	–3	43
Under 20	17.1	16.6	21.2	–3	40
20–24	23.7	22.3	28.0	–6	36
25–29	31.9	30.0	35.5	–6	29
30–34	37.9	36.3	43.4	–4	27
35–39	42.4	44.0	52.4	4	26
40–49[c]	51.1	52.8	61.8	3	21
Hispanic[d,e]					
Total	22.5	19.8	26.4	–12	33
Under 20	16.2	14.0	18.3	–14	31
20–24	21.5	19.0	24.2	–12	27
25–29	29.4	25.4	31.1	–14	22
30–34	35.2	31.8	39.5	–10	24
35–39	43.5	38.8	49.9	–11	29
40–49[c]	47.6	46.2	58.6	–3	27

Adapted and updated from Declercq, Menacker, and MacDorman (2006)
Source of update for 2005: Centers for Disease Control and Prevention (2008b)
[a] Number of cesareans per 100 live births
[b] Includes races other than white and black and Hispanic origin not stated
[c] Beginning in 1997, data are for women aged 40–54 years
[d] In 1991 New Hampshire did not report Hispanic origin. For 1991 all births in this State were assumed to be non-Hispanic (in 1993, the first year that data were available, 99.0% of births in New Hampshire were to non-Hispanic women)
[e] Includes all persons of Hispanic origin of any race

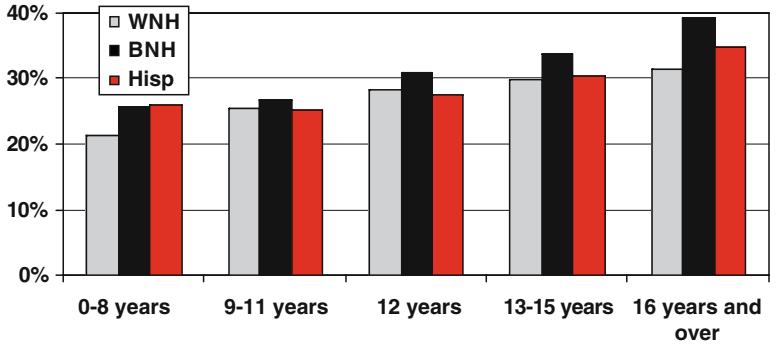


Fig. 16.3 Cesarean rates in singleton births to first-time mothers, by Education & Race/Ethnicity, U.S., 2005. *Source*: Centers for Disease Control and Prevention (2008b)

rates increasing substantially from 1996 to 2005. Notably, in 2005, one in five teens having their first birth had a cesarean and, given the constraints on VBACs, will likely continue to have cesareans in the future.

One of the constraints in using U.S. national data is that there are limited variables on the birth certificate to measure social status and its relationship to outcomes of interest, with maternal education being one of the few. Figure 16.3 presents data on c-section rates for singleton births to first time mothers (meaning by definition these are primary cesareans) by maternal race, ethnicity and education. Consistently, black non-Hispanic mothers have the highest primary cesarean rates at every education level. Among mothers with at least 16 years of education, black non-Hispanic mothers have a cesarean rate more than 8% points higher than that of white non-Hispanic or Hispanic mothers. Hispanic mothers had the lowest cesarean rates for the middle three education categories (9–15 years). Among those with 8 or fewer years of education, white non-Hispanic mothers have the lowest rate.

Another surrogate measure for the socio-economic status of mothers is whether their delivery is paid for by public or private insurance. While such data are not available nationally, Massachusetts (MA) does include such a measure on its birth certificate. Using these data, Figs. 16.4 and 16.5 present the overall cesarean rates by race/ethnicity for singleton births in Massachusetts from 1989 to 2006, demonstrating the same pattern in total cesarean rates as for the U.S as a whole. Figure 16.4 represents those mothers whose births were paid for through public sources, and Fig. 16.5 represents mothers using private insurance. Once again black non-Hispanic mothers have the highest rates in each case. Among those using public payment, white non-Hispanic mothers have higher rates than Hispanic mothers; while among those with private payments the differences between white non-Hispanic and Hispanic mothers are negligible. Looking across the two figurers it is also apparent that cesarean rates are consistently higher among those with private insurance than those with public. By 2006 the distinction in rates among those with private insurance had grown to 11% points higher among Hispanic mothers, seven among white non-Hispanic mothers and 6% points among black non-Hispanic mothers.

Maternal Perspectives and Attitudes Related to Childbirth

Mothers' perspectives on the birth process likely influence their openness to having a cesarean. The *Listening to Mothers II* survey included a series of questions concerning mothers' attitudes related to childbirth; select results are summarized in Table 16.11. Among first-time mothers

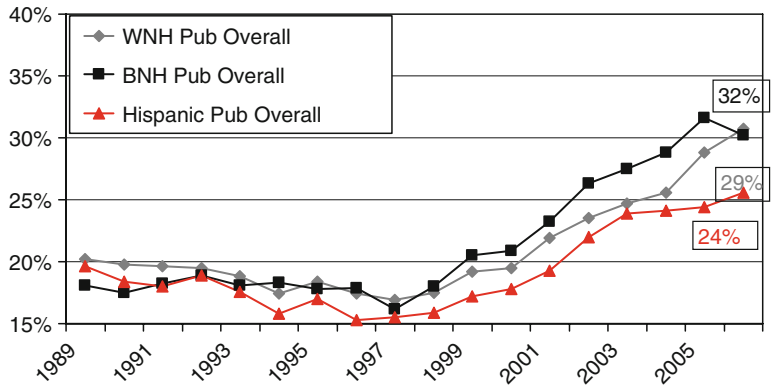


Fig. 16.4 Cesarean rates in singleton births by Public Payment & Race/Ethnicity, Mass., 1989–2006. *Source*: Massachusetts Natality files available through MassCHIP version 3.00r317

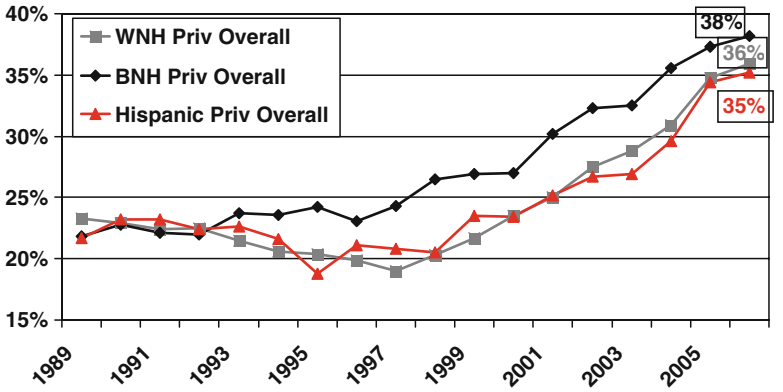


Fig. 16.5 Cesarean rates in singleton births by Private Payment & Race/Ethnicity, 1989–2006. *Source*: Massachusetts Natality files available through MassCHIP version. 3.r323. Massachusetts Department of Public Health

who had a vaginal birth, white non-Hispanic mothers were much less likely to report feeling capable, and black non-Hispanic mothers were least likely to report feeling frightened while giving birth. Mothers were presented with the statement, *Birth is a process that should not be interfered with unless medically necessary*; white non-Hispanic mothers were less likely to agree with that statement. Interestingly, in response to a hypothetical question about having a future elective primary cesarean, Hispanic mothers (14%) were more likely to respond positively, though the rates were low for all three groups. Mothers were also asked a series of questions about a mother's right to have a VBAC, cesarean or vaginal birth. The responses to the question on the right of a mother to have a cesarean by the women in all three groups were largely similar with about 46% agreement in each group. Mothers also were asked about the impact of malpractice on provider behavior and one of the statements involved the performance of unnecessary cesareans. White non-Hispanic mothers were much more likely than black or Hispanic mothers to agree that malpractice can lead to unnecessary cesareans. The survey also asked mothers if they felt pressured to have a cesarean. While only 9% of mothers overall

Table 16.11 Maternal attitudes toward birth process and policy by race/ethnicity

Attitudes	White non-Hispanic (n = 980) %	Black non-Hispanic (n = 191) %	Hispanic (n = 329) %	All (n = 1,573) %	p Values
Feelings while giving birth					
Capable (1st time mothers – vaginal birth)	39	47	55	43	(0.000)
Frightened (1st time mothers – vaginal birth)	38	26	46	37	(0.135)
Attitudes toward birth process and policy					
Birth process should not be interfered with unless medically necessary	46	58	58	50	(0.000)
Likely to have elective primary cesarean in future	6	6	14	8	(0.000)
Woman has a right to a cesarean	45	52	41	46	(0.118)
Malpractice concerns might lead doctors to perform unnecessary cesareans	45	34	28	42	(0.000)
(Had primary cesarean) Felt pressure to have a cesarean (n = 231)	24	10	50	26	(0.000)
U.S. maternity system good/excellent	86	76	84	84	(0.001)

Source: Adapted from Declercq, Sakala, et al. (2006)

reported feeling pressured, mothers who received a primary (26%) or repeat (25%) cesarean were much more likely to report feeling pressured. In the case of primary cesareans, responses varied widely by race/ethnicity, with half of the Hispanic mothers in that category (n = 40) indicating they felt pressured. Finally, mothers were asked to rate the U.S. maternity care system and while all rated it highly, black non-Hispanic mothers rated it less positively than either white non-Hispanic or Hispanic mothers.

As an example of the evidence base for outcomes associated with cesareans, Table 16.12 presents a summary of a large multi-center randomized clinical trial examining outcomes associated with cesareans for breech births. The three studies published from this one trial (Hannah et al., 2000, 2004; Whyte et al., 2004) report both short term and long term health outcomes and did not address disparities. The initial finding of greater safety for use of cesareans for breech presentations did impact obstetrical practice as the rate of cesareans for breech positions increased to 85% by 2000 (CDC, 2008b). However, later findings that these differences disappeared in the long run was not greeted with a decline in cesareans for breech presentation with the cesarean rate for breech births reaching 88% by 2005 (CDC, 2008b).

For comparability with other chapters in this book, the quality of the studies on induction and planned cesarean for breech position reviewed in Tables 16.3 and 16.12, are presented in Tables 16.13 and 16.14. All of the studies rated in the "good" category. However, two points should be made about the quality ratings. First, the results of studies with a 'good' quality rating may not be generalizable in the North American context of obstetrical practice given differences in labor management style as noted above in the discussion of the effect of epidurals on the cesarean rate. Secondly, although studies may be of good quality, whether their findings affect actual childbirth practices seems to depend largely on the willingness of providers to 'believe' the results; that is, often clinicians readily adopt those findings that confirm what they already are doing or want to do, and ignore those that require a distinct change in practice. It appears that external pressures are sometimes needed to change behavior, such as hospitals adopting certain benchmark practices for ongoing credentialing, or insurers using rates of certain interventions or outcomes in pay-for-performance (Perlman, 2006).

Table 16.12 Major outcomes associated with studies of planned cesarean for breech

Health status outcome #1 short-term morbidity/mortality

Author (year)	Study design	Study type	Description of intervention: what, how and where	Populations studied (ages included, race and ethnicity) and sample size	Address disparities (yes/no)	Key findings related to intervention effectiveness (OR with CI)	Caveats/biases	Findings support the intervention? Yes/no For which populations?
Hannah et al. (2000)	RCT	Published paper	Singleton, live fetus in frank or complete breech at term. Exclusion criteria: small pelvis, ≥4,000 g, hyperextension of vertex, contraindication to vaginal birth, lethal anomaly.	121 centers in 26 countries N/S. Am, Europe, Mideast, SE Asia, Africa – 16 low PMR – 10 high PMR Less than 10% assess pelvis with X-ray, CT, or MRI 30% assess head flexion clinically	**No**	**Planned CS** Perinatal/neonatal mortality: RR 0.23 (0.07–0.81) RD: –1% Serious neonatal morbidity: RR 0.36 (0.19–0.63) Maternal mortality/ morbidity: RR 1.24 (0.79–1.95)	Among vaginal birth deaths: 2 did not meet study criteria: stillbirth before enrollment 2 neonatal deaths from unrelated causes: SIDS, diarrhea Among 3 CS deaths, 1 death assoc w/heart rate abnormality and ruptured myeloeningocele	Corrected for 2 deaths not meeting inclusion criteria: 66% decreased risk of neonatal death or serious morbidity with planned CS. No difference for maternal morbidity outcomes by route of birth.

(continued)

Table 16.12 (continued)

Health status outcome #1 short-term morbidity/mortality								
Health status outcome #2 long term morbidity/mortality								
Author (year)	Study design	Study type	Description of intervention	Populations studied and sample size	Address disparities (yes/no)	Key findings related to intervention effectiveness (OR with CI or p values reflecting the intervention-outcome relationship)	Caveats/biases	Findings support the intervention? Yes/no For which populations?
Hannah et al. (2004) Whyte et al. (2004)	RCT	Published paper	Same study as above except outcomes assessed 2 years after birth	55% original study pop. delivered at a center doing 2 year follow-up; 85 centers in 18 countries Size available for analysis at 2 years: Infants CS N = 457 Vag N = 463 Mothers CS N = 457 Vag N = 460		Infant outcomes were death or neurodevelopmental Planned CS RR 1.09 (0.52–2.3) (Rate 3.1% vs. 2.8% in vaginal group) 6% more CS infants had reported medical problems in the past several months [RR 1.41 (1.05–1.89)] Maternal outcomes: No difference in breastfeeding, relationships, pain, incontinence, depression, menstrual or sexual problems, fatigue, or subsequent pregnancies. Planned CS assoc w/higher risk of constipation	Compared to original study population, follow-up population more likely to be >30, be birth in low PMR country, and less likely to be nulliparas FU sample size had 80% power to only detect a difference of 2.5%	Long-term findings show no difference between infant and maternal outcomes between planned CS for breech vs. vaginal birth

Table 16.13 Quality rating of studies associated with induction

Health status outcome: induction							
Author (year)	Reporting	External validity	Internal validity – bias	Internal validity – confounding	Power	Total quality score ≤14 = poor 15–19 = fair ≥20 = good	Suitability of study to assess effectiveness: greatest, moderate, least
Hannah et al. (1992)	12	2	6	6	2	28	Greatest
Heimstad et al. (2007)	12	3	6	6	2	29	Greatest

Table 16.14 Quality rating of studies associated with planned CS for breech

Health status	Author (year)	Reporting	External validity	Internal validity – bias	Internal validity – confounding	Power	Total quality score ≤14 = poor 15–19 = fair ≥20 = good	Suitability of study to assess effectiveness: greatest, moderate,least
Short-term outcomes	Hannah et al. (2000)	10	1	5	6	2	23	Greatest
Long-term outcomes mothers	Hannah et al. (2004)	11	2	5	6	0	24	Greatest
Long-term outcomes children	Whyte et al. (2004)	11	2	5	6	1	25	Greatest

Conclusion: Disparities in the Use and Effects of Intrapartum Interventions

The elimination of disparities in outcomes and the provision of health services based on the best available evidence are significant goals for public health practice. In this chapter we have examined contemporary maternity practice by focusing on five major interventions: induction, fetal monitoring, epidurals, episiotomy and cesarean birth (Table 16.15). Three clear findings emerge: (1) while there has been considerable research on each of these interventions, actual practice is not consistently related to its associated evidence base; (2) randomized trials have not examined the relationship of these interventions and disparities in outcomes; and (3) in all cases but fetal monitoring, which is virtually universally applied, there are differences in the application of the interventions to mothers from different race/ethnicity groups. However, there is also no clear pattern that would suggest that one group is more likely than any other to receive evidenced-base care. To a small degree, white non-Hispanic mothers are more likely and Hispanic mothers less likely to receive interventions, but these interventions are not necessarily evidence-based. For example, white non-Hispanic mothers have the highest rates of epidurals, medical inductions, and episiotomies while black non-Hispanic mothers have the highest rates of fetal monitoring and cesarean section, and Hispanic mothers are least likely to be induced or have a cesarean. There were substantial differences in maternal attitudes toward the birth process though not in a consistent pattern (e.g., one group always favoring more or less intervention).

It is unclear what effect evidence based research findings have had on the U.S. use of the obstetrical interventions described above. Major studies questioning the routine use of electronic fetal monitoring and episiotomy have been published for more than two decades. While the episiotomy rate has declined substantially in that time, one in four mothers with a vaginal birth in 2005 still experienced an episiotomy. Fetal monitoring is ubiquitous. Although a considerable amount of

Table 16.15 Summary of key findings on perinatal interventions and attitudes by Race/Ethnicity by data source

	Source	White non-Hispanic %	Black non-Hispanic %	Hispanic %	All %	p Values
Experienced medical induction	LTMII	48	38	26	41	(0.003)
Used fetal monitoring	LTMII	96	100	98	97	(0.135)
Had an epidural for a vaginal birth	LTMII	74	63	69	71	(0.009)
Episiotomy done in a vaginal birth	LTMII	27	18	21	25	(0.039)
Overall cesarean rate – 2006	NCHS	31	33	30	31	(0.000)
Agreed that "Birth process should not be interfered with unless medically necessary"	LTMII	46	58	58	50	(0.000)

Source: Adapted from Declercq, Sakala, et al. (2006) and Hamilton et al. (2007)

literature was published in the 1990s concerning VBACs, the primary cesarean rate, which was largely ignored in the research literature, also shifted dramatically, in a pattern precisely inverse to the VBAC rate. The disconnect between evidence and medical practice in general, and obstetrics in particular has been widely noted. Efforts to bridge this gap require that the public health community promote the use of evidence in setting practice guidelines to advance the quality, efficiency, and cost-effectiveness of the health care system.

In other circumstances one might now suggest remedies that would eliminate or minimize disparities to ensure that all groups experience the best possible care. In the case of maternity care, that becomes more complex since it is not at all clear who is receiving the most evidenced based care, and hence which way change should occur. Likewise, it is not clear that any differences across groups noted here are the result of discrimination by race/ethnicity as opposed to individual practice patterns and maternal differences in expectations. For many women, the state of pregnancy and the childbirth process provides their first adult contact with the health care system. Prenatal care, intrapartum and postpartum care present opportunities for the public health system to ensure access to evidence-based quality care for all mothers. There is an urgent need to open a dialogue with women and their health care providers. Women's views concerning the process of care during pregnancy and childbirth should be explored to enhance our understanding of the basis for the profound differences observed in attitudes toward a universal process such as childbirth. Providers' insights regarding obstetric care practice patterns and their inconsistent relationship to evidence are also required to improve perinatal outcomes for all women and infants and to reduce disparities between mothers and infants of different racial and ethnic groups.

References

Alfirevic, Z., Devane, D., & Gyte, G. M. L. (2006). Continuous cardiotocography (CTG) as a form of electronic fetal monitoring (EFM) for fetal assessment during labour. *Cochrane Database of Systematic Reviews*, (3), Art. No. CD006066.

American College of Obstetricians and Gynecologists Committee on Obstetric Practice. (2006a). ACOG Practice Bulletin No. 71. Episiotomy. *Obstetrics and Gynecology, 107*, 957–962.

American College of Obstetricians and Gynecologists Committee on Obstetric Practice. (2006b). ACOG Committee Opinion No. 342: Induction of labor for vaginal birth after cesarean delivery. *Obstetrics and Gynecology, 108*(2), 465–468.

Amin-Somuah, M. S. R., & Howell, C. (2005). Epidural versus non-epidural or no analgesia in labour. *Cochrane Database of Systematic Reviews*,(4), Art. No. CD000331.

Atherton, M. J., Feeg, V. D., & el-Adham, A. F. (2004). Race, ethnicity, and insurance as determinants of epidural use: Analysis of a national sample survey. *Nursing Economics, 22*(1), 6–13.

Banta, D., & Thacker, S. B. (1983). The risks and benefits of episiotomy: A review. *Birth, 9*(1), 25–30.

Beebe, L. A., Rayburn, W. F., Beaty, C. M., Eberly, K. L., Stanley, J. R., & Rayburn, L. A. (2000). Indications for labor induction. Differences between university and community hospitals. *Journal of Reproductive Medicine, 45*(6), 469–475.

Bell, J. S., Campbell, D. M., Graham, W. J., Penney, G. C., Ryan, M., & Hall, M. H. (2001). Do obstetric complications explain high caesarean section rates among women over 30? A retrospective analysis. *BMJ, 322*(7291), 894–895.

Bloom, S. L., Spong, C. Y., Thom, E., Varner, M. W., Rouse, D. J., Weininger, S., et al. (2006). Fetal pulse oximetry and cesarean delivery. *New England Journal of Medicine, 355*(21), 2195–2202.

Bump, R. C. (2002). Advising prospective mothers about the maternal morbidity of vaginal childbirth [see comment]. *American Journal of Obstetrics and Gynecology, 187*(4), 823.

Carroli, G., & Belizan, J. (2000). Episiotomy for vaginal birth. *Cochrane Database of Systematic Reviews*, (2), Art. No. CD000081.

Centers for Disease Control and Prevention, Division of Reproductive Health. (2008a). *Pregnancy Risk Assessment Monitoring System (PRAMS): Home*. Retrieved February 28, 2008, from http://www.cdc.gov/prams/.

Centers for Disease Control and Prevention. National Center for Health Statistics. (2008b). *VitalStats*. Retrieved February 28, 2008 and March 18, 2008, from http://www.cdc.gov/nchs/vitalstats.htm.

Chalmers, I., Enkin, M., & Keirse, M. J. N. C. (Eds.). (1989). *Effective care in pregnancy and childbirth*. New York, NY: Oxford University Press.

Crowley, P. (2000). Interventions for preventing or improving the outcome of delivery at or beyond term. *Cochrane Database of Systematic Reviews*, (2), Art. No. CD000170.

Davidoff, M. J., Dias, T., Damos, K., Russell, R., Bettegowda, V. R., Dola, S., et al. (2006). Changes in gestational age distribution among U.S. singleton births: Impact on rates of late preterm birth, 1992–2002. *Seminars in Perinatology, 30*, 8–15.

Declercq, E., Menacker, F., & MacDorman, M. (2006). Maternal risk profiles and the primary cesarean rate in the United States, 1991–2002. *American Journal of Public Health, 96*(5), 867–872.

Declercq, E. R., Sakala, C., Corry, M. P., & Applebaum, S. (2006). *Listening to Mothers II: Report of the second national U.S. survey of women's childbearing experiences* (pp. 1–104). New York: Childbirth Connection.

Devries, R., Benoit, C., van Teijlingen, E. R., & Wrede, S. (2001). *Birth by design: Pregnancy, maternity care, and midwifery in North America and Europe*. New York: Routledge.

Flamm, B. L., Newman, L. A., Thomas, S. J., Fallon, D., & Yoshida, M. N. (1990). Vaginal birth after cesarean delivery: Results of a 5-year multicenter collaborative study. *Obstetrics and Gynecology, 76*(5 Pt 1), 750–754.

Garcia, F. A., Miller, H. B., Huggins, G. R., & Gordon, T. A. (2001). Effect of academic affiliation and obstetric volume on clinical outcome and cost of childbirth. *Obstetrics and Gynecology, 97*(4), 567–576.

Graham, E. M., Petersen, S. M., Christo, D. K., & Fox, H. E. (2006). Intrapartum electronic fetal heart rate monitoring and the prevention of perinatal brain injury. *Obstetrics and Gynecology, 108*(3 Pt 1), 656–666.

Greene, M. F. (2001). Vaginal delivery after cesarean section – is the risk acceptable? [see comment]. *New England Journal of Medicine, 345*(1), 54–55.

Greene, M. F. (2006). Obstetricians still await a deus ex machina. *New England Journal of Medicine, 355*(21), 2247–2248.

Hamilton, B. E., Martin, J. A., & Ventura, S. J. (2007). *Births: Preliminary data for 2006*. National Vital Statistics Reports (Vol. 56, No. 7, pp. 1–18). Hyattsville, MD: National Center for Health Statistics.

Hannah, M. E., Hannah, W. J., Hellman, J., Hewson, S., Milner, R., Willan, A., et al. (1992). Induction of labor compared with serial antenatal monitoring in post-term pregnancy. *New England Journal of Medicine, 326*, 1587–1592.

Hannah, M. E., Hannah, W. J., Hewson, S. A., Hodnett, E. D., Saigal, S., & Willian, A. R. (2000). Planned caesarean section versus planned vaginal birth for breech presentation at term: A randomized multicentre trial. *Lancet, 356*, 1375–1383.

Hannah, M. E., Ohlsson, A., Farine, D., Hewson, S. A., Hodnett, E. D., Myhr, T. L., et al. (1996). Induction of labor compared with expectant management for prelabor rupture of the membranes at term. TERMPROM Study Group [see comment]. *New England Journal of Medicine, 334*(16), 1005–1010.

Hannah, M. E., Whyte, H., Hannah, W. J., Hewson, S. A., Amankwah, K., Cheng, M., et al. (2004). Maternal outcomes at 2 years after planned cesarean birth versus planned vaginal birth for breech presentation at term: The international randomized term breech trial. *American Journal of Obstetrics and Gynecology, 191*, 807–871.

Hartmann, K., Viswanathan, M., & Palmeri, R. (2005). Outcomes of routine episiotomy: A systematic review. *JAMA, 293*(17), 2141–2148.

Healthpeople.gov [homepage on the internet]. Retrieved December 22, 2006, from http://www.healthypeople.gov/Default.htm.

Heimstad, R., Skogvoll, E., Mattsson, L., Johansen, O. J., Eik-Nes, S. H., & Salvesan, K. A. (2007). Induction of labor or serial antenatal fetal monitoring in postterm pregnancy. *Obstetrics and Gynecology, 109*(3), 609–617.

Hodnett, E. D. (2002). Pain and women's satisfaction with the experience of childbirth: A systematic review. *American Journal of Obstetrics and Gynecology, 186*(5 Suppl. Nature), S160–S172.

Howden, N. L., Weber, A. M., & Meyn, L. A. (2004). Episiotomy use among residents and faculty compared with private practitioners. *Obstetrics and Gynecology, 103*(1), 114–118.

Huang, C., & Macario, A. (2002). Economic considerations related to providing adequate pain relief for women in labour: Comparison of epidural and intravenous analgesia. *Pharmacoeconomics, 20*(5), 305–318.

Irion, O., & Boulvain, M. (2000). Induction of labour for suspected fetal macrosomia. *Cochrane Database of Systematic Reviews*,(2), Art. No. CD000938.

Johnson, S. R. J. (1995). The effect of epidural anesthesia on the length of labor. *The Journal of Family Practice, 40*(3), 244–247.

Kaufman, K. E., Bailit, J. L., & Grobman, W. (2002). Elective induction: An analysis of economic and health consequences [see comment]. *American Journal of Obstetrics and Gynecology, 187*(4), 858–863.

Kirby, R. S. (2004). Trends in labor induction in the United States: Is it true that what goes up must come down? *Birth 31*(2), 148–151.

Kotaska, A. J., Klein, M. C., & Liston, R. M. (2006). Epidural analgesia associated with low-dose oxytocin augmentation increases cesarean births: A critical look at the external validity of randomized trials. *American Journal of Obstetrics and Gynecology, 194*, 809–814.

Kozak, L. J., DeFrances, C. J., & Hall, M. J. (2006). National Hospital Discharge Survey: 2004 annual summary with detailed diagnosis and procedure data. *Vital and Health Statistics Series 13: Data from the National Health Survey,162*, 1–209.

Leighton, B. L., & Halpern, S. H. (2002). The effects of epidural analgesia on labor, maternal, and neonatal outcomes: A systematic review [see comment]. *American Journal of Obstetrics and Gynecology, 186*(5 Suppl. Nature), S69–S77.

Leonhardt, D. (2006, December 20). To-do-list: Wrap gifts. Have baby. *New York Times*. Retrieved from http://www.nytimes.com/2006/12/20/business/20leonhardt.html?_r=1&pagewanted=all.

Lieberman, E., Davidson, K., Lee-Parritz, A., & Shearer, E. (2005). Changes in fetal position during labor and their association with epidural analgesia. *Obstetrics and Gynecology, 105*, 974–982.

Lieberman, E., & O'Donoghue, C. (2002). Unintended effects of epidural analgesia during labor: A systematic review [see comment]. *American Journal of Obstetrics and Gynecology, 186*(5 Suppl. Nature), S31–S68.

Lydon-Rochelle, M., Holt, V. L., Easterling, T. R., & Martin, D. P. (2001). Risk of uterine rupture during labor among women with a prior cesarean delivery [see comment]. *New England Journal of Medicine, 345*(1), 3–8.

Lydon-Rochelle, M., Holt, V. L., Martin, D.P., Easterling, T. R. (2000). Asocition between method of delivery and maternal rehospitalization. *JAMA*, 283, 2411–6.

Lydon-Rochelle, M. T., Holt, V. L., Nelson, J. C., Cardenas, V., Gardella, C., Easterling, T. R., et al. (2005). Accuracy of reporting maternal in-hospital diagnoses and intrapartum procedures in Washington State linked birth records. *Paediatric and Perinatal Epidemiology, 19*(6), 460–471.

Martin, J. A., Hamilton, B. E., Sutton, P. D., Ventura, S. J., Menacker, F., & Kimeyer, S. (2006). Births: Final data for 2004. *National Vital Statistics Reports, 55*(1), 1–102.

Martin, J. A., Hamilton, B. E., Sutton, P. D., Ventura, S. J., Menacker, F., Kimeyer, S., et al. (2007). Births: Final data for 2005. *National Vital Statistics Reports, 56*(6), 1–104.

Maslow, A. S., & Sweeny, A. L. (2000). Elective induction of labor as a risk factor for cesarean delivery among low-risk women at term. *Obstetrics and Gynecology, 95*(6 Pt 1), 917–922.

Massachusetts Department of Health. (2008). MassCHIP. Retrieved from http://www.mass.gov/?pageID=eohhs2subtopic&L=4&L0=Home&L1=Researcher&L2=Community+Health+and+Safety&L3=MassCHIP&sid=Eeohhs2.

Maternal and Child Health Bureau. (2006). *Maternal and child health services Title V block grant program: Guidance and forms for the Title V application/annual report* (pp. 1–218). Rockville, MD: Health Resources and Services Administration, U.S. Department of Health and Human Services.

Menacker, F., Declercq, E. R., & MacDorman, M. (2006). Cesarean delivery: Background, trends and epidemiology. *Seminars in Perinatology, 30*, 235–241.

Mozurkewich, E. L., & Hutton, E. K. (2000). Elective repeat cesarean delivery versus trial of labor: A meta-analysis of the literature from 1989 to 1999. *American Journal of Obstetrics and Gynecology, 183*(5), 1187–1197.

National Center for Health Statistics, Division of Vital Statistics. (2006). 2003 Natality Data Set [CD-ROM Series 21]. 17.

National Institute of Child Health and Human Development Research Planning Group. (1997). Electronic fetal heart rate monitoring: Research guidelines for interpretation: National Institute of Child Health and Human Development Research Planning Workshop [see comment]. *American Journal of Obstetrics and Gynecology, 177*(6), 1385–1390; *Journal of Obstetrical Gynecological and Neonatal Nursing, 26*, 635–640.

National Institute of Child Health and Human Development, Maternal Fetal Medicine Units (MFMU) Network. (n.d.). Retrieved December 22, 2006, from http://www.bsc.gwu.edu/mfmu/about.cgi.

National Institutes of Health. (1981). *Consensus Development Conference on Cesarean Childbirth, September 1980.* Sponsored by the National Institute of Child Health and Human Development. NIH Pub. No. 82-2067. Bethesda, MD: NIH.

National Institutes of Health. (2006). NIH state-of-the-science conference statement on cesarean delivery on maternal request. *Obstetrics and Gynecology, 107*(6), 1386–1397.
Norwitz, E. R., Snegovskikh, V. V., & Caughey, A. B. (2007). Prolonged pregnancy: When should we intervene? *Clinical Obstetrics and Gynecology, 50*(2), 547–557.
O'Boyle, A. L., Davis, G. D., & Calhoun, B. C. (2002). Informed consent and birth: Protecting the pelvic floor and ourselves [see comment]. *American Journal of Obstetrics and Gynecology, 187*(4), 981–983.
Obst, T. E., Nauenberg, E., & Buck, G. M. (2001). Maternal health insurance coverage as a determinant of obstetrical anesthesia care. *Journal of Health Care for the Poor & Underserved, 12*(2), 177–191.
Padden, M. O. (1999). HELLP syndrome: Recognition and perinatal management. *American Family Physician,60*(3), 829–836, 839.
Parer, J. T., & King, T. (2000). Fetal heart rate monitoring: Is it salvageable? [see comment]. *American Journal of Obstetrics and Gynecology, 182*(4), 982–987.
Perlman, M. (2006). Patinet safety in obstetrics and gynecology: An agenda for the future. *Obstetrics and Gynecology, 108*, 1266–1271.
Rand, L., Robinson, J. N., Economy, K. E., & Norwitz, E. R. (2000). Post-term induction of labor revisited. *Obstetrics and Gynecology, 96*(5 Pt 1), 779–783.
Rayburn, W. F., & Zhang, J. (2002). Rising rates of labor induction, present concerns and future strategies. *Obstetrics and Gynecology, 100*(1), 164–167.
Sachs, B. P., Kobelin, C., Castro, M. A., & Frigoletto, F. (1999). The risks of lowering the cesarean-delivery rate [see comment]. *New England Journal of Medicine, 340*(1), 54–57.
Sanchez-Ramos, L., Olivier, F., Delke, I., & Kaunitz, A. M. (2003). Labor induction versus expectant management for postterm pregnancies: A systematic review with meta-analysis [see comment]. *Obstetrics and Gynecology, 101*(6), 1312–1318.
Seyb, S. T., Berka, R. J., Socol, M. L., & Dooley, S. L. (1999). Risk of cesarean delivery with elective induction of labor at term in nulliparous women. *Obstetrics and Gynecology, 94*(4), 600–607.
Shapiro-Mendoza, C. K., Tomashek, K. M., Kotelchuck, M., Barfield, W., Weiss, J., & Evans, S. (2006). Risk factors for neonatal morbidity and mortality among "healthy," late preterm newborns. *Seminars in Perinatology, 30*(2), 54–60.
Thacker, S. B. (1987). The efficacy of intrapartum electronic fetal monitoring. *American Journal of Obstetrics and Gynecology,156*(11), 24–30.
Thacker, S. B., & Banta, H. D. (1983). Benefits and risks of episiotomy: An interpretative review of the English language literature, 1860–1980. *Obstetrical & Gynecological Survey, 8*(6), 322–338.
Thacker, S. B., & Stroup, D. F. (2002). Methods and interpretation in systematic reviews: Commentary on two parallel reviews of epidural analgesia during labor [see comment]. *American Journal of Obstetrics and Gynecology, 186*(5 Suppl. Nature), S78–S80.
Vintzileos, A. M., Nochimson, D. J., Guzman, E. R., Knuppel, R. A., Lake, M., & Schifrin, B. S. (1995). Intrapartum electronic fetal heart rate monitoring versus intermittent auscultation: A meta-analysis [see comment]. *Obstetrics and Gynecology, 85*(1), 149–155.
Weber, A. M., & Meyn, L. (2002). Episiotomy use in the United States, 1979–1997. *Obstetrics and Gynecology, 100*(6), 1177–1182.
Whyte, H., Hannah, M., Saigal, S., Hannah, W. J., Hewson, S., Amankwah, K., et al. (2004). Outcomes of children at 2 years after planned cesarean birth versus planned vaginal birth for breech presentation at term: The International Randomized Term Breech Trial. *American Journal of Obstetrics and Gynecology, 191*, 804–871.
Winter, S., Autry, A., Boyle, C., & Yeargin-Allsopp, M. (2002). Trends in the prevalence of cerebral palsy in a population-based study. *Pediatrics, 110*(6), 1220–1225.
Yasmeen, S., Romano, P. S., Schembri, M. E., Keyzer, J. M., & Gilbert, W. M. (2006). Accuracy of obstetric diagnoses and procedures in hospital discharge data. *American Journal of Obstetrics and Gynecology, 194*(4), 992–1001.
Zhang, J., Yancey, M. K., & Henderson, C. E. (2002). U.S. national trends in labor induction, 1989–1998. *Journal of Reproductive Medicine, 47*(2), 120–124.

Chapter 17
Regionalized Perinatal Care: An Evidence-Based Intervention in Development

Lindsay A. Thompson and David C. Goodman

Regionalized Perinatal Care: What Is It and How Does It Work?

Regionalized perinatal care is an on-going effort to systematically organize and coordinate health care for expectant mothers and newborns. In the 1960s and 1970s, academic clinicians and public health experts pioneered regionalized perinatal services in response to new neonatal care technologies and expanding but still limited expertise (Committee on Perinatal Health, 1976; Graven, Howe, & Callon, 1976). By design, regionalization prioritizes the distribution of clinical care and technologies to women and infants according to risk. The spectrum of regionalized services ideally spans from preconception planning to infant health, although the concept is more commonly associated with the inpatient obstetric and neonatal services that women and infants receive in the hours and days surrounding birth. This chapter provides an evidence-based evaluation of regionalized perinatal care, its role and potential to reduce racial and ethnic disparities in perinatal outcomes, and illustrates the need for a broader implementation if its goals are to be fully realized.

Regionalization has now been established for the past five decades, with high quality research and influential policies in its support. While market forces have dominated the evolution of the U.S. health care system overall, regionalized perinatal care, in contrast, attempts to temper competition with coordination of care across regions and hospitals. One example of regionalization well supported by evidence (and discussed in detail later) is the demonstration that the level of care at the hospital of delivery, meaning the available concentrations of technical and clinical skills, directly influences infant survival. Successful regionalization has thus been credited, in part, for the continued decline in newborn mortality over the past 40 years (Richardson et al., 1998), and is considered an essential component of effective perinatal services.

Since its inception, the aims of regionalization have been to prioritize in a systematic fashion the needs of the highest risk maternal–infant pairs. To best serve mothers and infants, regionalization historically sought to organize health care systems using four strategies: (1) improvement of access to maternal and newborn services within communities; (2) early identification of high-risk pregnancies; (3) referral to the appropriate level of hospital care; and (4) the implementation of mechanisms to ensure compliance, continuity and comprehensiveness across all communities serving women and infants (American Academy of Pediatrics & The American College of Obstetricians and Gynecologists (AAP/ACOG), 2007). Its original conceptualization and subsequent revisions included improved access to family planning and well-woman care (as part of Strategy #1) leading to a greater likelihood of intended pregnancies and healthy pregnant women, and consequently

L.A. Thompson (✉)
Division of General Pediatrics, University of Florida, Gainesville, Fl 32608, USA
e-mail: lathompson@peds.ufl.edu

A. Handler et al. (eds.), *Reducing Racial/Ethnic Disparities in Reproductive and Perinatal Outcomes: The Evidence from Population-Based Interventions*, DOI 10.1007/978-1-4419-1499-6_17,

Table 17.1 Framework of regionalized inpatient perinatal care

Level	Type of care (other references)	Main criteria for designation of hospital level[a]	Modifications[b]
I	Basic (community hospitals)	Surveillance, care of all admitted obstetric and newborn patients – High-risk patient triage system – Detect unanticipated maternal-fetal risk	
II	Specialty (regional hospitals)	Level I + treatment/ stabilization of moderately ill and/or preterm newborns – Stabilize severely ill – newborns for transfer – Recovery from severe illness or anticipated rapid resolution of illness	– Level IIa = temporary assisted ventilation – Level IIb = continuous positive airway pressure or temporary mechanical ventilation
III	Sub-specialty (tertiary hospitals)	Level II + manage extreme prematurity, surgical, or critical illness – Evaluation of new technologies – Data collection and retrieval	– Levels IIIa–IIId reflect access to surgical specialties (e.g., cardiac care, ECMO) – Level IV designation as regional sub-specialty care center for coordination, outreach, and evaluation

[a]The American Academy of Pediatrics (AAP) has published a set of guidelines for designated levels, but there is no licensing body to oversee this designation. Currently, there is a list of units by levels compiled by the AAP; however, there are some states and hospitals that choose to self-designate

[b]These modifications reflect an inconsistency in the adoption of these designated levels. In some states, health care systems or hospitals choose these modifications to further clarify their level of care

lower maternal and newborn risk. Most care for women and infants would be provided within community settings. However, should complications arise, regionalization would have adequate mechanisms to identify (Strategy #2), and appropriately refer such patients to secondary and tertiary services (Strategy #3). Finally, health care systems would monitor their processes (Strategy #4) to insure that all strategies (#1, #2, #3) were upheld to the highest standard.

Out of these four strategies emerged the concept of 'levels' of care (Table 17.1). This central component of regionalization links technology to increasing clinical risk, an effort originally developed to ration scarce perinatal resources. The clinical application of regionalization has at its foundation a hierarchy of levels of antepartum, peripartum and neonatal hospital care (levels I–III) to match the increasingly complex health needs of delivering mothers and infants. A level I hospital provides 'basic care' for routine deliveries, a level II offers specialty care, such as short term stabilization of infants, and level III offers subspecialty services such as high risk obstetrics and complex neonatal intensive care. An effective regionalization program requires systematic efforts between hospitals of different levels to coordinate the transportation of women with threatening high-risk deliveries or high-risk infants, as well as the 'back transfer' of convalescing infants once stable (American Academy of Pediatrics, 2004). The levels of care, and related measures such as patient volume per hospital unit and rates of maternal or high-risk infant transfers to higher levels of care, serve as markers of regionalization in most studies of perinatal services. Regionalization supports a hierarchy of care, inter-hospital agreements, and shared accountability across each regional health care system. In practice, this effort has increased the coordination of inpatient care while the broader health care goals to improve health before and during pregnancy for *all* expectant mothers, and thereby reduce racial/ethnic disparities in care and outcomes, have never been fully realized.

Historical Context of Regionalized Perinatal Services

Perinatal regionalization became a formal policy initiative in 1976 with the publication by the March of Dimes' *Toward Improving the Outcome of Pregnancy* (TIOP, referred to as TIOP I), written by the Committee on Perinatal Health (1976). This consensus statement was built on the best available research at the time, such as the Wisconsin analysis that showed infant survival to be enhanced with better access to perinatal health services (Graven et al., 1976). With support from the Robert Wood Johnson Foundation, and participation by the American Medical Association, the American College of Obstetrics and Gynecology, the American Academy of Pediatrics, and the American Academy of Family Physicians, TIOP I and the concept of regionalization was fully endorsed and promulgated over the following decades (Table 17.2). Implementation documents, such as *Guidelines for Perinatal Care*, now in its sixth edition (AAP/ACOG, 2007), provide visibility and professional guidance to regionalization. In addition to these publications, the Healthy People initiatives, funded and championed by the U.S. Department of Health and Human Services, established outcomes that incorporate key aspects of regionalization, such as the goal that 90% of very low birth weight births (<1,500 g) occur at tertiary care centers (level III) by 2010 (Koontz, 1984; U.S. Department of Health and Human Services, n.d.).

In 1993, a second March of Dimes Committee on Perinatal Care met and published *Towards Improving the Outcomes of Pregnancy: The 90s and Beyond* (TIOP II) (The Committee on Perinatal Health, 1993). The committee recognized the role of regionalization in improved newborn survival but noted that the broader concept of coordinated care from preconception to postnatal care had yet to develop. While regionalization has been established as the standard of care in the peripartum period in general, and for neonatal intensive care specifically, TIOP II recognized that an opportunity had been missed to address the underlying causes that prompt the need for higher levels of perinatal care, such as unintended pregnancy and inadequate access to prenatal care. As such, the committee urged a fuller development of preventive services along the full spectrum of reproductive care including regionalized services for preconception and inter-conception care as well as for ambulatory prenatal care. An expansion of regionalization efforts beyond the narrowly defined inpatient arena could generally lower the risk of mothers and infants, and reduce the need for tertiary care services.

Over a decade later, the TIOP II recommendation to broaden the scope of regionalization has been largely ignored, while regionalization of inpatient obstetrical and neonatal care has even been partly overturned. In the 1970s and 1980s, the limited availability of neonatal intensive care units and professionals demanded coordination of perinatal medical care by public entities, such as state maternal and child health programs. However, this era was quickly followed by a rapid expansion of maternal and newborn services, largely due to the financial incentives to hospitals of providing the full range of perinatal services. Further, the growth of NICUs and in the number of physicians trained as neonatologists (Pollack, Ratner, & Lund, 1998) mostly occurred in non-academic hospitals. At the same time, hospitals with level I and level II units upgraded themselves to higher levels of care (Yeast, Poskin, Stockbauer, & Shaffer, 1998). Competition began to gain the upper hand at the expense of regionalization.

Although regionalization was designed to match pregnant women with the needed level of care, the increasing numbers of high-level care centers have lead to fewer patients in any given NICU (Goodman, Fisher, Little, Stukel, & Chang, 2001; Goodman et al., 2002; Phibbs, Bronstein, Buxton, & Phibbs, 1996; Powell, Holt, Hickok, Easterling, & Connell, 1995). While access to care has improved, it has done so at the expense of patient volume, a factor that affects quality of care. Even though regionalized perinatal care remains a widely shared goal for perinatal health services, its implementation has varied markedly over time and across communities.

Table 17.2 Policy papers and the role of regionalization

Spectrum of perinatal services						
Policy documents	Major focus	Family planning services	Normal pregnancy care	High risk pregnancy care	Neonatal intensive care	Well infant and child care
Toward Improving the Outcome of Pregnancy I (TIOP I: 1976)	Introduces regionalization as a means for high quality care to ration scarce resources through hospital levels of care	Not mentioned	Mentions need for standard of care for 'uncomplicated maternity cases'	Introduces the need for risk-identification and maternal transfer to higher levels of care	Defines hospital levels • I = basic • II = specialty • III = subspecialty	Mentions need for immediate post-natal care
Toward Improving the Outcome of Pregnancy II (TIOP II: 1993)	Expands goals of TIOP I: ensure health for every woman and baby; includes regional accountability to increase effectiveness	Increase health promotion, education, reproductive awareness (decrease unintended pregnancy), preconception and interconception care	Universal ambulatory prenatal care	Increase ability to identify and transfer high risk mothers	Reinforces hospital levels; increase accountability, interhospital care	Extend perinatal services through the first year of life
Guidelines for Perinatal Care (6th ed., 2007)	Comprehensive guide with link to Guidelines for Women's Health Care[a]	Not mentioned	Mentions need for standard of care for 'uncomplicated maternity cases'	Risk identification along with Guidelines for Women's Health Care[a]	Reinforces hospital levels; increase accountability, interhospital care	Mentions need for immediate post-natal care
Healthy People 2010	Sets attainable public health goals (does not mention role of regionalization, only uses components of it)	Reduce unintended pregnancy	Increase early and adequate prenatal care	Increase the VLBW deliveries at level III care	Aim to reduce LBW, VLBW preterm births	Increase breast feeding rates

[a] A companion guide for Obstetrician/Gynecologists, called *Guidelines for Women's Health Care* has been available since 1989. However, it does not mention regionalization of services

Methods for Reviewing the Evidence for the Effectiveness of Regionalized Perinatal Care

There are numerous studies that grapple with the organization of perinatal care and its effects on neonatal outcomes. This chapter interprets the available literature on the effectiveness of regionalized perinatal care and its potential to reduce racial/ethnic disparities in perinatal outcomes. During the months of 9/06–12/07, we used Pubmed and multiple relevant MeSH subheadings to conduct a search of the relevant literature. We cite those articles that: (a) were noted by their frequent citations to be sentinel articles at the time of publication; (b) caused a shift in focus of policy or research; (c) addressed racial or ethnic disparities in perinatal health services; and/or (d) tested a critical aspect of regionalization in a new or effective way. We specifically included studies that share common outcomes, such as rates of neonatal mortality, to best compare the relative strengths of the evidence. Articles that use other outcomes may be mentioned but will not be critiqued within the tables of evidence presented. We include a comprehensive analysis of 'white papers' or articles issued after 1985 through December 2007 (we also include one paper from 1975).

Judging the Success of Perinatal Regionalization

Main Outcome Measures in Newborns that Reflect the Effectiveness of Regionalized Perinatal Care

Evaluation of regionalization has typically used vital statistics to measure population-based outcomes of newborns. The most salient outcome of interest is the *neonatal mortality rate* (NMR, the rate of death within the first 28 days of life divided by the total number of live births). Since this outcome reflects care given during the peripartum and neonatal period, it most closely estimates the effects of the hospital services that regionalization aims to organize. Another outcome of interest is *birth-weight specific mortality* (death within a given time frame grouped by birth weight categories) because it better compares outcomes within similar levels of risk. Some studies limit their population of interest to those infants most likely to need NICU care, and hence restrict the study population to infants with *very low birth weight* (VLBW, less than 1,500 g), or *extremely low birth weight* (ELBW, less than 1,000 g). Other outcomes of interest include *low birth weight rates* (LBW, a birth weight less than 2,500 g) and *infant mortality rates* (IMR, the rate of death within the first year of life divided by the total number of live births). Low birth weight and infant mortality rates, however, cumulatively reflect all risks, including the complex social, economic and biological factors along with those of ante-partum or pediatric care, making inferences about regionalization of neonatal intensive care more difficult.

Markers that Represent Aspects of Regionalization as an Intervention in Perinatal Care

Regionalization as an epidemiologic exposure is difficult to define or measure precisely. Unlike clinical trials, where interventions and responses can be precisely measured, the exposure to regionalized perinatal care is often ill-defined (e.g., what are the boundaries to a health care 'region?' What comprehensive data on 'regions' are routinely collected?). Even if data exist that would allow

for regional comparisons, hospital systems may be hesitant to share information for research, for confidentiality reasons, or because of competition with other hospitals. Further, there have been changes in the newborns thought to require higher levels of care, making on-going population comparisons difficult. In addition, complex statistical methods are often needed for evaluating population outcomes since regional characteristics must be combined with individual level data in multi-level models, techniques that were not fully developed in early study periods.

Nonetheless, there has been a large body of literature evaluating the effectiveness of regionalized perinatal services. While blinded, randomized trials have been impractical, and studies using historical controls are problematic given the continual advances in newborn care, researchers have been creative in finding ways to test the effects of regionalization. Methods vary from those that compare highly organized regions to 'less regionalized' areas to those where investigators test 'components' of regionalization. For example, many studies have compared infant survival by level of care in NICUs, and more recently the volume of NICU patients. The majority of these studies specifically measure newborn outcomes as evidence of the effectiveness of regionalized perinatal services. More recent studies, however, have built on this work, and study processes of care, such as rates of maternal and infant transfers in high-risk populations, without linkage to infant outcomes. Through these studies, regionalized care has come to mean almost exclusively inpatient peripartum care. This chapter summarizes these different approaches to evaluating regionalization (Tables 17.3 and 17.5) and its role in the reduction of racial and ethnic disparities in perinatal outcomes (Table 17.7).

Review of the Evidence for the Effectiveness of Regionalization in Improving Newborn Outcomes

Regionalization aims to match expectant women with needed services, and has focused on the referral of high-risk pregnant women and infants to hospitals with higher levels of neonatal intensive care. In this review, we examine several types of studies which focus on the effect of regionalization on adverse neonatal or perinatal outcomes. These studies are presented in Tables 17.3 and 17.4 in chronological order. In general, the chronological order matches the sequence of the sub-sections below.

Population Studies that Examine the Influence of Regionalization on Neonatal/Perinatal Outcomes

(The studies discussed in this section are found mostly toward the top of Tables 17.3 and 17.4.) In the 1970s the Robert Wood Johnson Foundation funded a demonstration project to "promote coordinated systems of care for entire geographical regions" (McCormick, Shapiro, & Starfield, 1985, p. 799). McCormick et al. evaluated this 'national demonstration program,' using a non-randomized interventional design to compare neonatal mortality rates (NMR) in the funded areas to those without explicit funds for regionalization. This sentinel study did *not*, in fact, find differences in neonatal mortality outcomes in funded compared to unfunded areas because coordinated care and services became increasingly standard within all regions after the publication of TIOP I in 1976. McCormick et al. did observe important changes in the pattern of service delivery as more LBW infants were born in tertiary hospitals accompanied by a decrease in NMR. Also published in 1985, Gortmaker, Sobol, Clark, Walker, and Geronimus assembled data from four states and found increased survival

Table 17.3 Health status outcome: effect of regionalization on neonatal or perinatal mortality

Author (year) Study design	Study type	Description of intervention/ outcome	Populations studied and sample size	Address disparities (yes/no)[a]	Key findings related to intervention effectiveness	Caveats/biases	Findings support the intervention? Yes/no For which populations?
Gortmaker et al. (1985) RC[d]	Published article	NMR[b] of VLBW[c] by hospital level (uses level III hospitals vs. 'other urban' vs. 'rural')	4 states' VLBW[c] population, 1978–1979, n = 9,021	Stratifies black and white; does not address way to improve disparities; Black infants have more access to level III	Improved VLBW[c] survival at level III, especially within first 12 h; Black infants without significant difference	Small sample size of black infants – may explain why black infants do not show significant difference in survival at higher hospital levels	Yes: supports higher hospital levels and maternal transports for VLBW[c] infants
McCormick et al. (1985) NR[f]	Published article	LBW[e], NMR[b] in n = 8 sites funded to promote regionalization vs. controls (n = 1–18 per region)	Singleton live births per region, 1970–1971; 1974–1975; 1978–1979, n = 2.4 million	No	NMR[b] lower in both groups over time, especially decreased in LBW[e] infants; assumes increased regionalization in both groups led to less NMR[b]	Assumes overall decline due to regionalization; no accounting for other causes of decreased NMR[b] over time	Yes: supports regionalization despite lack of findings in intervention group; cites strong overall change towards regionalized systems as cause of decline in NMR[b] in all LBW[e] births
Hein and Lathrop (1986) RC[d]	Published article	NMR[b] by hospital level (uses 3 levels)	Review of death certificates to identify: preventable (amenable to NICU[g] care) vs. non-preventable in Iowa, comparing 1982–1983 (357 reviewed cases) to 1978–1979 (459 reviewed cases)	No	Fewer deaths occurred in the 1983 cohort, more occurred at higher hospital levels, and the causes of death shifted from preventable to non-preventable	No accounting for other causes of decreased NMR[b] over time	Yes: supports higher hospital levels for high-risk infants

(continued)

Table 17.3 (continued)

Author (year) Study design	Study type	Description of intervention/ outcome	Populations studied and sample size	Address disparities (yes/no)[a]	Key findings related to intervention effectiveness	Caveats/biases	Findings support the intervention? Yes/no For which populations?
Lubchenco et al. (1989) RC[d]	Published article	VLBW[c] NMR[b] and morbidity when born in level I vs. level III hospitals	Fetal mortality and NMR[b] in Denver, CO, 1975–1978, n = 471	No	Fetal mortality was reduced at higher level of care, but NMR[b] was not	Small sample size	Yes, cautiously: claims enhanced regionalization (and transport services) might be reason for equivalent NMRs[b] in both groups
Langkamp et al. (1990) RC[d]	Published article	NICU[g] admissions and mortality among black and white newborns	9 counties in New York State, 1983–1984, n = 35,845 births	Yes	Higher NICU admission rates (OR 2.6), but higher mortality rates for blacks (OR 2.15)	Limited covariables; uses p value <1 in some cases	Yes, indirectly: causes of disparities in mortality for black infants predates NICU admissions
Mayfield et al. (1990) RC[d], R[i]	Published article	PMR[h] by hospital volume of births and hospital levels (uses 3 levels)	White singleton births in Washington State, n = 226,164	No (excluded non-white infants)	LBW[e] infants risk of mortality twofold better if born at level III vs. level I, II; volume of births less important for PMR[h]	Restricted population	Yes: supports LBW[e] births at level III hospitals; volume of births less influential
LeFevre et al. (1992) RC[d]	Published article	NMR[b] by hospital level (uses 5 levels including measures of patient volume)	Singleton live births in Missouri, n = 354,441	Stratifies by African-American, and white; does not address way to improve disparities	White (and some black) infants <2.25 kg born at level IA (low volume) hospitals have higher NMR[b] than at level III; no difference for higher birth weights (many BW-specific NMRs[b], CIs given)	Small sample size of black infants – may explain why black infants do not show significant difference in survival at higher hospital levels	Yes: supports LBW[e] births in level III hospitals and normal birth weight deliveries at low level hospitals

Howell and Vert (1993) RC[d]	Published article	Access to NICU services and PMR[h]	n = 133,293 (weighted) births in Michigan; 30,779 births in Lorraine, France	No: unable to obtain in Lorraine, France	Higher number of births in hospitals with NICUs in Michigan (31.1% vs. 25.0%) with lower PMR[h] (11.3 per 1,000 vs. 13.7)	Restricted population	Yes: supports all births in hospitals with NICUs and endorses policies of careful maternal screening
Bronstein et al. (1995) RC[d]	Published article	Compared 4 infant populations: VLBW[c] births in NICU hospitals; maternal transfers for impending VLBW[c] births; VLBW[c] transfers; and non-transferred VLBW[c]	VLBW infants in Alabama, 1988–1990, n = 2,596	Non white race entered into models as maternal characteristic	Maternal characteristics not considered part of a risk assessment, including race and insurance status, influenced transfer rates of VLBW infants; non-white women with early prenatal care *more* likely to give birth at hospital with NICU	Attempts to account for interactions of race with other sociodemographic characteristics; unable to generalize to other racial and ethnic groups	Yes, indirectly: as it targets barriers to perinatal regionalization
Powell et al. (1995) RC[d]	Published article	NMR[b] in LBW[e] by hospital level (uses 3 levels)	LBW[e] infants in Washington State, 1980–1991, n = 43,228	Maternal race recorded ("White, Black, other"); relation to outcomes not clear	Decline of LBW[e] births at level III; increased NMR[b] at non-level III [e.g., 1–1.5 kg: OR 2.1 (CI 1.3, 3.3)]	No risk adjustment	Yes: supports LBW[e] deliveries at high level, concerned over trend of fewer LBW[e] births at high levels at end of cohort

(continued)

Table 17.3 (continued)

Author (year) Study design	Study type	Description of intervention/ outcome	Populations studied and sample size	Address disparities (yes/no)[a]	Key findings related to intervention effectiveness	Caveats/biases	Findings support the intervention? Yes/no For which populations?
Kirby (1996) RC[d]	Published article	NMR[b] in LBW[e] by hospital level in an area without extensive regionalization (uses 3 levels)	Births in Arkansas, 1985–1989, n = 175,225	Race/ethnicity in risk adjustment; does not address way to improve disparities	Adjusted OR of NMR at a level III hospital 0.43 (0.29, 0.64)		Yes: supports LBW[e] deliveries at high level, even in areas without extensive regionalization
Phibbs et al. (1996) RC[d]	Published article	NMR[b] by hospital level and volume of (likely) NICU[g] admissions	Singleton, likely NICU[g] admissions 1990, n = 53,229	Race/ethnicity in risk adjustment; does not address way to improve disparities	Risk-adjusted mortality lower (OR 0.62, CI 0.47, 0.82) at level III, high volume (≥15 patients/day); little difference in cost		Yes: supports high-risk infant births at high level and high NICU[g] volume
Rosenblatt et al. (1996) RC[d]	Published article	NMR[b] by hospital level (uses 3 levels) in 2 countries and policies related to regionalization	n = 28 hospitals in Wales, n = 80 hospitals in Washington State; number of births not stated	No	Less regionalization in Wales (with more hospitals having more technology) vs. Washington state, but equal NMR[b]	2 diverse health care systems	Yes, indirectly: cites duplication of services in Wales and increased cost without improved NMR[b]
Dooley et al. (1997) RC[d]	Published article	NMR in non-tertiary hospitals to assess the relative contribution of maternal vs. hospital characteristics	n = 97 hospitals with >190,000 annual births, in Illinois, 1991–1993	Race/ethnicity in risk adjustment; does not address way to improve disparities	Maternal risk explained 73% of hospital variation in rates of NMR[b]; rates of transport and inborn VLBW[c] deliveries also contributed		Yes: the quality of regionalization needs to be measured against maternal risk differentials in areas

Horbar et al. (1997) RC[d]	Published article	Observed and standardized NMRs[b] at a range of NICUs to see if patient or hospital characteristics predicted mortality	VLBW[c] infants in n = 62 Vermont Oxford Network NICUs, n = 7,672	Race/ethnicity in risk adjustment; does not address way to improve disparities	Patient volume and presence of a pediatric residency do *not* predict increased mortality in multivariate analyses, but there are variations in mortality not due to patient severity	Only looks at highest levels of care (NICUs) and VLBW[c] infants	Yes, indirectly: effectiveness of care varies among NICUs
Menard et al. (1998) RC[d]	Published article	NMR[b] by hospital level in VLBW[c] births	VLBW[c] infants in South Carolina, 1993–1995, n = 2,375	Risk adjustment and stratification: African Americans higher mortality vs. whites within hospital levels	VLBW[c] infants survive better at level III hospitals: OR 1.66 (1.45, 1.90) vs. level I or II, adjusted for birth weight and race	Small sample; transfer biases (e.g., highest risk mothers not able to be transported = cause of lower mortality at higher levels)	Yes: supports VLBW[c] deliveries at level III, including black infants
Richardson et al. (1998) RC[d]	Published article	NMR[b] in VLBW[c] infants in two time cohorts	1989–1990 and 1994–1995, n = 739 infants	Race in risk adjustment; does not address way to improve disparities	VLBW[c] had improved risk-adjusted survival in the recent cohort; (OR 0.52, CI 0.29–0.96); estimate 2/3 of improvement due to better NICU care; 1/3 improvement from better obstetric care	Hard to generalize as it only took into account 2 NICUs	Yes: newborn care is more effective in the era of enhanced NICU technology that is centered in regionalized perinatal care

(continued)

Table 17.3 (continued)

Author (year) Study design	Study type	Description of intervention/ outcome	Populations studied and sample size	Address disparities (yes/no)[a]	Key findings related to intervention effectiveness	Caveats/biases	Findings support the intervention? Yes/no For which populations?
Yeast et al. (1998) RC[d]	Published article	NMR[b] by self-designated hospital level (uses 3 levels) and patient volume in 2 time cohorts	Missouri births, n = 381,866 (1982–1986) n = 390,741 (1990–1994)	Race in risk adjustment; does not address way to improve disparities	Shift to self-designated level II, III not related to acuity or census; VLBW[c] higher NMR[b] (OR 2.28, CI 1.33, 3.89) at level I		Yes: supports VLBW[c] deliveries at level III; warns that self-designation of hospital level does not equal effective regionalization
Moster et al. (1999) RC[d]	Published article	NMR[b] by hospital volume	Infants >2.5 kg in Norway, 1972–1995, n = 1,254,284	No	Annual deliveries 2,001–3,000 g had lowest NMR[b] (<2.0 kg = OR 2.1, CI 1.6, 2.8; >3.0 kg = OR 1.7, CI 1.4, 2.0)	Crosses time periods of critical advances in technology (e.g., surfactant)	Yes: supports volume as a modifying influence to levels of care
Chien et al. (2001) RC[d]	Published article	NMR and morbidity by 'outborn vs. inborn' births in NICUs[g]	Singletons ≤32 weeks GA in 17 Canadian NICUs[g] (≈level III hospitals), n = 3,769	No	Outborn infants adjusted NMR[b] OR = 1.7 (CI 1.2, 2.5)		Yes: supports transfer of high risk mothers (in utero fetus ≤32 weeks GA) to NICUs[g]
Cifuentes et al. (2002) RC[d]	Published article	NMR[b] by hospital level (uses 4 levels) and volume of deliveries (≥15 per day)	Singletons, BW <2 kg in California, 1992–1993, n = 16,732	Race/ethnicity in risk adjustment; does not address way to improve disparities	Birth in non-NICU[g] = adjusted OR 2.38 (1.81, 3.13); birth in intermediate NICU[g] = 1.92 (1.44, 2.54); low volume level III 1.42 (1.14, 1.76)		Yes: supports transfer of high risk moms and delivery of infants <2 kg at high level and high volume hospitals

Goodman et al. (2002) RC[d]	Published article	NMR[b] rates by number of neonatologists and number NICU beds per births	Births in 1995, n = 3.8 million	Race/ethnicity in risk adjustment; does not address way to improve disparities	NMR lower if an area had 4.3 neonatologists per 10,000 births vs. 2.7 neonatologists per 10,000 births. However, further increases (beyond 4.3) in number of neonatologists were not associated with a greater decrease in NMR	Uses standard methods to define 'area' but regional perinatal systems are not well defined	Yes, cautiously: model of scarce resources upon which RPS was grounded is no longer true
Heller et al. (2002) RC[d]	Published article	NMR[b] by hospital volume	95% of all births in Hesse, Germany, n = 582,655	No	OR = 3.5 (CI 2.6, 4.6) NMR[b] at low volume (<500 births) hospitals compared to high volume (>1,500 births)		Yes: supports deliveries at high volume hospitals, even for low risk (>2.5 kg) births
Rogowski et al. (2004) RC[d]	Published article	NMR[b] in VLBW[c] infants: estimates reduction in NMR if all infants born at best performing hospitals	Hospitals participating in Vermont Oxford Network, 1994–2000; approx. 40% of all US NICUs and approx. 50% of all US VLBW births	Yes: shows that black infants have as much variation as white infants in risk adjusted NMRs across hospitals; Black infants should benefit at least as much as white infants from improved quality	Threefold variation in NMR across hospitals; potentially large impact from improved quality across NICU hospitals in Vermont Oxford Network	Only uses NICU hospitals; excludes hospitals without NICUs	Yes: studies potentially could link RPS indicators, such as patient referrals, to quality indicators

(continued)

Table 17.3 (continued)

Author (year) Study design	Study type	Description of intervention/ outcome	Populations studied and sample size	Address disparities (yes/no)[a]	Key findings related to intervention effectiveness	Caveats/biases	Findings support the intervention? Yes/no For which populations?
Merlo et al. (2005) RC[d]	Published article	NMR[b] by hospital level (uses 4 levels)	99.5% all Swedish births, 1990–1995, n = 691,742	No	Low risk (birth weight >2.5 kg) and high risk (birth weight ≤2.5 kg) deliveries have lower NMR[b] with higher hospital level		Yes: supports deliveries at high level hospitals, even for low risk (>2.5 kg) births (policy may not be cost-effective)
Lui et al. (2006) RC[d]	Published article	Statewide coordinated changes in perinatal services. NMR[b] and level of hospital of birth in two time cohorts	23–28 week gestational age infants in New South Wales Australia, n = 1,778 (1992–1995) and n = 3,099 (1997–2002)	No	Fewer outborn infants in recent cohort (12.0% vs. 9.3%) and decreased outborn mortality (39.4% vs. 25.1%)	Australian study – not clear how to apply the US market pressures to their outcomes	Yes: coordinated telephone advice to optimize maternal transfers and centralization of the neonatal retrieval system showed marked improvements
Phibbs et al. (2007) RC[d]	Published article	NMR[b], volume, and level of hospital in VLBW[c]	VLBW[c] infants in California, 1991–2000, N = 48,237	Race/ethnicity in risk adjustment; does not address way to improve disparities	Mortality rates of VLBW[c] vary according to NICU volume *and* level of care	Unclear how findings will be implemented since market forces are expanding number of NICUs, not contracting them	Yes: supports VLBW[c] deliveries at high level, high volume units

[a] Articles that have embedded questions on racial/ethnic disparities and regionalization are noted here.
[b] *NMR* Neonatal mortality rate: the rate of death within the first 28 days of life divided by the total number of live births
[c] *VLBW* Very low birth weight: birth weight less than 1,500 g
[d] *RC* Retrospective cohort design
[e] *LBW* Low birth weight: birth weight less than 2,500 g
[f] *NR* non-randomized trial design
[g] *NICU* Neonatal intensive care unit
[h] *PMR* Perinatal mortality rate: the rate of fetal death plus early neonatal deaths (up to 1 week of life) divided by the total number of live births
[i] *R* review

Table 17.4 Quality rating of studies associated with regionalization: effect of regionalization on neonatal or perinatal mortality

Author (year)	Reporting	External validity	Internal validity – bias	Internal validity – confounding	Power	Total quality score	Suitability of study to assess effectiveness: greatest, moderate, least
Berger et al. (1975)	5	0	5	3	0	13	Moderate
Gortmaker et al. (1985)	8	3	4	4	0	19	Moderate
McCormick et al. (1985)	6	2	3	3	0	14	Moderate
Hein and Lathrop (1986)	5	3	3	4	0	15	Moderate
Lubchenco et al. (1989)	7	3	5	3	0	18	Moderate
Langkamp et al. (1990)	9	2	5	4	0	20	Moderate
Mayfield et al. (1990)	10	2	3	4	0	19	Greatest
LeFevre et al. (1992)	7	2	4	3	0	17	Greatest
Howell and Vert (1993)	8	2	4	3	0	17	Moderate
Bronstein et al. (1995)	8	3	6	4	0	21	Greatest
Powell et al. (1995)	9	3	7	5	0	24	Greatest
Kirby (1996)	8	3	6	4	0	21	Greatest
Phibbs et al. (1996)	11	3	7	5	0	26	Greatest
Rosenblatt et al. (1996)	8	3	5	3	0	19	Moderate
Dooley et al. (1997)	8	3	6	4	0	21	Greatest
Horbar et al. (1997)	10	3	7	4	0	24	Greatest
Menard et al. (1998)	10	2	5	4	0	21	Greatest
Richardson et al. (1998)	10	2	7	4	0	23	Greatest
Yeast et al. (1998)	7	3	6	3	0	19	Moderate
Moster et al. (1999)	8	3	7	3	0	21	Greatest
Chien et al. (2001)	11	3	7	5	0	26	Greatest
Cifuentes et al. (2002)	11	3	7	5	0	26	Greatest
Goodman et al. (2002)	11	4	7	4	0	26	Greatest
Heller et al. (2002)	9	2	6	5	0	22	Greatest
Rogowski et al. (2004)	8	1	7	4	0	20	Moderate
Merlo et al. (2005)	10	3	7	4	0	24	Greatest
Lui et al. (2006)	10	4	6	5	0	25	Greatest
Phibbs et al. (2007)	11	4	7	5	0	27	Greatest

Table 17.5 Health status outcome: effect of regionalization on location of VLBW births and the rates of maternal transfers

Author (year) Study design	Study type	Description of intervention/outcome	Populations studied and sample size	Address disparities (yes/no)[a]	Key findings related to intervention effectiveness	Caveats/biases	Findings support the intervention? Yes/no For which populations?
Handler et al. (1991) CS[b]	Published article	Maternal transfers to a pubic level III hospital	Maternal transfers in Chicago, IL, 1987, 1988, n = 970	No	Regionalized perinatal services are vulnerable to economic instability since private and public hospitals have different priorities	Case study design does not use statistical techniques	Yes: cautions that incentives outside of perinatal risk will undermine the goals of regionalization
Gould et al. (1999) RC[e]	Published article	VLBW[c] births by hospital level (uses 2 levels – level I and non-level I)	VLBW[c] births in California, 9 regions of perinatal care, 1989–1993, n = 24,094	Hispanic women have more VLBW[c] deliveries at level I; does not address way to improve disparities	Sevenfold variation in % VLBW[c] deliveries in level I hospitals across regions; race and ethnicity affect variation	Does not examine NMR[d]	Yes, supports more VLBW[c] at higher level hospitals, notes significant modifiers like race and ethnicity to outcomes of health services
Mehta et al. (2000) RC[e]	Published article	Site of delivery for VLBW[c] and infants with major congenital anomalies	Infants born 1990–1995 in Ohio, N = 9,032	No	59.8% of VLBW and 36.1% with malformations born in level III; variations over time and regions	Does not examine NMR[d]	Yes: the decrease in VLBW level III births is an indicator of failing RPS
Wall et al. (2004) RC[e]	Published article	Hospital factors and rate of non-transfer of ≤1.25 kg to level III hospitals	500–1,249 g births in Illinois, n = 2,904	African-American OR = 0.81 (CI 0.68, 0.88) for non-transfer; Hispanic OR = 0.82 (CI 0.65, 0.97) in adjusted analyses	Level II+, high Medicaid revenues, high HMO revenues and status as teaching hospital lead to decreased transfer rates	Unclear why African-American and Hispanic more likely to be transferred	Yes: findings challenge goals of regionalization; identified non-clinical determinants of transfers

Sinkin et al. (2005) RC[e]	Published article	Effect of managed care on rates of maternal transport	Upstate NY, n = 106,466	Race 'other than white' had OR 0.52 (CI 0.42, 0.63) for transfer	Adjusted OR of transfer = 0.72 (0.62, 0.82) for Medicaid managed care vs. FFS	Unclear why non-whites are less likely to be transferred	Yes: documents influence of payment on transport systems, a key aspect of regionalization

[a] Articles that have embedded questions on racial/ethnic disparities and regionalization are noted here
[b] *CS* Comparitive Study
[c] *VLBW* Very low birth weight: birth weight less than 1,500 g
[d] *NMR* Neonatal mortality rate: the rate of death within the first 28 days of life divided by the total number of live births
[e] *RC* Retrospective cohort

for high-risk infants, especially within the first 12 hours of birth, in hospitals with level III nurseries. Hein and Lathrop (1986) reviewed death certificates in Iowa and inferred that improved regionalization contributed to the decrease in NMR in Iowa in the 1970s, especially since the majority of deaths occurred at higher level hospitals, and that the causes of death shifted to causes less likely to be helped by perinatal care.

In addition to these early studies, others have illustrated the relative value of regionalized systems by comparing NMRs in areas with varying implementation of regionalization. For example, Howell and Vert (1993) compared perinatal outcomes in the State of Michigan to those in Lorraine, France, and found that an increase in high-risk births in Michigan NICUs coincided with a decrease in the perinatal mortality rate (PMR). In contrast, Lorraine, France, had fewer infants born within hospitals with NICUs and had a higher PMR. Rosenblatt et al. (1996) compared outcomes in Wales, UK, where hospitals were less likely to have established referral patterns for transportation of the smallest infants to level III centers, to Washington State, where regionalization and levels of care had been established for over a decade. Despite access to national health insurance in the UK, which might imply better access and improved outcomes, the neonatal mortality rates were equivalent. The authors speculate that there may have been unnecessary duplication of services in the UK, whereas regionalization in Washington State was a more efficient model of care. These studies of the systematic organization of health services, while initially of only moderate quality (Table 17.4), uniformly supported regionalization as a means for disbursing new, scarce, and expensive neonatal intensive-care technologies to those patients with the highest need, which in turn predicted improved newborn outcomes. Indeed, a study by Richardson et al. (1998) estimated that approximately two-thirds of the improvement in neonatal mortality was likely due to advances in all aspects of NICU care (including advances in technology as well as in regionalization), while only one-third of the improvement might be attributed to obstetric care.

Studies of Hospital Levels of Care and Effects on Neonatal Outcomes

Because of the challenges in measuring the overall effect of regionalized systems, most studies beyond the initial studies noted above, focus on measurable components of regionalization, specifically inpatient characteristics of the hospitals and associated newborn outcomes. (These studies are in Tables 17.3 and 17.4, clustered in the middle of the citations, which are listed chronologically.) Since TIOP I and subsequent documents concentrate on defining levels of care, and inpatient newborn datasets are comparatively easy to obtain, studies comparing levels of care and infant outcomes have been central to research on regionalization. LeFevre, Sanner, Anderson, and Tsutakawa (1992) and Mayfield, Rosenblatt, Baldwin, Chu, and Logerfo (1990) independently noted higher NMR among high-risk infants (defined by birth weight) born at lower level neonatal centers, although Lubchenco et al. (1989) found conflicting results. LeFevre et al. also established the relative safety of low-risk (normal birth weight) infants' birth at level I centers, noting that the phenomenon of closing lower level nurseries was unwarranted, and could decrease access to obstetric services overall. Powell et al. (1995) also showed that increasing rates of LBW births at hospitals with higher levels of care coincided with decreased NMR. They warned, however, that at the end of their study period (1991) there was a trend toward fewer LBW births at level III centers, forecasting future studies where de-regionalization became more common. Kirby (1996) found births in Arkansas hospitals with higher levels of care to be associated with better birth-weight specific survival, especially for the smallest infants, despite Arkansas' lack of a formal structure of regionalization at that time. Similar findings were reported by Menard, Liu, Holgren, and Sappenfield (1998), and Chien et al. (2001). Even recently, studies continue on the benefits of higher levels of care. Merlo et al. (2005) confirmed higher levels of NICU care in Sweden to be associated with improved survival for

high-risk deliveries. They also suggest that low-risk deliveries may have increased survival in hospitals with higher levels of neonatal care, even if such a policy is not cost-effective. These high quality studies (Table 17.4), among others (e.g., Lubchenco et al., 1989), used various inpatient data sources to universally endorse high-risk deliveries at hospitals with high levels of care.

Studies of Hospital Delivery Volume and Effects on Neonatal Outcomes

During the development of regionalization, the emphasis was on ways to ensure that expectant mothers and high-risk infants received the appropriate level of care. With the proliferation of high level NICUs, the significance of patient volume in each hospital has received greater attention. (These studies are in Tables 17.3 and 17.4, clustered closer to the bottom as these citations are listed chronologically.) While too many patients may stretch hospital resources, hospitals with too few may never develop the expertise to deliver high quality care. Phibbs et al. (1996) showed that mortality decreased as patient census increased, even within Level III NICUs. In contrast, Horbar, Badger, Lewit, Rogowski, and Shiono (1997) did not find an association in their examination of VLBW infants according to hospital volume of delivery in a group of hospitals in the Vermont Oxford Network. Moster, Lie, and Markestad (1999) found that low-risk deliveries were safest when roughly 2,000–3,000 births occurred per unit. Cifuentes et al. (2002) confirmed the findings of Phibbs et al. by also finding a strong association between volume, level of care and improved survival. Likewise, Heller, Richardson, and Schnell (2002) demonstrated in Germany that the dissemination of higher level care was accompanied by a lower volume per unit. Low volume units had persistently higher mortality rates, even for low-risk infants, raising the question of whether low volume units were providing quality care.

While the match between higher risk infants and higher levels of care has remained the goal of regionalization, clearly the volume of patients significantly modifies its effect of on newborn outcomes. Collectively, these findings have translated into policy recommendations. Beginning in 1990, for example, the effective and timely transportation of high-risk mothers prior to delivery has become a Healthy People benchmark through the stated goal of increasing VLBW deliveries at level III centers (Healthy Children 2000, n.d.). Most recently, a definitive article by Phibbs et al. (2007) found VLBW deliveries to have the lowest mortality at high level *and* high volume nurseries, a finding which will require further policy response in the coming years. Combined with the results from studies on ‘de-regionalization’ (see below), these articles, employing novel, high quality methods, show that both high level and high volume should be a dual goal of all communities for regionalized perinatal services.

Studies that Examine “De-Regionalization” and Its Effects on Neonatal Outcomes

De-regionalization, as a trend, merits separate mention. Beginning in the 1980s and 1990s, market forces challenged the goals of perinatal regionalization. Rising health care costs led to changes in payment mechanisms, such as managed care, that often discouraged inter-hospital referrals. Further, the concurrent expansion of available perinatal workforce and technologies made these once scarce resources much more accessible. Growth in neonatal intensive care capacity occurred without prospective targets. As early as 1985, the Committee on the Fetus and Newborn of the American Academy of Pediatrics suggested that there was already an adequate supply of neonatologists to fill the current level II and III nurseries (American Academy of Pediatrics, 1985). Despite this

assessment, the number of units and neonatologists continued to increase, changing the paradigm from limited access to higher levels of perinatal care to competition for patients among hospitals. In 1998, Pollack et al. reported that neonatologists were increasingly employed in well-child nurseries instead of higher-level NICUs, although the benefits to newborns of this practice are unknown. The growth of reproductive technology and the consequent increase in multiple births requiring NICU care (Hashimoto, Lindsell, Brewer, Eichel, & Donovan, 2004), the financial advantages of NICU services, and the increased national attention to the needs of premature infants (March of Dimes, n.d.), fueled the expansion of NICU care.

Studies of de-regionalization mark new approaches to evaluations of regionalization. Case reports called attention to the disintegration of perinatal regionalization due to the market forces discussed above (Gagnon, Allison-Cooke, & Schwartz, 1988; Richardson et al., 1995). These studies found, for example, that the decision to transport mothers and infants was based more on insurance status than clinical risk. These papers stimulated the creation of TIOP II (1993), the aims of which were to reinforce the importance of appropriate perinatal inpatient care and to expand regionalization efforts to include preconception and prepartum care. Despite these high quality studies and efforts to counterbalance economic trends, de-regionalization continued. For example, for-profit, publicly held corporations that also manage inter-hospital perinatal services, began employing neonatologists and obstetricians (e.g., Pediatrix). Yeast et al. (1998) showed that high-risk deliveries were shifting away from tertiary/level III centers, with a greater retention of high-risk mothers at level II centers. While it may be argued that increased knowledge and technology had disseminated into level II hospitals, Yeast confirmed that VLBW infants had higher mortality if born anywhere but level III nurseries (see Tables 17.3 and 17.4).

Studies that Examine Location of VLBW Births and Maternal Transfer

Other evidence mounted that the original applications of regionalization were being challenged using outcomes such as location of birth and maternal transfers (the majority of studies in this section are in Tables 17.5 and 17.6). Gould, Sarnoff, Liu, Bell, and Chavez (1999) used the delivery of VLBW infants (<1,500 g) at level I hospitals as a non-controversial indicator of de-regionalization. They observed a wide variation in rates of maternal transport of impending VLBW deliveries (more than sevenfold) across perinatal regions in California, and also noted that race, ethnicity, and other socio-economic risk factors independently influenced these rates. Further studies showed that hospital characteristics or individual insurance status influenced maternal

Table 17.6 Quality rating of studies associated with regionalization: effect of regionalization on location of VLBW births and rates of maternal transfers

Author (year)	Reporting	External validity	Internal validity – bias	Internal validity – confounding	Power	Total quality score	Suitability of study to assess effectiveness: greatest, moderate, least
Handler et al. (1991)	6	2	3	3	0	14	Moderate
Gould et al. (1999)	11	3	8	4	0	26	Greatest
Mehta et al. (2000)	8	3	4	4	0	19	Moderate
Wall et al. (2004)	10	3	7	4	0	24	Greatest
Sinkin et al. (2005)	10	3	6	4	0	23	Greatest

transport beyond the identifiable risks of expectant mothers [Dooley, Freels, & Turnock, 1997; Handler et al., 1991; Sinkin, Fisher, Dozier, & Dye, 2005; Wall, Handler, & Park, 2004 (Note: Dooley et al. can be found in Tables 17.3 and 17.4)]. Because transport systems rely on voluntary agreements between hospitals, they are difficult to maintain when other forces, particularly economic, are also involved.

While one might expect that more overall resources would enhance newborn survival, Goodman et al. (2002) (Tables 17.3 and 17.4), however, showed otherwise. By measuring outcomes in various regions to reflect the distribution of newborn health services, they found that increased neonatal resources did not reduce NMR beyond a relatively low threshold of neonatologists (4.3 neonatologists per 10,000 births). Further, areas with greater neonatal intensive care capacity (increased numbers of providers, beds, hospitals, etc.) were not located in places of highest perinatal need (Goodman et al., 2001). The original intent of regionalization – to create regional systems that could meet the clinical demands of patients most in need of specialized perinatal services – had been replaced by an expanded and expensive workforce competing for a perinatal population.

At present, the debate continues about the role of regionalization in perinatal services. On the one hand, there is continual, high-quality evidence in support of regionalization, such as the 2006 study by Lui et al. (Tables 17.3 and 17.4) in Australia that showed improved survival of VLBW infants in areas with the introduction of statewide coordinated strategies for referrals including perinatal advice to promote transfers prior to delivery. On the other hand, there are continued economic pressures, such as the expansion of highly technical neonatal care into low-risk settings, that continue to challenge the development of regionalization. The studies on de-regionalization, by providing evidence that the decline of regionalization leads to worse outcomes, support policies that prioritize regionalized perinatal services for all infants.

Review of the Evidence of the Effectiveness of Regionalization for Reducing Racial and Ethnic Disparities in Newborn Outcomes

Based on the evidence summarized in this chapter, regionalized perinatal systems have been credited, in part, with the significant reductions in infant and neonatal mortality over the past five decades. (The additional studies discussed in this section are found in Tables 17.7 and 17.8; relevant studies from Tables 17.3 and 17.5 are listed in the footnote in Table 17.7.) At the same time, epidemiological research has unequivocally demonstrated persistent disparities in newborn outcomes by race and ethnicity (Allen, Alexander, Tompkins, & Hulsey, 2000; Berger, Udry, & Hendricks, 1975; Carmichael & Iyasu, 1998; Morse et al., 2006). For example, African-American infants have twice the risk of LBW as compared to their European-American counterparts. Paradoxically, African-American infants also have the lowest birth-weight specific mortality for preterm infants, conferring a relative advantage in survival for the NICU population (Morse et al.). Among full term infants, African-Americans have the highest IMR (Thompson, Goodman, & Little, 2002). In addition, the favorable newborn outcomes within Hispanic populations, with high rates of newborn survival despite socioeconomic disparities, remain unexplained. These advantageous outcomes have recently declined, apparently due to influences of acculturation and the "American" environment (Jenny, Schoendorf, & Parker, 2001). Disparities in perinatal outcomes are profound and have persisted over the past decades, and logically should serve as important outcomes in studies of the effectiveness of regionalization of perinatal services.

The majority of studies that examine regionalization as a tool for reducing IMR or NMR do not directly measure the contribution of health services, and regionalization in particular, to reducing

Table 17.7 Health status outcome: effect of regionalization on racial/ethnic differences in newborn outcomes

Author (year) Study design	Study type	Description of intervention/ outcome	Populations studied and sample size	Address disparities (yes/no)[a]	Key findings related to intervention effectiveness	Caveats/biases	Findings support the intervention? Yes/no For which populations?
Berger et al. (1975) RC[c]	Published article	Models how regionalization might decrease racial disparities in PMRs[b]	North Carolina, 1970–1972, n = 285,191	Yes	Birth weight and gestation account for almost all of risk of death; likely 5% reduction in PMR disparity between white and black infants	Theoretical analyses using prior data	Yes, supports regionalization, but notes that it would need to focus on prevention of prematurity to improve mortality rates in Black infants
Kugler et al. (1990) RC[c]	Published article	LBW[d] and NMR[e] by black vs. white, military vs. civilian health systems	Infants in Washington State, 1982–1985, n = 29,848	Yes	LBW[d] doubled, but NMR[e] for normal birth weight black infants = NMR for white infants in military hospitals (OR 1.08, CI 0.6, 2.0). NMR[e] in LBW[d] in black military infants elevated vs. whites (OR 1.8, CI 1.5, 2.2)	Selection bias, hard to control for other variables	Yes, spectrum of care provided in military health care systems narrows disparities in normal weight infants; no effect on LBW[d] and NMR[e] in LBW[d]
Morales et al. (2005) RC[c]	Published article	Influence of hospital characteristics (minority-serving) on black and white NMR[e] in VLBW[f] infants	1995–2000, n = 332 NICUs[g], n = 74,050 VLBW[f]	Yes	Minority serving hospitals have higher NMR[e] for black (OR 1.29, CI 1.01, 1.64); white (1.30, CI 1.09, 1.56); and pooled (1.28, CI 1.10, 1.50)	Unclear why these minority serving hospitals lead to increased mortality	Yes, cautiously: hospital factors influence NMR[e] beyond hospital level or volume

[a] Articles from Tables 17.3 and 17.5 with extensive evaluations of racial/ethnic disparities noted in their analyses were not re-listed here. Authors Gortmaker, Schwartz, Langkamp, Yeast, Menard, Phibbs, Cifuentes, LeFevre, Gould, Sinkin, and Wall contribute findings that examine the intersection of racial/ethnic disparities and regionalization of perinatal services

[b] *PMR* Perinatal mortality rate: the rate of fetal death plus early neonatal deaths (up to 1 week of life) divided by the total number of live births

[c] *RC* Retrospective cohort

[d] *LBW* Low birth weight: birth weight less than 2,500 g

[e] *NMR* Neonatal mortality rate: the rate of death within the first 28 days of life divided by the total number of live births

[f] *VLBW* Very low birth weight: birth weight less than 1,500 g

[g] *NICU* Neonatal intensive care unit

Table 17.8 Quality rating of studies associated with regionalization: effect of regionalization on racial/ethnic differences in newborn outcomes

Author (year)	Reporting	External validity	Internal validity – bias	Internal validity – confounding	Power	Total quality score	Suitability of study to assess effectiveness: greatest, moderate, least
Berger et al. (1975)	5	0	5	3	0	13	Moderate
Kugler et al. (1990)	9	2	5	3	0	19	Moderate
Morales et al. (2005)	12	2	7	5	0	26	Greatest

racial and ethnic disparities. Perhaps it has been assumed that through the systematic adoption of an organized health system, regionalization would benefit all infants, above all those at highest risk. Vital statistics have shown otherwise (Centers for Disease Control and Prevention, 2005; Mathews et al., 2004). As early as 1975, Berger et al. hypothesized that regionalization's focus on the clinical needs of inpatient, peripartum care would not address the underlying *causes* of maternal and newborn illness. By estimating the potential effects of regionalization on outcomes of African-American infants, they predicted that regionalization would not reduce disparities in outcomes of African-American infants compared to other populations because this systematic organization of perinatal care could not adequately affect the elevated LBW rates in African-Americans. Additional studies have supported this prediction, such as an economic analysis that showed that hospitals offering tertiary maternal and newborn services often do not provide sufficient care, as defined by regionalized perinatal services, for lower income pregnant women at high risk (Perkins, 2004). Others hypothesize, however, that the disparities in newborn outcomes in some racial and ethnic groups may have been attenuated since high level perinatal services tend to be more prevalent in urban areas, providing increased access in the same areas where minorities tend to live (Bronstein, Capilouto, Carlo, Haywood, & Goldenberg, 1995; Gortmaker et al., 1985; Langkamp, Foye, & Roghmann, 1990; Schwartz, Muri, Overpeck, Pezzullo & Kogan, 2000). Rogowski, Staiger, and Horbar (2004) and Rogowski, Horbar, et al. (2004) showed that African-American infants had as much variation in NMRs as other infants, and that they would benefit at least as much as white infants if hospitals performed at the highest levels.

Researchers have tried to grapple with this problem in different ways. Some studies on the effectiveness of regionalization as an intervention, especially those in the last decade, have used statistical techniques to adjust for racial and ethnic differences, instead of directly examining disparities (Cifuentes et al., 2002; Menard et al., 1998; Phibbs et al., 1996; Yeast et al., 1998). Others present secondary analyses with racial and/or ethnic estimates included with their overall hypotheses. Although not the focus of these studies, the results have raised questions yet to be answered about the influence of health services on disparities in racial and ethnic minorities. For example, Gortmaker et al. (1985) and LeFevre et al. (1992), despite showing improved infant survival at higher levels of care overall, were unable to demonstrate significant decreases in NMRs in African-Americans at level III hospitals compared to level I hospitals. While Menard et al. found a benefit for VLBW African-American infants at level III vs. level I hospitals, they also found that the rate of birth-weight *adjusted* mortality (not birth weight-specific mortality) for African-American infants was higher than white infants within similar levels of care. Other researchers have shown that the organization and delivery of health services varies by individual race and ethnicity (Gould et al., 1999; Sinkin et al., 2005; Wall et al., 2004). These studies support the possibility that regionalization, defined in terms of hospital levels and maternal transports, may not benefit African-American infants to the same extent as white infants.

Morales et al. (2005) in a novel and explicit way, showed that hospitals that predominantly serve minority infants had higher NMR for both white and African-American infants than hospitals with

low levels of minority patients. They further showed that white and African-American infants had similar mortality rates within hospitals with similar levels of minority patients. Since VLBW African-American infants are more likely to be treated at hospitals in which the majority of patients are of color, this study suggests that improving quality of care at "minority serving hospitals" may actually address disparities in neonatal outcomes. Importantly, this study includes characteristics of the patient population (such as the percent of patients from a particular demographic) as a hospital-level factor, similar to patient volume and level of care, which are also associated with infant survival and, more specifically, racial disparities. Additional research is necessary to validate the few efforts that focus directly on the effects of regionalization on racial/ethnic disparities.

Other studies have directly examined the interplay of health services in general and racial and ethnic disparities in perinatal outcomes. While these studies do not directly mention regionalization, these high-quality studies address aspects of health services that are linked to regionalization. Kugler, Connell, and Henley (1990) compared care in military and civilian hospitals and found equivalent NMRs in normal birth weight white and African-American infants in Washington State when infants were born in military hospitals. They concluded that the overall military health care offered before and throughout pregnancy appeared to infer increased survival. They further showed that even though all African-American infants in this cohort had an elevated rate of LBW and subsequent increased NMR in these LBW infants, only those normal birth weight African-American infants born at *civilian* hospitals had a doubled mortality similar to rates continually observed in national studies. In addition, Langkamp et al. (1990) showed that while African-American infants in New York State experienced higher neonatal mortality rates (OR 2.15), they had even higher rates of NICU admissions (OR 2.6), concluding that disparities in outcomes predate NICU admissions. These data suggest that interventions that start prior to the peripartum (inpatient) period will be necessary to narrow racial/ethnic disparities in newborn outcomes.

Relevance of Evidence for Practitioners

In sum, regionalization as a public health policy is an effective systematic intervention for inpatient perinatal health services, and health care providers and policy makers should continue to support its implementation. Individual practitioners may find it difficult to observe the benefits of public health efforts, but should uphold its principles nonetheless to enhance population-level improvements in newborn outcomes. High-risk infants have better outcomes when born at high level, high volume hospitals. However, despite its current application, the concepts of regionalization remain untested in three important and inter-related ways. First, regionalization has only been implemented for the peripartum period, most likely because health care financing favors inpatient services as opposed to preventive care. At this time, the loci of interventions need to expand outside of the NICU. Second, there have yet to be effective ways to monitor regionalized perinatal systems to insure that the goals are being adhered to. Finally, the role of regionalization as an intervention has not reduced racial and ethnic disparities since the focus on high-risk inpatient care does not address the causes of perinatal disease that lead to NICU admissions. As summarized in this chapter, preliminary evidence suggests that there is a differential impact of regionalization by race/ethnicity but the majority of the evidence is inconclusive. Some supportive evidence appears to be due to factors at the time of hospitalization and some due to factors that predate hospitalization. As such, there remains a larger, likely more effective opportunity to address racial and ethnic disparities in newborn outcomes through regionalization of the full spectrum of care for women and infants, starting with family planning and preconception counseling and continuing with primary care for newborns.

Research, Regionalization, and the Future

Reducing disparities in newborn outcomes through regionalization requires major changes in the structure, coordination, and financing of health services for women and children. State agencies and providers need to define and coordinate their goals and then monitor population outcomes of their entire birth cohort. The current system, in which statewide Title V agencies often play a role, seems to lack sufficient capacity for fully assessing and implementing change (Johnson & Little, 1999), but perhaps it is possible that future modifications could reach this goal. Payers will need to reward preventive services in addition to NICU care. While it is unknown exactly how to circumvent these current financial and organizational barriers, the opportunities to reduce disparities are clear. For example, regionalized services that include access to family planning would likely reduce unintended pregnancy, a dominant risk factor for poor newborn outcomes. Current recommendations to expand and standardize preconception health from the Centers for Disease Control (Johnson et al., 2006) are making important efforts in this direction and include a central recommendation for enhanced regional surveillance of related public health indicators (Johnson et al.). It is now possible that the balance to which Richardson alluded to in 1998, with two-thirds of improvements in neonatal mortality due to all aspects of NICU care, might shift further towards policies to address the preconception and prenatal aspects of perinatal care. It is expected that this shift will finally begin to address the on-going racial and ethnic disparities in newborn outcomes.

Nonetheless, the importance of regionalization to advances in perinatal care cannot be underestimated as it is a rare example of coordination of services across health care systems. While the original goals of regionalization are being continually challenged by the expansion of scarce resources to a situation of plenty and the emergence of hospital competition, the central goals of regionalization of perinatal services have persisted, even if in practice it remains focused on inpatient care. There are now opportunities for practitioners and policy makers alike to apply the concepts of regionalization to domains of maternal and newborn care beyond the inpatient peripartum period; it is hoped that this effort will lead in turn to a reduction in racial and ethnic disparities in perinatal outcomes.

References

Allen, M. C., Alexander, G. R., Tompkins, M. E., & Hulsey, T. C. (2000). Racial differences in temporal changes in newborn viability and survival by gestational age. *Paediatric and Perinatal Epidemiology*, *14*(2), 152–158.

American Academy of Pediatrics. Committee on Fetus and Newborn. (1985). Manpower needs in neonatal pediatrics. *Pediatrics*, *76*(1), 132–135.

American Academy of Pediatrics. Committee on Fetus and Newborn. (2004). Levels of neonatal care. *Pediatrics*, *114*(5), 1341–1347.

American Academy of Pediatrics & The American College of Obstetricians and Gynecologists. (2007). *Guidelines for perinatal care* (6th ed.). Washington, DC: Author.

Berger, G. S., Udry, J. R., & Hendricks, C. H. (1975). Regionalized perinatal care: An estimate of potential effect on racial differences in perinatal mortality in North Carolina. *North Carolina Medical Journal*, *36*(8), 476–479.

Bronstein, J. M., Capilouto, E., Carlo, W. A., Haywood, J. L., & Goldenberg, R. L. (1995). Access to neonatal intensive care for low-birthweight infants: The role of maternal characteristics. *American Journal of Public Health*, *85*(3), 357–361.

Carmichael, S. L., & Iyasu, S. (1998). Changes in the black-white infant mortality gap from 1983 to 1991 in the United States. *American Journal of Preventive Medicine*, *15*(3), 220–227.

Centers for Disease Control and Prevention (CDC). (2005). "Racial/ethnic disparities in infant mortality–United States, 1995–2002." *MMWR Morb Mortal Wkly Rep*. 10, *54*(22), 553–556.

Chien, L., Whyte, R., Aziz, K., Thiessen, P., Matthew, D., & Lee, S. (2001). Improved outcome of preterm infants when delivered in tertiary care centers. *Obstetrics and Gynecology*, *98*(2), 247–252.

Cifuentes, J., Bronstein, J., Phibbs, C., Phibbs, R., Schmitt, S., & Carlo, W. (2002). Mortality in low birth weight infants according to level of neonatal care at hospital of birth. *Pediatrics*, *109*, 745–751.

Committee on Perinatal Health. (1976). *Toward improving the outcomes of pregnancy: Recommendations for the regional development of maternal and perinatal health services. (TIOP I)*. White Plains, NY: March of Dimes.

Committee on Perinatal Health. (1993). *Toward improving the outcomes of pregnancy: The 90s and beyond. (TIOP II)*. White Plains, NY: March of Dimes.

Dooley, S., Freels, S., & Turnock, B. (1997). Quality assessment of perinatal regionalization by multivariate analysis: Illinois. *Obstetrics and Gynecology*, *89*, 193–198.

Gagnon, D., Allison-Cooke, S., & Schwartz, R. (1988). Perinatal care: The threat of deregionalization. *Pediatric Annals*, *17*, 447–452.

Goodman, D. C., Fisher, E. S., Little, G. A., Stukel, T. A., & Chang, C. H. (2001). Are neonatal intensive care resources located according to need? Regional variation in neonatologists, beds, and low birth weight newborns. *Pediatrics*, *108*(2), 426–431.

Goodman, D. C., Fisher, E., Little, G., Stukel, T., Chang, C., & Schoendorf, K. (2002). The relation between availability of neonatal intensive care and neonatal mortality. *New England Journal of Medicine*, *346*, 1538–1544.

Gortmaker, S., Sobol, A., Clark, C., Walker, D. K., & Geronimus, A. (1985). The survival of very low-birth weight infants by level of hospital of birth: A population study of perinatal systems in four states. *American Journal of Obstetrics and Gynecology*, *152*(5), 517–524.

Gould, J. B., Sarnoff, R., Liu, H., Bell, D. R., & Chavez, G. (1999). Very low birth weight births at non-NICU hospitals: The role of sociodemographic, perinatal, and geographic factors. *Journal of Perinatology*, *19*(3), 197–205.

Graven, S. N., Howe, G., & Callon, H. (1976). Perinatal health care studies and program results in Wisconsin, 1964–1970. In P. A. Swyer & J. A. Stetson (Eds.), *Neonatal intensive care*. St. Louis, MO: Charles Green Publisher.

Handler, A., Rosenberg, D., Driscoll, M., Cohen, M., Swift, E., Garcia, P., et al. (1991). Regional perinatal care in crisis: A case study of an urban public hospital. *Journal of Public Health Policy*, *12*(2), 184–198.

Hashimoto, L. N., Lindsell, C. J., Brewer, D. E., Eichel, M. M., & Donovan, E. F. (2004). Contributions of infertility treatment to very-low-birth-weight multiple birth infants receiving neonatal intensive care. *American Journal of Obstetrics and Gynecology*, *190*(2), 401–406.

Healthy Children 2000. (n.d.). *National health promotion and disease prevention objectives*. Retrieved August 22, 2007, from http://odphp.osophs.dhhs.gov/pubs/hp2000/.

Hein, H. A., & Lathrop, S. S. (1986). The changing pattern of neonatal mortality in a regionalized system of perinatal care. *American Journal of Diseases of Children*, *140*(10), 989–993.

Heller, G., Richardson, D. K., & Schnell, R. (2002). Are we regionalized enough? Early neonatal death in low-risk births by the size of delivery units in Hesse, Germany. *International Journal of Epidemiology*, *31*, 1061–1068.

Horbar, J. D., Badger, G. J., Lewit, E. M., Rogowski, J., & Shiono, P. H. (1997). Hospital and patient characteristics associated with variation in 28-day mortality rates for very low birth weight infants. Vermont Oxford Network. *Pediatrics*, *99*(2), 149–156.

Howell, E., & Vert, P. (1993). Neonatal intensive care and birth weight-specific perinatal mortality in Michigan and Lorraine. *Pediatrics*, *91*, 464–469.

Jenny, A. M., Schoendorf, K. C., & Parker, J. D. (2001). The association between community context and mortality among Mexican-American infants. *Ethnicity & Disease*, *11*, 722–731.

Johnson, K. A., & Little, G. A. (1999). State health agencies and quality improvement in perinatal care. *Pediatrics*, *103*(1 Suppl. E), 233–247.

Johnson, K., Posner, S. F., Biermann, J., Cordero, J. F., Atrash, H. K., Parker, C. S., et al. (2006). Recommendations to improve preconception health and health care – United States. A report of the CDC/ATSDR Preconception Care Work Group and the Select Panel on Preconception Care. *MMWR. Recommendations and Reports*, *55*(RR-6), 1–23.

Kirby, R. (1996). Perinatal mortality: The role of hospital birth. *Journal of Perinatology*, *16*, 43–49.

Koontz, A. M. (1984). Pregnancy and infant health: Progress toward the 1990 objectives. *Public Health Reports*, *99*(2), 184–192.

Kugler, J. P., Connell, F. A., & Henley, C. E. (1990). Lack of difference in neonatal mortality between blacks and whites served by the same medical care system. *Journal of Family Practice*, *30*(3), 281–288.

Langkamp, D., Foye, H., & Roghmann, K. (1990). Does limited access to NICU services account for higher neonatal mortality rates among blacks? *American Journal of Perinatology*, *7*, 227–231.

LeFevre, M., Sanner, L., Anderson, S., & Tsutakawa, R. (1992). The relationship between neonatal mortality and hospital level. *Journal of Family Practice*, *35*(3), 259–264.

Lubchenco, L. O., Butterfield, L. J., Delaney-Black, V., Goldson, E., Koops, B. L., & Lazotte, D. C. (1989). Outcome of very-low-birth-weight infants: Does antepartum versus neonatal referral have a better impact on mortality, morbidity, or long-term outcome? *American Journal of Obstetrics and Gynecology*, *160*, 539–545.

Lui, K., Abdel-Latif, M. E., Allgood, C. L., Bajuk, B., Oei, J., Berry, A., et al. (2006). Improved outcomes of extremely premature outborn infants: Effects of strategic changes in perinatal and retrieval services. *Pediatrics*, *118*, 2076–2083.

March of Dimes. (n.d.). *March of Dimes Prematurity Campaign*. Retrieved August 30, 2007, from http://www.marchofdimes.com/prematurity/5408_5576.asp.

Mathews, T. J., Menacker, F., & MacDorman, M. F. (2004). *Infant mortality statistics from the 2002 period linked birth/ infant death data set. National Vital Statistics Reports* (Vol. 53). Hyattsville, MD: National Center for Health Statistics.

Mayfield, J., Rosenblatt, R., Baldwin, L., Chu, L., & Logerfo, J. (1990). The relation of obstetrical volume and nursery level of perinatal mortality. *American Journal of Public Health*, *80*, 819–823.

McCormick, M., Shapiro, S., & Starfield, B. (1985). The regionalization of perinatal services. Summary of the evaluation of a national demonstration program. *New England Journal of Medicine*, *253*, 799–804.

Mehta, S., Atherton, H., Schoettker, P., Hornung, R., Perlstein, P., & Kotagal, U. (2000). Differential Markers for Regionalization. *Journal of Perinataology*, *20*(6), 366–372.

Menard, M., Liu, Q., Holgren, E., & Sappenfield W. (1998). Neonatal mortality for very low birth weight deliveries in South Carolina by level of hospital perinatal service. *American Journal of Obstetrics and Gynecology*, *179*(2), 372–381.

Merlo, J., Gerdtham, U., Eckerlund, I., Hakansson, S., Olausson, O. P., Pakkanen, M., et al. (2005). Hospital level of care and neonatal mortality in low- and high-risk deliveries. *Medical Care*, *43*(11), 1092–1100.

Morales, L. S., Staiger, D., Horbar, J. D., Carpenter, J., Kenny, M., Geppert, J., et al. (2005). Mortality among very low-birthweight infants in hospitals serving minority populations. *American Journal of Public Health*, *95*(12), 2206–2212.

Morse, S. B., Wu, S. S., Ma, C., Ariet, M., Resnick, M., & Roth, J. (2006). Racial and gender differences in the viability of extremely low birth weight infants: A population-based study. *Pediatrics*, *117*(1), 106–112.

Moster, D., Lie, R., & Markestad, T. (1999). Relation between size of delivery unit and neonatal death in low risk deliveries: Population based study. *Archives of Disease in Childhood. Fetal and Neonatal Edition*, *80*, F221–F225.

Pediatrix Medical Group. (n.d.). Retrieved August 22, 2007, from http://www.pediatrix.com/.

Perkins, B. B. (2004). Perinatal regionalization and economic order. In *The medical delivery business: Health reform, childbirth and the economic order* (pp. 85–100). New Brunswick, NJ and London: Rutgers University Press.

Phibbs, C. S., Baker, L. C., Caughey, A. B., Danielsen, B., Schmitt, S. K., & Phibbs, R. H. (2007). Level and volume of neonatal intensive care and mortality in very-low-birth-weight infants. *New England Journal of Medicine*, *356*(21), 2165–2175.

Phibbs, C., Bronstein, J., Buxton, E., & Phibbs, R. (1996). The effects of patient volume and level of care at the hospital of birth on neonatal mortality. *Journal of the American Medical Association*, *276*, 1053–1059.

Pollack, L. D., Ratner, I. M., & Lund, G. C. (1998). United States neonatology practice survey: Personnel, practice, hospital and neonatal intensive care unit characteristics. *Pediatrics*, *101*, 398–405.

Powell, S., Holt, V., Hickok, D., Easterling, T., & Connell, F. (1995). Recent changes in delivery site of low-birth-weight infants in Washington: Impact on birth weight-specific mortality. *American Journal of Obstetrics and Gynecology*, *173*, 1585–1592.

Richardson, D. K., Gray, J. E., Gortmaker, S. L., Goldmann, D. A., Pursley, D. M., & McCormick, M. C. (1998). Declining severity adjusted mortality: Evidence of improving neonatal intensive care. *Pediatrics*, *102*(4 Pt 1), 893–899.

Richardson, D. K., Reed, K., Cutler, J. C., Boardman, R. C., Goodman, K., Moynihan, T., et al. (1995). Perinatal regionalization versus hospital competition: The Hartford example. *Pediatrics*, *96*(3 Pt 1), 417–423.

Rogowski, J. A., Staiger, D. O., & Horbar, J. D. (2004). Variations in the quality of care for very-low-birthweight infants: Implications for policy. *Health Affairs (Millwood)*, *23*(5), 88–97.

Rosenblatt, R., Macfarlane, A., Dawson, A., Cartilidge, P., Larson, E., & Hart, L. (1996). The regionalization of perinatal care in Wales and Washington State. *American Journal of Public Health*, *86*, 1011–1015.

Schwartz, R. M., Muri, J. H., Overpeck, M. D., Pezzullo, J. C., & Kogan, M. D. (2000). Use of high-technology care among women with high-risk pregnancies in the United States. *Maternal and Child Health Journal*, *4*(1), 7–18.

Sinkin, R., Fisher, S., Dozier, A., & Dye, T. (2005). Effect of managed care on perinatal transports for the publicly funded in upstate New York. *Journal of Perinatology*, *25*, 79–85.

The Vermont Oxford Network. (n.d.). Retrieved August 24, 2007, from http://www.vtoxford.org/.

Thompson, L. A., Goodman, D. C., & Little, G. A. (2002). Is more neonatal intensive care always better? Insights from a cross-national comparison of reproductive care. *Pediatrics*, *109*(6), 1036–1043.

U.S. Department of Health and Human Services. Healthy People 2010. (n.d.). Retrieved August 22, 2007, from http://www.healthypeople.gov/.

Wall, S., Handler, A., & Park, C. (2004). Hospital factors and nontransfer of small babies: A marker of deregionalized perinatal care? *Journal of Perinatology*, *24*(6), 351–359.

Yeast, J., Poskin, M., Stockbauer, J., & Shaffer, S. (1998). Changing patterns in regionalization of perinatal care and the impact on neonatal mortality. *American Journal of Obstetrics and Gynecology*, *178*(1), 131–135.

Index

A. Handler et al. (eds.), *Reducing Racial/Ethnic Disparities in Reproductive and Perinatal Outcomes: The Evidence from Population-Based Interventions*, DOI 10.1007/978-1-4419-1499-6,