

Introduction to the Special Issue: Subgroup Analysis in Prevention and Intervention Research

Lauren H. Supplee · Brendan C. Kelly ·
David M. MacKinnon · Meryl Yoches Barofsky

Published online: 23 October 2012
© Society for Prevention Research (outside the USA) 2012

Prevention scientists, intervention developers, patients, providers, and clients are continually seeking more effective and efficient treatments for a wide range of social, behavioral, and public health problems. Across this range of problems, there is a common interest in developing a better understanding of impacts of interventions on specific subgroups. Among policy-makers, an interest in the question “What works?” is now often accompanied by “What works for whom?” For example, the Obama administration has been emphasizing the use of rigorous research as part of evidence-based policy-making. In a 2009 memorandum to federal agencies and departments, the Office of Management and Budget emphasized both the importance of rigorous research on program effectiveness as well as evidence aimed at improving the life outcomes of individuals. In medicine and health policy, there has been a strong push toward comparative effectiveness research; the purpose of which is “to inform patients, providers, and decision-makers, responding to their expressed needs about *which interventions are most effective for which patients under specific circumstances*” (Federal Coordinating Council for Comparative Effectiveness Research 2009, p 3). One implication of this emphasis on what works has been greater attention to the research and methodologies associated with specific populations and subgroups.

The analysis of subgroups can matter a great deal in prevention science and intervention research. First, many prevention scientists use subgroup findings to unpack significant main effects or to investigate why there was a lack of significant main effects. Prevention scientists frequently want to explore whether a program was more or less effective for a segment of the target population and why it may have been less effective for another subgroup. Rothwell (2005) argues that the importance of subgroup analysis is not in differential response to treatment, but in identifying how to maximize benefits from treatments and mitigate risk. Rothman (2012) highlights that moderated effects can help prevention scientists to refine theory, tailor prevention to specific contexts or to the needs of specific populations, or target intervention. Policy-relevant research around prevention and intervention science is regularly challenged to answer the question of what works for whom. Subgroup analysis can be seen to be directly linked to policy decisions around programmatic aims (e.g., Upward Bound), funding decisions (e.g., Even Start), and new initiatives targeting funding towards evidence-based programs (e.g., teen pregnancy and home visitation).

Subgroup analysis, broadly, aims to measure change within and between groups. Subgroups are defined by characteristics measured at baseline. Subgroups can be characterized by variables that are easier to define, such as age, to those less well-defined, such as risk status. Subgroups can be continuous or categorical; variables with low, moderate, or high measurement error; and variables that are measured, latent, or estimated based on response to treatment. Subgroups can include individual characteristics or site-level variables. Subgroups may occur with more regularity in a population, such as gender, more infrequently such as families with multiple risk factors, or more difficult to represent in large numbers in prevention trials such as rural communities.

Work by Rothwell (2005) and Wang et al. (2007) highlighted the issues with subgroup analysis within medical research. Rothwell (2005) specified best practices in study design,

L. H. Supplee (✉) · B. C. Kelly
Office of Planning, Research and Evaluation,
Administration for Children and Families,
370 L'Enfant Promenade SW, 7th Fl West,
Washington, DC 20447, USA
e-mail: lauren.supplee@acf.hhs.gov

D. M. MacKinnon
Arizona State University,
Phoenix, USA

M. Y. Barofsky
University of Maryland,
College Park, USA

analysis, and interpretation of subgroups in medical research. On the surface, these recommendations are reasonable and based on sound methodology around subgroup analysis. However, implementing the recommendations has inherent challenges, especially in the context of prevention and intervention research. For example, one of the recommendations is that trials should be designed and powered appropriately for confirmatory subgroup analysis. In medical trials where sample enrollment, time, and treatment dosage may be significantly less costly than prevention research, increasing sample sizes substantially may not be as challenging as it can be in prevention science. Similarly, medical trials may facilitate replication of exploratory subgroup findings more easily than in prevention and intervention research.

Given some of the unique challenges around subgroup analysis in prevention science and the greater attention to the use of evidence to inform policy decision-making, the Administration for Children and Families along with other federal partners (the Office of the Assistant Secretary for Planning and Evaluation in the U.S. Department of Health and Human Services, the Division of Violence Prevention in the Centers for Disease Control and Prevention, the Institute for Education Sciences, the National Institute on Drug Abuse, and the National Institute of Mental Health) convened a meeting of experts to discuss the state of the field. Well-designed research that analyzes subgroup analyses appropriately has the potential to inform the field on how to maximize treatment and steer limited public and private resources with informed decision-making.

This Special Issue contains papers that draw from this interagency meeting and highlight cutting-edge thinking about the topic of subgroups. The papers naturally fell into two categories: (a) design and analysis and (b) reporting and interpreting. During the meeting, there was some consensus from the participants about some core topics, such as the need to specify confirmatory and exploratory subgroups, but there was less agreement on other topics; for example, whether to design studies assuming heterogeneity or homogeneity in a study sample. Since subgroup analysis is of broad interest across a number of fields, the intent of the interagency meeting was to begin a conversation. This Special Issue continues that dialogue. In this article, we summarize the discussion at the meeting and highlight the papers in the current Special Issue that begin to address key issues in subgroup analysis.

Design and Analysis

During the meeting, Wang (2010) and others identified four of the most important challenges that arise in identifying and analyzing subgroup impacts. These challenges include: clearly specifying confirmatory versus exploratory tests; limited power to detect subgroup effects in analyses; pre-specification

of hypotheses and subgroups versus post hoc analysis and contextualization of results; and appropriately adjusting for multiple comparisons increasing Type I error rates. There was consensus about the definition of confirmatory and exploratory subgroups and the importance of specification at the outset. Confirmatory tests are pre-specified based on baseline characteristics, whereas exploratory subgroups include most other subgroup analyses. Exploratory subgroups, in general, tend to be more subject to bias and harder to interpret.

Questions remain about the field's ability to power trials with large enough sample sizes to conduct more than one or two confirmatory tests. Some speakers suggested that sample sizes might need to be increased almost fourfold to have appropriate power to detect effects. Since funding, resource, and time constraints naturally restrict the size of most trials, discussions occurred around how to prioritize subgroup tests to conduct meaningful tests with sufficient power to detect effect. One proposal raised at the meeting suggested weighing three factors in determining which subgroup tests to conduct: theoretical rationale, empirical support, and policy importance (Bloom and Michalopoulos 2012). The consensus was that most trials need to be selective of the types and number of confirmatory tests conducted. There was also consensus that results from post hoc analyses should always be contextualized as exploratory. One presenter noted that on occasion, post hoc subgroup analyses may not be as problematic; for example, if a post hoc subgroup naturally contains 50 % or more of the total sample. However, caution is still warranted in reporting and interpreting the results. Finally, a lack of corrections for multiple comparisons is a pervasive problem. The discussion highlighted the work of Schochet (2008) as a good reference for understanding and appropriately accounting for this problem.

Papers Focused on Study Design in This Issue Wang and Ware (2012) present a summary of sound methodology for conducting subgroup analysis. The paper recommends best practices of conducting tests of heterogeneity, choices of metrics for the subgroup and outcome variables, accounting for multiplicity, and the cautions involved in conducting post hoc analyses. The paper by Farrell et al. (2012) provides an illustration of how different levels of an ecological model can be used to explore subgroup effects from individual variation to school- or community-level variation. Subgroup impacts at one contextual level can have the potential to moderate the impacts of programs found in other trials to be effective. Farrell et al. highlight the challenges around exploring these alternate variables and note how authors can sometimes be less than thoughtful in their analysis of subgroups. They identify a number of common mistakes, including a lack of a theoretical rationale for the inclusion of contextual subgroups, vulnerability to the challenges of multiple comparisons when multiple contextual levels could be included, and having sufficient power to detect effects. Rothman (2012)

provides a strong argument for the benefits of pre-specifying mediators and moderators to guide researchers to power the study and carefully measure those variables to move from exploratory to confirmatory analysis.

Papers Focused on Analysis in This Issue Four papers in the Special Issue present specific analytic techniques for subgroups. In response to the noted challenges around the power to detect effects in trials interested in conducting subgroup analysis, three papers specifically highlight methods that have the potential to maximize the power available to explore subgroup differences. First, Borenstein and Higgins (2012) recommend the use of meta-analysis to provide context and the ability to see patterns of effects for subpopulations or conditions across trials. One advantage of meta-analysis is that it allows conclusions to be drawn for subgroups across a set of studies, providing a richer context for findings than can be achieved with just one study. In a related paper, Brown et al. (2012) discuss how conducting parallel analyses on multiple datasets can increase the ability to detect moderation impacts. One of the advantages of parallel analysis is the ability to explore moderation of impacts both within and across trials. Finally, to mitigate some of the challenges associated with a lack of power and potential bias from multiple comparisons, Lanza and Rhoades (2012) explore latent class analysis as a method for exploring underlying classifications of participants along multiple dimensions. A potential application of Lanza and Rhoades' approach arises in the context of the current home visitation programs being implemented across states and territories. In this effort, localities are asked to implement evidence-based programs in their communities and asked to serve communities with a number of risk factors. Applying a latent class analysis might allow for the investigation of differential impacts of home visiting across a set of baseline characteristics that are more complex, but also potentially more relevant, for program administrators.

The work of Almirall et al. (2012) provides an entirely different view of subgroup analysis. This paper discusses how causal effect moderation can be examined in time-varying settings and why this requires analytical methods other than traditional regression. The paper is unique compared to the others based on the authors' perspective on the relationships between subgroups, moderators of treatment effects, and treatment impact. Almirall et al. assert that subgroups and moderators (such as motivation to seek treatment or current substance use) can have significant variation over time which must be taken into consideration. In addition to variation over time, there may also be variation induced by treatment exposure and an interaction between exposure, impact, and subgroup definition that ought to be recognized and analyzed. The causal inference approaches outlined in the paper are a promising new tool for addressing other complexities of subgroup analysis, such as adjustment for confounding variables.

Reporting and Interpreting

Wang et al. (2007) conducted a review of articles published in the *New England Journal of Medicine* during 1 year and found that a large percentage of these articles had gaps or incorrect presentation of subgroup analyses. The gap in the published literature around subgroup analysis included: (1) authors not providing enough detail in the manuscript to determine whether the analyses were pre-specified or post hoc and (2) lack of specification of the analytic technique used for presenting subgroup findings. Though it is unknown whether a review of published prevention and intervention research would come to the same conclusions, it is reasonable to believe that the state of the field in prevention science and intervention research is similar.

Discussions at the meeting scratched the surface around issues of reporting and interpretation. Howard Bloom led an open conversation with participants around reporting and interpreting subgroup findings. The discussion highlighted the challenges in these tasks but provided few definitive recommendations. There was no consensus around a number of critical issues including when conclusions about program effectiveness can be drawn on exploratory findings or how to interpret subgroup findings in the context of main effects that are not aligned with subgroup impacts. Additional methodological work is needed in these areas.

Participants agreed that in order for advances in the analysis of subgroups to continue, journal editors, grant reviewers, academic associations, and peer reviewers need to pay greater attention to the issues around subgroups and to apply the tools of scientific investigation, such as clearly stating hypotheses and including sufficient information necessary to make judgments about findings, across a variety of disciplines. Given the increasing importance in utilizing subgroup findings, it is important that the field begins to carefully consider planning studies with appropriate power to detect effects, clearly specifying exploratory versus confirmatory subgroups and discussing them in the conclusions with the appropriate caveats.

Papers Focused on Reporting Subgroup Analysis in This Issue This Special Issue of *Prevention Science* includes one paper that explicitly focused on issues around reporting and interpreting subgroup findings (see Bloom and Michalopoulos 2012); however, many of the papers touch on the topic throughout. Farrell et al. (2012) noted a number of factors related to reporting subgroup analyses in the series of papers they examined. These included: authors drawing conclusions of homogeneity of treatment effects when the variable in the sample had restricted range; the lack of inclusion of confidence intervals in the presentation of effect sizes; and only one paper in the series examined stood out for its explicit contextualization of results given the multiple comparison problems. The Bloom and Michalopoulos paper is important in its

careful consideration of how subgroup results are reported and interpreted. The authors propose guidelines to consider to ensure that the subgroup effects reported and those potentially influencing decision-making around program impact and implementation be done in the most defensible manner. Bloom and Michalopoulos propose this approach to reduce the potential for emphasizing spurious effects and the potential consequences for inappropriate interpretation in policy-relevant research. Rothman (2012) argues that investigators should be motivated to frame theoretically grounded research questions that include mediators or moderators and appropriately design and power studies to move those analyses to a confirmatory framework.

Application and Implications Across Prevention Science and Intervention Research

Finally, the commentary by Massetti and Haegerich (2012) provides a useful perspective across the issues of subgroup analysis through the lens of an application to violence prevention. They describe the substantial interest in the field for knowing what works and for whom and how the collection of papers in this Special Issue can provide for a better understanding to the field across a number of contexts and levels. The authors make a strong argument that the issue and importance of subgroup analysis are not just an esoteric or methodological concern, but rather are of critical importance for what actually occurs in the practice of public health and violence prevention. Through the specific example of violence prevention, the paper illuminates a number of insights that are prevalent across the fields of prevention science and intervention research,

Conclusion

It is the goal of this Special Issue to present multiple perspectives on subgroup analysis and continue the discussion that was begun in the interagency meeting. We hope that the Special Issue begins to address some of the steps for the future that emerged from the meeting. These include the need to continue to develop and refine methods for analyzing subgroups. There needs to be continued emphasis on the common pitfalls in analyzing and reporting subgroups for prevention scientists including students, researchers, peer reviewers, and editors. If the goal of prevention research is to conduct applied, policy-relevant research that has impacts on the target population, we need to ensure the conclusions drawn about program effectiveness are done so appropriately.

Acknowledgments We would like to thank all of the authors and reviewers who have contributed to this Special Issue. In addition, we would like to recognize our co-funders: the Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services; the Division of Violence Prevention, Centers for Disease Control and Prevention; the Institute for Education Sciences; the National Institute on Drug Abuse; and the National Institute of Mental Health. Videos from the meeting are available at: <http://www.researchconnections.org/childcare/datamethods/interagency-meeting.jsp>.

References

- Almirall, D., McCaffrey, D. F., Ramchand, R., & Murphy, S. A. (2012). Subgroups analysis when treatment and moderators are time-varying. *Prevention Science*. doi:10.1007/s11121-011-0208-7.
- Bloom, H. S., & Michalopoulos, C. (2012). When is the story in the subgroups? *Prevention Science*. doi:10.1007/s11121-010-0198-x.
- Borenstein, M., & Higgins, J. P. T. (2012). Meta-analysis and subgroups. *Prevention Science*.
- Brown, C. H., Sloboda, Z., Faggiano, F., Teasdale, B., Keller, F., Burkhart, G., et al. (2012). Methods for synthesizing findings on moderation effects across multiple randomized trials. *Prevention Science*. doi:10.1007/s11121-011-0207-8.
- Farrell, A. D., Henry, D. B., & Bettencourt, A. (2012). Methodological challenges examining subgroup differences: Examples from universal school-based youth violence prevention trials. *Prevention Science*. doi:10.1007/s11121-011-0200-2.
- Federal Coordinating Council for Comparative Effectiveness Research. (2009). *Report to the President and Congress*. Available at: <http://www.hhs.gov/recovery/programs/cer/ceannualrpt.pdf>. Accessed 22 Jan 2011.
- Lanza, S. T., & Rhoades, B. L. (2012). Latent class analysis: An alternative perspective on subgroup analysis in prevention and treatment. *Prevention Science*. doi:10.1007/s11121-011-0201-1.
- Massetti, G. M., & Haegerich, T. M. (2012) Commentary on subgroup analysis in intervention research: Opportunities for the public health approach to violence prevention. *Prevention Science*. doi:10.1007/s11121-012-0331-0.
- Rothman, A. (2012). Exploring connections between moderators and mediators: Commentary on subgroup analyses in intervention research. *Prevention Science*. doi:10.1007/s11121-012-0333-y.
- Rothwell, P. M. (2005). Subgroup analysis in randomized controlled trials: Importance, indication and interpretation. *Lancet*, 365, 176–186.
- Schochet, P. Z. (2008). *Technical methods report: Guidelines for multiple testing in impact evaluations*. NCEE 2008-4018. National Center for Education Evaluation and Regional Assistance. Available at: <http://ies.ed.gov/ncee/pubs/>. Accessed 10 Jan 2011.
- Wang R. (2010) Overview of subgroup analyses in medical and public health research. *Presentation at the Interagency meeting on subgroup analyses*, Washington DC.
- Wang, R., & Ware, J. H. (2012). Detecting moderator effects using subgroup analyses. *Prevention Science*. doi:10.1007/s11121-011-0221-x.
- Wang, R., Lagakos, S. W., Ware, J. H., Hunter, D. J., & Drazen, J. M. (2007). Statistics in medicine—Reporting of subgroup analyses in clinical trials. *The New England Journal of Medicine*, 357, 2189–2194.