Chapter 8
Studying Family Transitions from a Systems Perspective: The Role of Biomarkers

Carolyn Tucker Halpern, Kathleen Mullan Harris and Eric A. Whitsel

Family processes—and the individual developmental transitions embedded within them—reflect complex interactions among multilevel factors including genetic, hormonal, and neural influences; higher-level cognitive, experiential, and behavioral processes; as well as the physical, social, and cultural environments in which they operate. Family influence and process are by definition, partly biological. Genetic endowments and the shared environment of the family confer both sensitivities and vulnerabilities to family members, as well as developmental opportunities to foster resilience. To understand the complex intersections of these diverse factors, a multilevel dynamic systems approach is needed. By multilevel, we do not intend a statistical definition but rather a conceptualization that encompasses the broad range of factors noted above and their coactional contributions to biological and social processes (Gottlieb and Halpern 2002). Here, we discuss biological factors as contributors to the family system and biomarker collection in large scale studies framed through our experiences in the National Longitudinal Study of Adolescent Health (Add Health). We illustrate field, laboratory, and data dissemination challenges for a selection of common biomarkers, leading to best practice recommendations. We also present illustrative findings from research integrating biomarker, social and behavioral data that provide novel insights into social and behavioral phenomena.

". . . . sociological problems are better understood when a biosocial theory is brought to bear."
J. Richard Udry, 1988, page 717

C. T. Halpern (✉)
Department of Maternal & Child Health, University of North Carolina
Chapel Hill, Chapel Hill, NC, USA
e-mail: carolyn_halpern@unc.edu

K. M. Harris
Department of Sociology, University of North Carolina Chapel Hill,
Chapel Hill, NC, USA
e-mail: kathie_harris@unc.edu

E. A. Whitsel
Departments of Epidemiology and Medicine in the Gillings School of Global Public Health
and School of Medicine, University of North Carolina, Chapel Hill, NC, USA
e-mail: eric_whitsel@unc.edu

S. M. McHale et al. (eds.), *Emerging Methods in Family Research,*
National Symposium on Family Issues 4, DOI 10.1007/978-3-319-01562-0_8,

Finally, we offer a rationale for incorporating biomarkers into social science research, despite the challenges, and highlight future possibilities for expanded multilevel research capitalizing on intergenerational study designs.

Why Consider Including Biomarkers?

Family process is inextricably linked to health, and the study of health is fundamentally a study of biological processes and outcomes. Social science and health surveys have traditionally relied on self-report to identify health outcomes. For example, at several waves of data collection, Add Health presented a list of chronic conditions to respondents (Rs) and asked, "Has a doctor, nurse or other health care provider ever told you that you have or had. . . .(cancer or lymphoma or leukemia; high blood cholesterol or lipids; high blood pressure or hypertension; high blood sugar or diabetes; . . .)." Rs self-reported "yes" or "no;" affirmative responses were followed by a question about age of onset.

Although self-reported health measures vary in quality, they generally underestimate health risks that go undetected or for which symptoms do not appear early or consistently; this is especially true among young, otherwise healthy populations (Kehoe et al. 1994). Moreover, socioeconomic status, race, ethnicity, nativity, and language proficiency influence knowledge of health, symptoms, and disease; access to diagnostic services; and understanding of health questionnaires. Consequently, the accuracy of self-reported health survey data can vary in ways unrelated to pathophysiological mechanisms of disease. Objective measures are therefore the gold standard for reliable and valid measurement of health.

In a survey field setting, objective health measures are those derived from the collection of biospecimens or through physical measurement by trained and certified personnel (e.g., interviewers or phlebotomists) following standardized protocols. We refer to these objective measures as "biomarkers" or "biological markers" of (ab)normal biological states resulting from underlying (patho)physiological processes. Biomarkers were once limited to clinical patient samples, or community-based epidemiological studies. However, in the early 1990s several international and/or aging studies began collecting biomarker data to strengthen and complement self-reported data on health and aging (Finch et al. 2001; Weinstein et al. 2008). For example, the Indonesian Family Life Survey (IFLS), the National Study of Midlife Development in the United States (MIDUS), and the Social Environment and Biomarkers of Aging Study (SEBAS, i.e. Taiwan Biomarker Project) were forerunners in this regard. However, these studies focused on later life stages when illness is typically manifest. Add Health is one of the first large, longitudinal, nationally representative studies to collect biomarker data from a young adult population, when the data may be most informative about pre-disease pathways and cumulative physiological risk. Add Health also has the strengths of extensive interpersonal and physical contextual data, which allow for multilevel analyses of health trajectories from youth into adulthood.

Biomarker Feasibility for Large Scale Studies: Add Health as an Example

Below we describe the theoretical foundation and unique design features of Add Health that facilitate a multilevel systems approach to family process research, followed by our theoretical choices for biological data collection at Add Health, Wave IV.

Theoretical Foundation

The scientific purpose of Add Health is to study developmental and health trajectories across the life course from early adolescence into adulthood using an integrative approach that combines social, behavioral, and biomedical sciences in its research objectives, design, data collection, and analysis. We developed an Integrative Life Course Theoretical Model that specifies three broad conceptual domains of longitudinal and reciprocal influences in trajectories of health and human development: context, behavior, and biology. Here we focus on the biology domain (described below).

Add Health Design

Add Health is an ongoing longitudinal study of a nationally representative sample of more than 20,000 US adolescents who were in grades 7–12 in 1994–95 and have been followed for 15 years through adolescence and the transition to adulthood. In-home interviews occurred in 1995 (Wave I), 1996 (Wave II), 2001–02 (Wave III), and 2008–09 (Wave IV) when Rs were aged 24–32 years. Add Health used a school-based design to obtain direct independent measures of the multiple contexts of young people's lives, including school, peer network, dyadic relationships (friends, romantic and sexual partners), family, neighborhood, and community. Add Health oversampled by ethnicity, physical disability, school (i.e., all enrolled students in 16 schools were selected for in-home interviews), and biological relatedness to increase the diversity of potential research. For example, it sampled > 3,000 pairs of individuals with varying biological relatedness (aka the "genetic sample:" identical/fraternal twins, full & half siblings, cousins, and biologically unrelated youth raised in the same household) to facilitate genetic and environmental research. Thus, Add Health is nationally representative of young people in every race, ethnic, immigrant, geographic, and socioeconomic subgroup now living in all 50 United States. (See Harris 2011 for details.)

Add Health's multilevel and longitudinal data provide unprecedented opportunities to characterize the social, physical, and health environments in the Context Domain of the Integrative Life Course Theoretical Model. Almost 8,000 data elements on the social and physical environment at multiple spatial levels are available

across waves (e.g., poverty rates, sexually transmitted infection (STI) prevalence, welfare policies, cigarette taxes, and the proximity and number of physical features such as parks and recreation centers). Add Health also contains extensive longitudinal information for the Behavior Domain measured at multiple levels (i.e., family, school, peer group, relationship dyad, neighborhood, and community), including life histories of sexual and risk behavior, civic engagement, fertility, cohabitation, marriage, and work.

Add Health's design features furthermore enabled measurement in the Biology Domain, including the genetic sample and longitudinal self-reports of physical development, general health, chronic illness, physical activity, mental health, and disability. Add Health has always collected objective measures of health as well, including anthropometrics to identify overweight and obesity, biospecimens to measure STI and HIV status, genetic markers, and more recently, biomarkers of cardiovascular health, metabolic processes, immune function, inflammation, and medication use.

Thus, Add Health provides unique opportunities to study how environments and behaviors are linked to biological and family processes in their influence on health and well-being across the life course. Further, the longitudinal design and diverse sampling strategy enable the examination of bidirectional associations between family characteristics and within-individual change across the life course to support dynamic multilevel system analysis for multiple segments of the US population.

Biology Domain

The choice of biological data in Add Health was driven by scientific knowledge of reasonably prevalent health conditions at a given developmental stage of the Add Health cohort, especially data with implications for future health, the role of specific biological processes in causation, and the ability of specific measures to characterize these processes. Within these scientific criteria, choices were constrained by the feasibility of methods used to obtain valid and reliable physical measurements and biological specimens in a large non-clinical field setting, and measurement strategies and assay techniques used to capture biological phenomena of interest in each life stage. Within these constraints, we chose noninvasive, innovative, cost-efficient, and practical methods for collecting biomarkers appropriate for population research (McDade et al. 2007).

Thus, we measured height and weight in adolescence, during the transition to adulthood, and in adulthood to map the obesity epidemic in Add Health (see Harris 2010; Gordon-Larsen 2010). At Wave III (ages 18–26), the Add Health cohort was at relatively high risk of contracting STIs, so we collected urine and saliva to test for STIs and HIV (see Miller et al. 2004; Morris et al. 2006). We expanded biomarker collection at Wave IV (ages 24–32), when the cohort was settling into adulthood and diverging along pre-disease pathways. We identified obesity, health risk behavior, and stress as the leading health concerns at Wave IV, and collected biological data to provide information about the associated consequences.

Such consequences include hypertension, diabetes, hyperlipidemia, inflammation, immune dysfunction, and cardiovascular disease (Ferraro and Kelley-Moore 2003; McEwen 1998). Therefore, blood pressure and pulse were measured during the Wave IV interview. Metabolic and inflammatory processes associated with obesity also pose significant risks for renovascular, cardiovascular, and cerebrovascular disease (Blake and Ridker 2002; Khaw et al. 2001). Accordingly, we collected capillary whole blood spots from a finger prick onto filter paper, dried, punched and then assayed them for glucose, hemoglobin A1c (HbA1c), lipids (total cholesterol (TC), high-density lipoprotein cholesterol (HDL-C), triglycerides (TG)), and a marker of inflammation, high sensitivity C-reactive protein (hsCRP)—important markers of current health status and future risk for diabetes, kidney disease, peripheral artery disease, heart disease, and stroke (Fagot-Campagna et al. 2001; Nissen et al. 2005). Moreover, the immune system and the hypothalamic-pituitary-adrenal (HPA) axis may be compromised by stressful events, adverse social environments, or health risk behaviors, which in turn can lead to infectious, autoimmune and cardiovascular disease (Herbert and Cohen 1993). We collected saliva to measure a stress hormone, cortisol (pretest only), and assayed dried blood spots for Epstein Barr viral capsid antigen IgG (EBV)—an indirect measure of chronic stress (Glaser et al. 1991). Because genes influence health, behavior, and the moderation of contextual and behavioral effects as they relate to future health, we also collected DNA from all Rs at Wave IV. In combination with the longitudinal contextual and behavioral data, these biological measures inform inter-relationships between biological processes, social, and behavioral trajectories.

Cost of Biomarkers

The financial cost of conducting research involving biomarkers reflects the necessity of many resources. Equipment and supplies are needed for measuring, collecting, storing, shipping, and processing. For example, scales with acceptable levels of accuracy and upper range bounds–and possibly portability, depending on study design–are needed to measure weight. Tubes and syringes are needed for venous blood collection, and materials to keep specimens cold may be needed for packaging and shipment. For some specimens, such as urine or blood, different supplies and/or protocols may be required, depending on the biomarker of interest. Some specimens may be collected relatively easily by Rs themselves (e.g., saliva for buccal cell DNA extraction). However, other types require either specialized field personnel (e.g., nurses or phlebotomists to collect venous blood) or lay interviewers with extensive training (e.g., to conduct skinfold measurements or collect blood spots), which increases costs of measurement. Collection costs also reflect whether specimens are collected at home, at clinic visits, or via mail. Most biomarkers entail laboratory costs to receive, store, and process specimens and to measure the marker of interest. These costs can vary substantially depending on the biomarker and the technology used. In general, the larger the study sample size, the greater the cost. For some biomarkers,

such as genetic markers or established hormone assays, the cost typically equals a fixed cost of genotyping times the number of Rs who provided a specimen; for other biomarkers, such as blood pressure or weight, there may be economies of scale if personnel can use the same monitor or scale to obtain measures from multiple Rs in the field. However, costs of other biomarkers can vary substantially, depending partly on whether assays are well-established (e.g., lipids in plasma) or developmental (e.g., lipids in dried blood spots).

Challenges in the Field

Even with extensive research into optimal materials and protocols, Add Health encountered challenges in the field. For example, to avoid keypunch error, digit preference, and data fabrication, we planned to automatically download electronic data from blood pressure monitors. However, electronic downloading was unreliable during the pretest so blood pressures were manually entered in duplicate on laptops. Such trade-offs in performance may be related to the reliance on lower cost equipment required in large scale studies. Clinical monitors from which blood pressure data are reliably downloaded can be ten to twenty times more expensive than the blood pressure monitor used at Add Health Wave IV. Assuming 350 field interviewers, choosing between the two monitor types reflected a difference in total cost of approximately $ 200,000–400,000. Despite such a difference, the monitor used at Wave IV (Model BP3MC1-PC-IB; MicroLife USA, Inc; Dunedin, FL) was approved by the British Hypertension Society and advertised as having an accuracy of 3 mm Hg. Moreover, it performed extremely well in field work (Nguyen et al. 2011).

With non-medical field staff there may also be challenges related to the "ick" factor. Collection of saliva and urine was pre-tested at Wave II of Add Health, but interviewers performed poorly in the field. They were uncomfortable collecting specimens, were afraid to handle them, and may have sabotaged the effort. Given the poor pretest performance, urine collection was abandoned for the main study.

Some specimen collection requires extensive training and adherence to complex protocols. For example, the Wave IV protocol for finger prick whole blood collection required the interviewer to: clean the R's middle or ring finger with an alcohol prep pad, apply a tourniquet to the arm, prick the finger with a lancet, and wipe away the first drop of blood. Next, the interviewer was to drop (ideally) seven blood spots onto a special collection card, ensuring that the finger did not touch the card and that the blood spot saturated the collection circle on the card. Samples were then to be air dried over desiccant for three hours, packaged, and shipped via Federal Express to the lab on the same day as collection. Interviewers had to ensure that they were sending the specimen to the correct lab (Add Health used several labs for different specimens) with the correct R biospecimen ID (different from field interview IDs). The protocol required multiple supplies (plastic gloves, band-aids, gauze, alcohol prep pads, a tourniquet, Lancets, 7-spot collection cards (Whatman 903® *Protein Saver*, Whatman Inc., Piscataway, NJ), Chux pads, and a biohazard container), all of which interviewers carried in the field along with other supplies and a laptop computer.

Seemingly simple saliva collection can also be complicated. For example, although salivary concentrations of most steroid hormones are not affected by saliva flow rate, or mouth collection site, salivary concentrations of some proteins such as alpha-amylase and secretory IgA (SIgA) vary by site (Beltzer et al. 2010; Veerman et al. 1996). Concentrations of some analytes (e.g., SIgA, dehydroepiandrosterone sulfate) vary inversely with saliva flow rates (Kugler et al. 1992; Vining et al. 1983). Further, use of stimulants to increase saliva flow may interfere with assay performance, a concern when Rs self-collect saliva.

Other challenges to specimen integrity over which investigators have little control are exogenous, historical, or cultural events that impact normal field operations. For example, Wave III was in the field during 9/11 (i.e., terrorist attacks on September 11, 2001). This tragic event and its aftermath disrupted the scheduling and keeping of interview appointments (especially on the east coast), but the more enduring effect was severe delays in shipping urine and saliva to labs, resulting in significant biomarker loss. Either specimens were lost in transit, or they arrived at the lab beyond the 48-hour window required for a valid test according to FDA regulations.

Complexity of Protocols for Participants

Stress experiences, and their short- and long-term health implications, were themes of the Wave IV Add Health program project. In the pretest we evaluated a protocol to collect saliva for cortisol measurement. Cortisol is an endogenous corticosteroid that affects multiple physiological systems; it has been implicated in multiple physical and psychological illnesses. Circulating cortisol concentrations can be influenced by stress, but it also has a strong diurnal profile, rising shortly after awakening and then falling throughout the day. Thus, multiple samples are needed, and adherence to the protocol for timing of sample collection is critical. Of particular interest in the context of field collection is the cortisol response to awakening (CRA; a large, rapid increase within a 20–30-minute period after awakening), thought to be a reliable indicator of the acute reactivity of the HPA axis.

Based on available literature, we developed and tested a low burden and affordable protocol to maximize consent and protocol adherence, given that specimen collection would be unsupervised on a day after the in-home interview. We requested that the R collect three specimens on a single day: upon awakening, 30 min later, and just before bed. Rs completed a brief checklist linked to each specimen, noting time of collection, stressful daily events, and recent consumption of food, beverages, or drugs before specimen collection, all needed for proper interpretation of assay results.

Analyses of pretest specimen receipt and protocol adherence were illuminating. Virtually everyone agreed to provide samples, but about a quarter of self-reported collection times were missing for the 76 % of Rs who returned samples, suggesting large amounts of collection time data would be missing in the main field work. Although

there was some suggestion that higher incentives ($ 40) might improve specimen return, the increase was cost prohibitive. Given our final sample size of 15,701, applying the 97 % consent and 81 % return rates yields a cost close to $ 494,000 in *incentives alone*. Further, through embedded experiments we determined that collection protocol adherence was poor, thus explaining the awakening sample's intra-class correlation (ICC) of 0.06. Given likely poor data quality, we eliminated saliva collection for cortisol measurement from the main field work (Halpern et al. 2012).

Challenges in the Laboratory

Social scientists may mistakenly assume that, unlike the inherent error and potential bias in survey or observational data, biological measurement is cut and dry, and accuracy is a given. Hormone molecules get "counted," gene allele variations are clear, and all procedures are standardized and scrupulously followed by infallible laboratory technicians. Alas, this is simply not true. Below we illustrate some of the potential pitfalls that can occur.

Technology Change, Time, and Processing Delays

Equipment, techniques, reagents, etc. evolve. Whereas DNA and hormones were once measured only in blood from venous draws, many are now obtained via buccal cells and saliva. However, data derived from new technology must be appropriately scrutinized. For example, when salivary radio-immunoassays (RIAs) for steroids were introduced, many assumed these early assays had the same accuracy and reliability as RIAs using blood samples. In practice, this was not necessarily true. In one experiment, correlations of salivary testosterone (T) measurement from the same specimens across labs ranged from 0.05 to 0.48 (Halpern and Udry 1992). Similarly, correlations of T values between plasma and saliva samples collected from the same person at the same point in time ranged across labs from −0.009 to 0.86; even within a lab the correlation was only 0.52 (Halpern and Udry 1992). Perhaps such differences simply introduce noise and bias biological associations with behavior toward the null; however, this is not necessarily true. To illustrate, correlations between male adolescents' reports of "frequency of thinking about sex" and salivary T ranged from an insignificant 0.12 to a statistically significant 0.40, depending on lab measurement used (Halpern and Udry 1992).

There are other challenges related to state-of-the-art laboratory methods. For example, we assayed lipids in dried whole blood spots at Wave IV. Originally, the lipids were colorimetrically assayed using procedures that measure change in color (optical density) reflective of increases in plasma lipid concentrations. However, the colorimetric assays were replaced with fluorimetric alternatives during specimen

collection and assay. The anticipated advantage (e.g., validity) of the fluorimetric assays—which involve ultraviolet excitation of and spectroscopic measurement of light emitted from fluorochromes—led to their adoption. That said, subsequent examination of the temporal variation, short-term reliability, inter-convertibility, and internal/external validity of the two assays did not consistently substantiate the purported advantage of the fluorimetric assays.[1]

Factors that delay assaying or genotyping samples (e.g., overwhelmingly large numbers) make consistency of laboratory procedures and uniformity of supplies/reagents an issue. Indeed, assay performance may vary over time and among supply/reagent lots while measurement tools such as genotyping platforms may be retired. These possibilities can be difficult to avoid if field work is prolonged, or if a lab has inadequate automation to accommodate incoming sample volume.

Art and Ambiguity

Genotyping is another domain challenged by the state-of-the-art of measurement. There are issues related to the genotyping chemistry and software used to differentiate true alleles from various sources of background error. (An allele is one of two or more forms of a gene or genetic locus.) Variation in allele calls, and therefore the labeling of genotypes, is a good example. Calling is based on a pattern of peaks, points, or bands on a computer generated "image" of an electropherogram or gel. Although most laboratories use software programs for sizing (repeat polymorphisms) and clustering (single nucleotide polymorphisms or SNPs), some human judgment is still involved. To assist judgment software packages include detailed sections on manual editing of automated allele calls. For sizing polymorphisms, the two most common problems are "stutter bands" (strong peaks that are one or more repeat units smaller than the actual allele) and allele dropout (large alleles poorly amplified and missed). The editing process involves two independent reviewers of peak profiles who may offer discordant calls that must be reconciled. If not reconciled, the genotype will be judged to be "missing" (Hill et al. 2004).

Virtually all SNP genotyping, from a single Taqman assay to multi-million-SNP chips, is based on assigning genotypes to clusters of points visualized on a Cartesian plane. Ideally, one homozygote forms a tight cluster on the X axis, the opposite homozygote clusters on the Y axis, and the heterozygotes cluster neatly in between. Manual editing is required when the software cannot define clusters accurately or cannot assign genotypes to points not clearly in one of the three clusters. For relatively small arrays this is feasible, but for large arrays it is impossible. Multiple software applications have been developed that claim to call genotypes more accurately than the manufacturer-supplied software (e.g., Browning and Yu 2009), but the implication of these improvements is clear: The genotypes for any set of samples genotyped on very large arrays will vary depending on the software package used to define them.

[1] However, a separate inter-conversion process for HbA1c (necessitated by a post-pretest lab closing) was successful (see Table 8.1).

Table 8.1 Reliability of Add Health Wave IV Biomarkers Based on IIV Study (Details of Add Health Wave IV equipment and measurement protocols are available at http://www.cpc.unc.edu/projects/addhealth/data/guides)

Type	Measure	ICC (95 % CI)[a]
Anthropometric	Weight	1.00 (1.00–1.00)
	Height	0.98 (0.98–0.99)
	BMI	0.99 (0.99–1.00)
	Waist	0.98 (0.97–0.99)
Cardiovascular	SBP	0.81 (0.74–0.88)
	DBP	0.68 (0.57–0.79)
	PR	0.47 (0.31–0.63)
Metabolic	HbA1C	0.97 (0.96–0.98)
	Glucose (non-fasting)	0.39 (0.21–0.58)
	TG	0.71 (0.60–0.81)
	TC	0.40 (0.22–0.58)
	HDL-C	0.31 (0.12–0.51)
Inflammatory	hsCRP	0.70 (0.59–0.81)
Immune	EBV	0.97 (0.96–0.98)

[a] *ICC* (*95 % CI*) intra-class correlation coefficient, 95 % confidence interval

Challenges of Scale

In a large study with concentrated data collection, questions of scale become critical. Factors that delay processing, assaying, or genotyping samples (e.g., overwhelmingly large numbers, inadequate automation) make consistency of laboratory procedures and uniformity of supplies/reagents an issue. Assay performance may vary over time and among supply/reagent lots; measurement tools such as genotyping platforms may be retired. An inadequately staffed or automated lab will be hard pressed to process the *hundreds* of specimens that will arrive *daily*. Processing issues may result in labeling/storage errors, contamination of samples, errors in data entry, or exceeding the time window for storage outside of a −70 C freezer. Further, depending on the size and number of specimen aliquots, simple storage space may be an issue, both for the lab and for the investigator, if the specimens are to be archived.

Quality Control

One unique aspect of the Add Health Wave IV design was to include an intra-individual variation (IIV) study. Inclusion of an IIV study involved repeated collection of biomarkers on the same individuals over a short time interval to estimate their reliability. Approximately 100 IIV Rs were interviewed twice, one to two weeks apart. The first visit included a full interview and full set of biomarkers. The second visit included an abbreviated interview, mainly capturing information needed for biomarker interpretation, and a full set of biomarkers. Labs and technicians were blinded to IIV specimens. We computed ICCs as measures of reliability, then used them to monitor biomarker data quality and correct for measurement error.

Table 8.1 shows reliability information derived from the IIV study. The ICCs of course vary depending on the stability and susceptibility of a given biomarker to behavioral and/or contextual change. Thus, height, weight, and EBV, for example, have excellent test-retest reliability, while more labile measures such as pulse rate were lower. Comparisons with external standards were conducted to monitor biomarker validity (see http://www.cpc.unc.edu/projects/addhealth/data/guides/WaveIV).

Biomarker Consent and Compliance

Investigators may wonder what level of biomarker consent to expect, and whether requests for specimen collection will affect consent rates for study participation. In Add Health, the longitudinal design has reinforced consistent rapport with Rs, who trust the rigorous security system we maintain to ensure our original pledge of confidentiality to them. An average of 99 % of Rs agreed to height and weight at all waves (including waist circumference at Wave IV). Compliance for Wave IV blood pressure readings was also 99 %. Consent rates at Wave III were also high: 92 % provided urine for STI testing; 95 % provided saliva for HIV testing; and 83 % provided buccal cell saliva for DNA testing. The lower compliance for DNA may be due to not providing a separate monetary incentive for DNA at Wave III as was done for urine and saliva.

At Wave IV we obtained signed consent separately for the collection of salivary buccal cell DNA, blood spots, and salivary cortisol (pre-test only) at the beginning of the interview when we obtained signed consent for the Wave IV survey administration. This allowed Rs to change their minds about biomarker consent at the end of the survey and before biomarker collection. (A few Rs did, in both directions). We also used a two-tiered consent process for (1) currently planned Program Project research and (2) future research "related to long term health." The latter provided a biospecimen archive for future testing. Consent to biospecimen collection was uniformly high: 96 % for DNA and 95 % for blood spots as part of the planned program project research. Consent to archival for future analysis also was high: 78 % for DNA and 80 % for blood spots. Black and Asian Rs were somewhat less likely to consent to biospecimen archival than Hispanic and white Rs.

Biospecimen Archive

A significant scientific advantage of collecting biospecimens in a large-scale survey is the potential for assembling and maintaining an archive or "bio-repository." The Add Health Biospecimen Archive includes urine (Wave III), DNA (Wave III and Wave IV), and dried blood spots (Wave IV). Planning for an archive requires additional tasks and costs including (1) a separate consent process for archival, (2) budgeting for long-term storage and additional shipping costs and logistics for returning samples from laboratories, (3) greater field and training demands to procure enough specimen

to archive, (4) development of dissemination policies for sharing specimens, and (5) plans for the archive if/when the study ends. However, the scientific benefits of maintaining an archive far outweigh the costs. The archive allows study investigators and outside investigators, through an ancillary study mechanism, to capitalize on rapidly changing technology to produce additional biomarker information that can be linked with the existing survey, geographic, and biological longitudinal data.

In the last 10 years new methods have been rapidly developed to analyze saliva and dried blood spot samples for an increasing array of biomarkers (McDade et al. 2007) and to explore a widening range of specimens collected in field settings (e.g., hair, fingernails, vaginal swabs). During the last decade enormous advances in genotyping technology, including chip arrays that accommodate in excess of 2 million genetic markers, have been made. In pace with technological change has been a parallel decline in cost and an explosion of new research and knowledge. As both technology and scientific discovery advance, an archive offers opportunities to test new hypotheses in both cross-sectional and longitudinal studies.

Dissemination Challenges

A biospecimen archive raises numerous dissemination issues including data sensitivity (e.g., diabetes markers, STI results, genetic data), deductive disclosure risks when sensitive biomarker data are merged with survey and geographic data, security requirements for access, analysis, and publication of the data with which users must comply, and determining who should have access and under what conditions (e.g., what forms of the data will be released, what merged files are permissible, what level of review and oversight of user access is needed)[2]. Biomarker data also increase knowledge demands (and therefore demands for training and experience) on users in terms of interdisciplinary understanding, research design, and analytic methods that enable users to develop meaningful research questions, appropriately use and model the data, and properly interpret their findings. Wide access to biomarker data will undoubtedly increase the possibility that inexperienced users will misuse it or misinterpret results.

Fortunately, there is growing research and information technology expertise on how to securely release sensitive data to minimize disclosure risks. Add Health pioneered an innovative dissemination policy, later adapted by other studies, to release data using a four-tiered access plan whereby the security requirements for access and use become stricter as the data requested represent greater disclosure risks or require more monitoring of use (e.g., genetic data versus blood pressure or glucose concentrations). Providing hands-on workshops or didactic sessions on biomarker data (e.g., as done at the Add Health Users Conferences) can improve the proper use of diverse types of biomarkers.

[2] Space limitations preclude elaboration here but we also want to alert readers to the complex issues surrounding the obligation to report (especially urgent and emergent) laboratory values with established clinical utility to respondents or their designees, and the simultaneous obligation not to harm while doing so.

Recommendations for Best Practices

Our experiences suggest the following practices to optimize biomarker data quality:

- Read available literature and consult with experts to evaluate the biomarkers, and their collection and processing protocols that are best suited for your research aims and design.
- Use standardized collection protocols and pretest them exactly as they are to be implemented.
- Collect more specimen than you need and avoid storing or shipping all specimens together. Assays may need to be repeated and natural disasters can wipe out an entire archive.
- Implement uniform training, certification, and retraining of staff; use identical supply/reagent/assay types and sources; and conduct regular equipment testing and calibration.
- Never assume that your lab has caught all the problems or issues, especially when assays or techniques are relatively new, or are new to the lab. Never assume that field staff and/or Rs are following the protocol correctly. Introduce your own real-time, masked, external quality control and assurance procedures to assess reliability and validity. A measuring tape placed in the wrong location on the body, even just slightly off, can change measures dramatically.
- Be clear about how long specimens are to be retained, and under what storage conditions.

Illustrative Insights from Biomarkers

As large scale studies begin to collect biomarkers, research opportunities to integrate new biological data with social, behavioral, and environmental data in a multi-system model are being realized. Here we provide a few examples of this integrative research. MIDUS is an ongoing longitudinal aging study of adults aged 25–74 years in 1995 designed to investigate the role of behavioral, psychological, and social factors in accounting for age-related variations in health and well-being in a national sample (http://midus.wisc.edu/index.php). Integrating biological data used to measure metabolic syndrome in adulthood (i.e., waist circumference, blood pressure, lipids, blood glucose) with retrospective survey data on socioeconomic status and parents' behavior in childhood, Miller et al. (2011) report a buffering effect of maternal nurturance in the influence of childhood poverty on metabolic syndrome such that high levels of maternal nurturance offset the metabolic consequences of childhood disadvantage.

Another recent study examined how transitions to fatherhood impact testosterone (T) levels using longitudinal survey data and biological samples collected in a large sample of men participating in the Cebu Longitudinal Health and Nutrition Survey (CLHNS), a representative 1-year birth cohort study begun in the Philippines in 1983

Table 8.2 Prevalence estimates of selected health conditions using survey, biomarker and pharmacologic data, young adults ages 24–32 (2008–09) (national longitudinal study of adolescent health (Wave IV))

Health Condition	Percentage (%)
Hypertension ($N = 14{,}252$)	
Use medication	3.4
Self-reported	11.1
Use medication or self-reported	11.8
Use medication, self-reported, SBP $\geq$ 160 or DBP $\geq$ 100[a]	14.0
Use medication, self-reported, SBP $\geq$ 140 or DBP $\geq$ 90[a]	26.1
Diabetes ($N = 12{,}224$)	
Use medication	1.3
Self-reported	2.5
Use medication or self-reported	2.8
Use medication, self-reported, or glucose $\geq$ 200[b]	3.4
Use medication, self-reported, glucose $\geq$ 200, or HbA1c $\geq$ 6.5[b]	5.5

[a] Stage 2 hypertension is classified as SBP $\geq$ 160 or DBP $\geq$ 100; Stage 1 hypertension is classified as SBP $\geq$ 140 or DBP $\geq$ 90 ("The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation and Treatment of High Blood Pressure (JNC 7)." 2003. *Hypertension* 42: 1206)

[b] Random (non-fasting) glucose values $\geq$ 200 mg/dL and HbA1c values $\geq$ 6.5 % are cut-offs for classification of diabetes (American Diabetes Association. 2007. "Diagnosis and Classification of Diabetes Mellitus." *Diabetes Care* 30(S1): S 42–47)

(http://www.cpc.unc.edu/projects/cebu). Salivary T was assessed when the cohort was aged 21.5 (in 2005) and again at age 26 (in 2009). Men with the highest levels of T in 2005 were more likely to become committed partners and fathers by 2009, but those who did become fathers showed steeper drops in T compared to their single, childless counterparts (Gettler et al. 2011). Testosterone levels were lowest in men who spent the greatest amount of time caring for their children. These findings demonstrate the bi-directional relationships between family and biological processes.

As mentioned earlier, self-reports tend to underestimate the population prevalence of health conditions, especially among young people who are in a healthy life stage and for conditions that are asymptomatic. Table 8.2 shows prevalence estimates of hypertension and diabetes at Wave IV, when the Add Health cohort was in their mid-20s and early 30s. By combining self-reported, biological, and medication data, we obtain more sensitive estimates. In this young adult population, 11.1 % self-report hypertension. Combined with antihypertensive medication use, prevalence rises slightly to 11.8 %. Bringing in blood pressures (BPs) and using conventional thresholds to define hypertension according to Joint National Committee (JNC) 7 guidelines (Chobanian et al. 2003), the prevalence of stage 2 hypertension rises to 14.0 %; and stage 1 hypertension rises to more than 25 %. We see similar measurement gains in prevalence of diabetes as defined by the American Diabetes Association (ADA 2012). These data illustrate the insensitivity of self-reported health data, and related measurement error present in modeling these disease outcomes, especially among seemingly healthy young populations.

The integration of genetic and social data to understand how the environment moderates gene expression in behavioral outcomes has captured the attention of the social science research community. Guo et al. (2010) examined how the dopamine transporter gene (*DAT1*) interacts with age (or life course stage) in relation to risk behavior (delinquency, number of sex partners, substance use, and seatbelt use) from adolescence into young adulthood, using data from Waves I, II and III of Add Health. They reported a protective effect of the 9R/9R genotype in the VNTR of *DAT1* on risky behavior. However, this protective effect varied according to age/life course stage, such that genetic protection is evident when the risk behavior is illegal (e.g., alcohol use and smoking in adolescence) but vanishes when the behaviors are legal or more socially tolerated (e.g., alcohol use and smoking in adulthood). This important research demonstrates how legal, as well as social, contexts can moderate genetic associations for diverse behaviors.

A final example relates to STIs. Bruckner and Bearman (2005) examined the effectiveness of adolescent virginity pledges (i.e., a pledge to remain a virgin until married) in reducing STI rates among young adults aged 18–24 years in Add Health. Urine specimens collected at Wave III were tested for human papilloma virus, chlamydia, gonorrhea, and trichomoniasis; positive results indicated infection. Pledgers were consistently less likely to be exposed to risk factors across a wide range of indicators, but their STI prevalence did not differ from non-pledgers. Bruckner and Bearman (2005) hypothesized that pledgers were less likely to use condoms at sexual debut and less likely to be tested, and therefore diagnosed, with STIs. They concluded that any health advantages accruing from pledging do not appear to stem STI acquisition among young adults.

In an unpublished response to this research, Rector and Johnson (2005) re-examined the linkage between adolescent virginity pledging and STIs among young adults using self-reported data on diagnosis or symptoms at Wave III. Their results are opposite to Bruckner and Bearman, finding that virginity pledging predicts lower STI prevalence among young adults when STIs are measured by self-report. However, as illustrated above, self-reports fail to capture many health conditions that are better indexed by objective biological measures. These two sets of findings are a compelling example of how different public policy implications can be depending on the measures used.

Now, Tell me Again Why I Should Consider Biomarkers?

After learning about the expense, challenges, and complications inherent in biomarker collection on a large scale, one might ask again, "Why use biomarkers?" Despite the challenges, there are many advantages to using biomarkers as measures of health and family process. First, biomarker data provide information about health conditions that may be unknown to the individual. Being unaware of potential health problems is especially likely in young adulthood when, being without disease symptoms, young people assume they are healthy and are unlikely to have regular medical

check-ups. A second advantage is the opportunity to archive biospecimens for future analysis to capitalize on advances in technology, new research knowledge, and potentially declining costs for testing. This is especially relevant for DNA archives where genotyping techniques are changing at escalating speed while costs decline.

Third, repeated biomarker assessments in longitudinal studies allow for analysis of change, including disease onset and progression, and the ability to map predisease pathways. This is extremely valuable in young populations before disease is manifest, because identifying the precursors to disease will lead directly to policy and program interventions to improve health and lower disease prevalence. Fourth, biomarkers reflect different time metrics and therefore offer different types of insight into health status and biological processes. For example, some markers measure current status (e.g., STI, blood glucose), while others measure cumulative health (e.g., HIV status, diabetes) or contextually dependent change (e.g., cortisol). A fifth advantage of using biomarkers is that biomarkers are not influenced by recall bias or by individual characteristics that tend to bias self-reports of health.

As J. Richard Udry advocated throughout his long career, biosocial theory and research have already made, and will continue to make, important contributions to our knowledge of sociological problems particularly issues related to family functioning, and the future holds even greater opportunities for biosocial family research. For example, Add Health is expanding its family system design with intergenerational investigation of health and development across three generations: the parents of Add Health Rs (G1); Add Health Rs (G2); and the children of Add Health Rs (G3). The long-term goal is to collect parallel social, environmental, behavioral, biological, and genetic data on three generations to enable unprecedented multilevel dynamic systems research on the intergenerational linkages in health and behavior in family systems. We encourage other family researchers who have not considered the inclusion of biomarkers in their research to broaden their scientific reach as well.

Authors' Note We dedicate this work to J. Richard Udry, a pioneer in biosocial research, the original Director of Add Health, and a mentor to Halpern, Harris, and many others who had the good fortune to work with him.

References

American Diabetes Association. (2012). Diagnosis and classification of diabetes mellitus. *Diabetes Care, 35,* 64–71.

Beltzer, E. K., Fortunato, C. K., Guaderrama, M. M., Peckins, M. K., Garramone, B. M., & Granger, D. A. (2010). Salivary flow and alpha-amylase: Collection technique, duration, and oral fluid type. *Physiology and Behavior, 101*(2), 289–296.

Blake, G. J., & Ridker, P. M. (2002). Inflammatory biomarkers and cardiovascular risk prediction. *Journal of Internal Medicine, 252*(4), 283–294.

Browning, B. L., & Yu, Z. (2009). Simultaneous genotype calling and haplotype phasing improves genotype accuracy and reduces false-positive associations for genome-wide association studies. *American Journal of Human Genetics, 85*(6), 847–861.

Brückner, H., & Bearman, P. (2005). After the promise: The STD consequences of adolescent virginity pledges. *Journal of Adolescent Health, 36*(4), 271–278.

Chobanian, A. V., Bakris, G. L., Black, H. R., Cushman, W. C., Green, L. A., Izzo, J. L., National High Blood Pressure Education Program Coordinating Committee (2003). Seventh report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure. *Hypertension, 42*(6), 1206–1252.

Fagot-Campagna, A., Saaddine, J. B., Flegal, K. M., & Beckles, G. L. (2001). Diabetes, impaired fasting glucose, and elevated HbA1c in U.S. adolescents: The third national health and nutrition examination survey. *Diabetes Care, 24*(5), 834–837.

Ferraro, K. F., & Kelley-Moore, J. A. (2003). Cumulative disadvantage and health: Long-term consequences of obesity? *American Sociological Review, 68*(5), 707–729.

Finch, C. E., Vaupel, J. W., & Kinsella, K. (Eds.), Committee on Population, National Research Council. (2001). *Cells and surveys: Should biological measures be included in social science research?* Washington: National Academies Press.

Gettler, L. T., McDade, T. W., Feranil, A. B., & Kuzawa, C. W. (2011). Longitudinal evidence that fatherhood decreases testosterone in human males. *Proceedings of the National Academy of Sciences, 108*(39), 16194–16199.

Glaser, R., Pearson, G. R., Jones, J. F., Hillhouse, J., Kennedy, S., Mao, H. Y., & Kiecolt-Glaser, J. K. (1991). Stress-related activation of Epstein-Barr virus. *Brain, Behavior, and Immunity, 5*(2), 219–232.

Gordon-Larsen, P., The, N. S., & Adair, L. S. (2010). Longitudinal trends in obesity in the United States from adolescence to the third decade of life. *Obesity, 18*(9), 1801–1804.

Gottlieb, G., & Halpern, C. T. (2002). A relational view of causality in normal and abnormal development. *Development and Psychopathology, 14*(3), 421–435.

Guo, G., Cai, T., Guo, R., Wang, H., & Harris, K. M. (2010). The Dopamine Transporter Gene, a Spectrum of Most Common Risky Behaviors, and the Legal Status of the Behaviors. *PLoS ONE, 5*(2), e9352. doi:10.1371/journal.pone.0009352.

Halpern, C. T., & Udry, J. R. (1992). Variation in adolescent hormone measures and implications for behavioral research. *Journal of Research on Adolescence, 2*(2), 103–122.

Halpern, C. T., Whitsel, E. A., Wagner, B., & Harris, K. M. (2012). Challenges of measuring diurnal cortisol concentrations in a large population-based field study. *Psychoneuroendocrinology, 37*(4), 499–508.

Harris, K. M. (2010). An integrative approach to health. *Demography, 47*(1), 1–22. Accessed 17 August 2012.

Harris, K. M. (2011). Design features of add health. http://www.cpc.unc.edu/projects/addhealth/data/guides/design%20paper%20WI-IV.pdf.

Herbert, T. B., & Cohen, S. (1993). Stress and immunity in humans: A meta-analytic review. *Psychosomatic Medicine, 55*(4), 364–379.

Hill, S. Y., Shen, S., Zezza, N., Hoffman, E. K., Perlin, M., & Allan, W. (2004). A genome wide search for alcoholism susceptibility genes. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 128B*(1), 102–113.

Kehoe, R., Wu, S. Y., Leske, M. C., & Chylack L. T. Jr., (1994). Comparing self-reported and physician-reported medical history. *American Journal of Epidemiology, 139*(8), 813–818.

Khaw, K. T., Wareham, N., Luben, R., Bingham, S., Oakes, S., Welch, A., & Day, N. (2001). Glycated hemoglobin, diabetes, and mortality in men in Norfolk cohort of European Prospective Investigation of Cancer and Nutrition (EPIC-Norfolk). *BMJ, 322,* 15–28.

Kugler, J., Hess, M., & Haake, D. (1992). Secretion of salivary immunoglobulin A in relation to age, saliva flow, mood states, secretion of albumin, cortisol, and catecholamines in saliva. *Journal of Clinical Immunology, 12*(1), 45–49.

McDade, T. W., Williams, S., & Snodgrass, J. J. (2007). What a drop can do: Dried blood spots as a minimally invasive method for integrating biomarkers into population-based research. *Demography*, *44*(4), 899–925. doi:10.1353/dem.2007.0038.

McEwen, B. S. (1998). Stress, adaptation and disease. Allostasis and allostatic load. *Annals of the New York Academy of Sciences, 840,* 33–44.

Miller, W. C., Ford, C. A., Morris, M., Handcock, M. S., Schmitz, J. L., Hobbs, M. M., Udry, J. R. (2004). The prevalence of chlamydial and gonococcal infections among young adults in the United States. *The Journal of the American Medical Association, 291*(18), 2229–2236.

Miller, G. E., Lachman, M. E., Chen, E., Gruenewald, T. L., Karlamangla, A. S., & Seeman, T. E. (2011). Pathways to resilience: Maternal nurturance as a buffer against the effects of childhood poverty on metabolic syndrome at midlife. *Psychological Science, 22*(12), 1591–1599. doi:10.1177/0956797611419170.

Morris, M., Handcock, M. S., Miller, W. C., Ford, C. A., Schmitz, J. L., Hobbs, M. M., & Udry, J. R. (2006). Prevalence of HIV infection among young adults in the United States: Results from the Add Health study. *The American Journal of Public Health, 96*(6), 1091–1097.

Nissen, S. E., Tuzcu, E. M., Schoenhagen, P., Crowe, T., Sasiela, W. J., Tsai, J., Ganz, P. (2005). Statin therapy, LDL cholesterol, C-reactive protein, and coronary artery disease. *New England Journal of Medicine, 352*(1), 29–38.

Nguyen, Q. C., Tabor, J. W., Entzel, P. P., Lau, Y., Suchindran, C., Hussey, J. M., Whitsel, E. A., et al. (2011). Discordance in national estimates of hypertension among young adults. *Epidemiology, 22*(4), 532–541.

Rector, R., & Johnson, K. A. (2005). *Adolescent virginity pledges, condom use and sexually transmitted diseases among young adults*. Paper presented at the National Welfare Research and Evaluation Conference, Washington, DC.

Udry, J. R. (1988). Biological predispositions and social control in adolescent sexual behavior. *American Sociological Review, 53*(5), 709–722.

Veerman, E. C., van den Keybus, P. A., Vissink, A., & Nieuw Amerongen, A. V. (1996). Human glandular salivas: Their separate collection and analysis. *European Journal of Oral Sciences, 104*(4), 346–352.

Vining, R. F., McGinley, R. A., & Symons, R. G. (1983). Hormones in saliva: Mode of entry and consequent implications for clinical interpretation. *Clinical Chemistry, 29*(10), 1752–1756.

Weinstein, M., Vaupel, J. W., & Wachter, K. W. (Eds.), Committee on Advances in Collecting and Utilizing Biological Indicators and Genetic Information in Social Science Surveys, National Research Council. (2008). *Biosocial surveys*. Washington: National Academies Press.