A Note from the Director

Welcome back to The Methodology Center Perspective!

It has been a busy winter and spring so far. In February the Center underwent an extremely helpful external review organized by the College of Health and Human Development's Associate Dean for Research, Dr. Neil Sharkey. On behalf of the entire Center I would like to extend thanks to the external reviewers, Dr. Stephen West from Arizona State University and Dr. Nancy Wolff from Rutgers University. I have never seen reviewers work so hard and do such a thorough job. Their report gave us some wonderful suggestions for future directions.

We welcomed a new research associate in January. Dr. Inbal Shani joined us after earning her doctorate in Behavioral Sciences at the Faculty of Industrial Engineering and Management, Technion—Israel Institute of Technology, Haifa, Israel. She is working with Susan Murphy and me on our research related to designing clinical trials and engineering more effective behavioral interventions, as well as pursuing her own interests in behavioral research.

Brittany Rhoades, a former trainee in the Prevention and Methodology Training program (PAMT) who is currently finishing up her graduate work, has accepted a position at Penn State as an Evaluation Research Specialist in the Prevention Center. She will be associated with a new endeavor, the Evidence-based Prevention and Intervention Support Center (EPISCenter), funded by the State of Pennsylvania. She will also continue to work in The Methodology Center as part of Stephanie Lanza's project.

Another former PAMT trainee, Dr. Bethany Bray, will be leaving the Center this summer to take a position as an Assistant Professor in the Psychology Department at Virginia Tech. Bethany has been a part of the Center for seven years, and we will miss her. We wish her the very best in her new position, and hope she comes back to visit often.

Every spring semester the Center organizes three one-credit graduate courses. Each course is five weeks long and offers an advanced introduction to a focused methodological topic. This spring the courses were: Latent Class Analysis for Cross-Sectional and Repeated Measures Data, offered by Stephanie Lanza, Bethany Bray and me; Emerging Methods for Identifying Subgroups, offered by Lisa Dierker, who is a Visiting Professor in the Center this year; and Missing Data: Methods and Designs, offered by John Graham. Each course was full, reflecting the continuing keen interest in methodology at Penn State.

It is hard for me to believe that it is already time for us to submit a renewal application for the T32 grant from NIDA that supports the Prevention and Methodology Training Program. Ed Smith, Mark Greenberg, and I are working on this now. We have enjoyed this training program tremendously and are optimistic that it will continue.

We hope you enjoy this issue of our newsletter!

Linda M. Collins, Ph.D.
Director, The Methodology Center
Penn State University

New Study Offers Practical Recommendations for Causal Inference Methods

Correlation is not causation. If X was not randomly assigned, then a relationship between X and Y does not describe how intervening on X would affect Y. In the social and behavioral sciences, however, we continually ask causal questions for which randomized experiments would be untenable. For example: How does dieting impact physical or psychological health? Dieting may be encouraged or discouraged by researchers, but an intervention that directly assigned the behavior to participants would be ethically or practically infeasible. If we set aside causal questions whenever the putative cause was not randomized, our domain of scientific inquiry would be severely limited.

The crucial issue in non-randomized studies is comparability. Dieters and non-dieters have different physical and psychosocial profiles. The traditional way to adjust for these differences is by analysis of covariance (ANCOVA), which treats confounders—variables on which the groups tend to differ—as predictors in a linear regression. Despite the popularity of ANCOVA, it may be a very poor method of adjusting for confounders, because results may be sensitive to departures from the assumed model. It also leads us to ask the wrong question. In ANCOVA, the effect of X is defined as the average difference in Y between groups of individuals who experienced different values of X. Causal inferences, however, are statements about hypothetical changes in Y if different values of X were experienced by the same individuals. Instead of asking, “How do outcomes differ for dieters and non-dieters?” we should be asking, “What would the dieters’ outcomes have been if they had not dieted?”

A quarter-century ago, Rubin (1974) introduced a novel notation in which each participant has two possible results: one if she exhibits the behavior, another if she does not. The difference between these “potential outcomes” is the causal effect of the behavior on the outcome. Only one potential outcome is seen for each person, so causal effects cannot be observed directly. By making certain assumptions, however, it is possible to estimate an average causal effect (ACE) in a population or subgroup. A vast and growing literature on causal inference from the standpoint of potential outcomes can now be found in statistics, epidemiology, and economics. The topic has become so popular that social and behavioral scientists who wish to delve into this material
New Web-based Technologies are Now Available

The Methodology Center’s website recently underwent an aesthetic renovation. Visit http://methodology.psu.edu to view the new design and navigation. The changes to the site allow for new web-based technologies to be easily utilized and readily available. These include:

New Searchable Publications Database
As part of the renovation, the Center’s publication database has been updated and is now searchable! Powered by Refworks, a citation software and bibliographic management tool, the new database features tools allowing impressive organization, searchability, and sharing.

RSS Feeds of Recent Articles
An RSS feed of the newest published articles available in the Refworks publications database is highlighted on the main page of the Center’s website. Visitors to the site can subscribe to the feed and receive the dynamic content to their own RSS readers.

The Methodology Center on Facebook and Wikipedia
Facebook.com is a free social network that connects people with friends and others who work, study and live around them. The Center’s new page on Facebook allows its members to connect with each other in a unique way. It will be utilized for dispensing Center news and events as well as facilitating conversation between its members. Visit the page by clicking on the “Find us on Facebook” icon on the main page of the Center’s site. In order to join, you must be a member of Facebook.

Wikipedia.com is a free digital encyclopedia project written collaboratively by volunteers around the world, and almost all of its articles can be edited by anyone with access to the Wikipedia website. Search for “Methodology Center” on Wikipedia and you’ll find the Center’s page.

The Center is continually looking to incorporate new web-based technologies to carry out its mission. Keep visiting The Methodology Center’s website to stay up to date on the newest endeavors!

Featured Scientist

Jacquelyn Wiersma

Jacquelyn D. Wiersma is a post-doctoral fellow in the Prevention and Methodology Training Program at Penn State University. Jackie received her Ph.D. in Human Development and Family Studies from Texas Tech University in 2008. Jackie’s research examines the associations of alcohol use and adolescent and young adult romantic relationships. She has focused on the congruency and discrepancy in young adult drinking patterns and how this contributes to couples’ satisfaction, commitment, alcohol-related problems, alcohol abuse (Wiersma, Fischer, & Fitzpatrick, 2009), and intimate partner violence in romantic relationships (Herrera, Wiersma, & Cleveland, 2008). Recently, she has been focusing on the processes of selection and socialization as it relates to young adult drinking in romantic relationships. For example, her dissertation found that romantic partner drinking was significantly associated with young adult drinking within unmarried relationships, and both selection and socialization effects played a role in young adults’ continued use of drinking in romantic relationships.

Currently, Jackie is working with Jennifer Maggs, Associate Professor of Human Development and Family Studies (Penn State University), on the University Life Study (ULS). The ULS was designed to examine links between motivations, daily activities, college experiences, and risk behaviors using semester baseline data and a series of 14 consecutive daily surveys from the fall of freshman to senior year. Jackie is investigating the associations among heavy alcohol use, casual sex behaviors, and romantic relationship sex behaviors. She is also examining differences in substance use for individuals with romantic partners and individuals with nonromantic partners.

Jackie continues her four-year work on the Add Health dataset, working alongside Bo Cleveland, Associate Professor of Human Development and Family Studies (Penn State University), to examine patterns of drinking in adolescent and young adulthood romantic relationships. Using the Add Health data, Jackie analyzes couple-level data with multilevel models to understand selection and socialization processes in romantic partner drinking. She is preparing to submit an RO3 this year to examine these processes in adolescent romantic drinking partnerships.


Both selection and socialization play a role in young adult romantic partners’ drinking.
now face a bewildering array of concepts and choices: propensity scores, G-estimation, marginal structural models, doubly robust estimators, and so on. Entree to the world of modern causal inference is difficult not only because of the sheer volume of material, but also because of the lack of consensus about which methodologies are best.

Two statisticians from The Methodology Center, Joe Schafer and Joseph Kang (now at Northwestern University), have produced a 35-page article for Psychological Methods that attempts to display the forest through the trees. Starting from the beginning, they discuss the difficulties of causal inference in non-randomized studies. Avoiding complicated notation and jargon, they introduce potential outcomes and ACE’s and explain how this framework differs from classical ANCOVA. They motivate and compare a wide variety of old and new methods for estimating ACE’s. Throughout the article, they offer insights and practical guidance for researchers who attempt causal inference with observational data.

This is not the first overview of causal inference written for non-experts. For example, a recent book by Stephen Morgan and Christopher Winship (2007) presents concepts and methods from a sociologist’s perspective, and in some respects the two publications overlap. One new feature of the article by Schafer and Kang, however, is the extensive use of simulated data to illustrate and compare methods. They created an artificial population of one million adolescent girls to represent a hypothetical observational study of the effects of dieting on emotional distress. Distributions used to generate the data mimic actual variables from the National Longitudinal Study of Adolescent Health (Add Health). They drew 5,000 simulated samples from this population and applied nine different estimation techniques to each sample. Another noteworthy feature of this article is a technical appendix with detailed formulas for standard errors. General approaches for computing standard errors have been described in statistics and biostatistics journals, but to our knowledge the detailed standard-error formulas for specific ACE estimators have not been compiled or presented elsewhere.

For researchers who wish to try out these methods, our website provides access to one sample from the artificial population and R code to reproduce each estimate and standard error: http://www.stat.psu.edu/~jls/causal.html


### Recent Activity in The Methodology Center

Donna Coffman received funding from the National Institute on Drug Abuse (NIDA) for an R03 project to study causal inferences for mediational processes in prevention studies. The project began April 1 and will study the robustness of causal inferences to violations of assumptions.

Michael Cleveland received a National Institute on Drug Abuse (NIDA) Travel Award to participate in the Early Career Poster Session that NIDA is co-sponsoring with the National Institute on Alcohol Abuse and Alcoholism (NIAAA) and the American Psychological Association (APA) Divisions 28 (Psychopharmacology and Drug Abuse) and 50 (Addictions) during this year’s APA Annual Convention.

A paper written by Michael Zhang with Runze Li and another coauthor at University of California Davis was selected as one of the twenty recipients of the Distinguished Student Paper Awards at the 2009 Meeting of the Eastern North American Region (ENAR) of the International Biometric Society. Michael presented a talk “Regularization Parameter Selections via Generalized Information Criterion” based on this paper at the ENAR conference in San Antonio, TX.

Michael Zhang has accepted a position as Senior Biostatistician at Novartis Oncology located in Florham Park, NJ. At Novartis, Michael will conduct research related to design and analysis of clinical trials.

Linda Collins, John Dziak, and Runze Li wrote an overview paper comparing different kinds of experimental design trial design for studying multivariate treatments and interventions. They compared the relative costs and benefits of full factorial, fractional factorial, and one-at-a-time strategies for examining the importance of different treatment factors in this context. The article will appear in a forthcoming issue of Psychological Methods.


A paper by Susan Murphy and a colleague from Simon Fraser University will appear in an upcoming issue of the Journal of the American Statistical Association. This paper combines ideas from fractional factorial experimental designs for developing multicomponent interventions with full factorial experimental designs for developing adaptive interventions to consider more than one treatment component at any given time.


Rhonda BeLue and Stephanie Lanza have an article scheduled to appear in Ethnicity and Disease which employed latent class analysis to investigate the relationship between high cholesterol status and diet and exercise behavior.


Lisa Dierker, Linda Collins, and colleagues from the Tobacco Etiology Research Network used time series methods to evaluate substance use trends across the first year of college and determine the predictive value of day-to-day patterns. The article appeared in the journal Substance Use and Misuse.


Michael Cleveland, along with colleagues from the Prevention Research Center at Penn State, recently had a manuscript appear in the Journal of Youth and Adolescence. The authors used hierarchical logistic regression to examine the influence of family-based protective factors across different school contexts.


Donna Coffman and Rhonda BeLue have a manuscript scheduled to appear in the Journal of Community Psychology. This study illustrates the use of item response theory to detect racial differences in the Sense of Community Index.


---

Instead of asking, “How do outcomes differ for dieters and non-dieters?” we should be asking, “What would the dieters’ outcomes have been if they had not dieted?”
Ask a Methodologist

I would like to use growth mixture trajectory analysis to study the developmental progression of both alcohol and cigarette use. However, I cannot decide how I should examine the relationship between these two variables. Should I treat each one as a time-varying covariate for the other process or is it better to use a dual trajectory approach? — Mixed Up

Dear Mixed Up,

This is a very interesting question. Although there are many excellent technical papers that describe growth mixture trajectory analysis and an increasing number of substantive papers that have applied this method to substance use data, very little information currently exists that evaluates the impact of several important design and measurement decisions on the results of such models. One of these decisions involves how to examine the association between time-varying independent variables and trajectories. Opportunities for considering these sorts of proximal effects can be handled through two approaches in growth mixture trajectory models. The first approach treats time-varying variables as covariates and indicates the extent to which deviation from the average group trajectory is related to the covariates. Alternately, by treating the time-varying variable as another developmental process, the second approach (dual trajectory analysis or joint trajectory analysis) aims to analyze the connections between the trajectories of two processes that are evolving contemporaneously (such as depression and alcohol use) or that evolve over different time periods that may or may not overlap (such as prosocial behavior in childhood and school achievement in adolescence).

A recent study (Tan, Dierker, Rose, & Li, under review) found that each of the two approaches suggests something unique about the relationship between cigarette smoking and alcohol use. The trajectory analysis with a time-varying covariate evaluates how much change in smoking one would expect if we observe certain changes in alcohol use, within each specific subgroup. In contrast, dual trajectory analysis describes whether or not a person is more likely to be in certain smoking trajectory group given that he/she is in a specific alcohol use trajectory group. These two methods could be useful for different situations and for answering different research questions. For example, the trajectory analysis with covariate can be useful in generating hypotheses regarding the group-specific mechanism of substance use development, consequently shedding light on appropriate intervention. On the other hand, the dual trajectory analysis approach can be helpful in identifying high-risk subgroups for prevention treatment. In this example, subjects having high stable or increasing smoking trajectory were most likely to also have a high stable or increasing alcohol use trajectory and were likely most in need of intervention.