Incorporating Partially Validated Data into Analyses of Error-Prone Datasets to Improve Estimation.

Bryan Shepherd, Ph.D.
Associate Professor
Department of Biostatistics
Vanderbilt University School of Medicine

There is growing interest in using observational data from clinical, administrative, and electronic health records as cost-effective resources to support medical research. An increasing number of articles have highlighted data quality concerns with these types of datasets; spurious or unrecognized associations driven by incorrectly recorded exposures and outcomes could misguide clinical decision-making. In our own work using data from local and international HIV observational databases, we have found error rates as high as 50% for some key study variables. However, given budget constraints, many researchers question and often ignore the necessity of data validation/auditing. I will discuss our current research on statistical methods that can improve estimation by incorporating validation information obtained on a sub-sample of records. These methods are related to those in the statistical literature on missing data and measurement error, except in our setting errors are likely in multiple variables (including outcomes) and the presence and magnitude of the errors between variables are often correlated. I will demonstrate the use of these methods with data from the Caribbean, Central, and South American Network (CCASAnet) for HIV Research and the Vanderbilt Comprehensive Care Clinic (VCCC). In CCASAnet, we are often left wondering what to do after discovering errors during on-site data audits: ignore them, have the researchers re-enter all data, or exclude sites. Our approach offers a compromise by incorporating the audit data into analyses. At the VCCC, an endpoints committee performs comprehensive chart reviews to validate key variables. As we have data before and after validation, we examine the cost-savings and feasibility of obtaining unbiased results by validating only a sub-sample of records and applying our methods.

FRIDAY, APRIL 17TH, 2015
12:00 PM TO 1:00 PM
898-J PRB

*Vanderbilt University School of Medicine is accredited by the Accreditation Council for Continuing Medical Education to provide continuing medical education for physicians.

Vanderbilt University School of Medicine designates this live activity for a maximum of 1.0 AMA PRA Category 1 Credit(s)™. Physicians should claim only the credit commensurate with the extent of their participation in the activity.

For more information, please contact Jill Shell at jill.shell@vanderbilt.edu or 6-6760.
A schedule of the series and additional information is available at: https://medschool.vanderbilt.edu/cqs/workshop2015