Training Description and Curriculum

December 13 -14, 2010
8am-5pm

Topic
Using Applied Multilevel Modeling for MCH EPI Data Analysis

Background
Multilevel modeling has become a mainstay of public health analysis where subjects are often observed clustered within larger units (e.g., clinics, providers, service areas) or where subjects are repeatedly measured across time. Traditional approaches are insufficient for accounting for the clustering of the data, resulting in underestimation of measures of uncertainly, and therefore invalid statistical conclusions. Recently, the introduction of multilevel modeling has given epidemiologists a new tool for producing adjusted measures that can be used to lay out a set of risk factors according to their impact on a health outcome, and therefore inform the prioritization of prevention and intervention strategies. In this workshop, attendees will learn about the use of multilevel models (aka mixed or hierarchical models) for analysis of clustered and longitudinal datasets, continuous and dichotomous outcomes, and other related types of multilevel datasets.

Training Goal
The purpose of this training is to provide an in depth conceptual and methodological overview of data analysis using multilevel modeling, along with hands-on experience in calculating and interpreting measures using public health datasets.

Training Objectives
Participants in this training will:
1. Gain an understanding of the different perspectives of collecting and analyzing MCH datasets using multilevel modeling;
2. Become familiar with the multilevel logistic regression models for analysis of clustered and longitudinal datasets, continuous and dichotomous outcomes, and other forms of multilevel datasets used in common epidemiological studies;
3. Gain experience in using extensions of the multilevel logistic model for ordinal and nominal outcomes;
Target Audience and Prerequisites
This workshop is aimed at public health professionals such as epidemiologists and other statisticians who work in public health agencies. Attendees should have a good working knowledge of regression analysis, both for continuous and dichotomous outcomes. Registration will be limited to 30 participants.

Training Methods and Approaches
The training presentations will be interactive and will include the application of multilevel models, with direct application illustrated using standard statistical software (i.e., SAS). The training participants will be asked to share state examples of identified best practices for state-multilevel MCHEPI datasets. Participants will be provided with a laptop for the duration of the training.

Course Curriculum
Day one
Lecture – Multilevel models for clustered continuous outcomes
- Description and structure of multilevel data
- Multilevel model description, meaning of random cluster effects and intraclass correlation (ICC)
- 2- and 3-level examples, model comparisons, R squared measures
Exercise I – running 2- and 3-level models with SAS PROC MIXED

Lecture – Multilevel models for clustered dichotomous outcomes
- Extension of logistic regression
- Latent variable formulation, meaning of regression coefficients, ICC
- 2- and 3-level examples, comparison of models
Exercise II – running 2- and 3-level models for clustered dichotomous data with SAS PROC GLIMMIX and SAS PROC NLMIXED

Lecture – Multilevel models for clustered ordinal and nominal outcomes
- Proportional odds models for ordinal outcomes
- Non proportional odds models for ordinal outcomes
- Nominal models with reference cell
Exercise III - running 2- and 3-level models for clustered ordinal and nominal data with SAS PROC GLIMMIX and SAS PROC NLMIXED; testing the proportional odds assumption
Wrap-up for Day One: Discussion of Conceptual and Technical Issues in Analyzing Clustered Data
Day two
Lecture – Multilevel models for longitudinal continuous outcomes
- Population and subject-specific time trends
- Multilevel model description, meaning of random effects and variance parameters
- Treatment of time-varying covariates
Exercise IV – running longitudinal models with SAS PROC MIXED

Lecture – Multilevel models for longitudinal dichotomous outcomes
- Extension of logistic regression
- Latent variable formulation, subject-specific and marginal regression coefficients
- Examples, comparison of models, fitting marginal proportions
Exercise V – running models for longitudinal dichotomous data with SAS PROC GLIMMIX and SAS PROC NLMIXED

Lecture – Missing data in longitudinal studies
- Missing data mechanisms
- Testing missing completely at random assumption
- Sensitivity analysis to missing data assumptions (pattern-mixture and selection models)
Exercise VI - running pattern-mixture models with SAS PROC MIXED

Wrap-up for Day Two: Discussion of Conceptual and Technical Issues in Analyzing Longitudinal Data

**Trainer Information**
Don Hedeker, PhD
Professor
Division of Epidemiology and Biostatistics
School of Public Health
University of Illinois at Chicago
1603 W. Taylor St., room 955
Chicago, IL  60612-4336
Phone: (312) 996-4896
Fax: (312) 996-0064
Email: hedeker@uic.edu

Donald Hedeker, PhD., is a Professor of Biostatistics in the School of Public Health. He received his Ph.D. in Quantitative Psychology from The University of Chicago, IL. Don's main expertise is in the development and use of advanced statistical methods for clustered and longitudinal data, with particular emphasis on mixed-effects models. He is the primary author of four freeware computer programs for mixed-effects analysis: MIXREG for normal-theory models, MIXOR for dichotomous and ordinal outcomes, MIXNO for nominal outcomes, and MIXPREG for counts. These programs and manuals are available at http://www.uic.edu/~hedeker/mix.html.

Don has received numerous awards including named a Fellow of the American Statistical Association, the highest honor in his field in 2000. He was also recognized as a University Scholar by the University of Illinois that same year. Don is a past Associate Editor for Journal of
Educational and Behavioral Statistics and is a Consulting Editor for Statistics for Bipolar Disorders. He has also been the PI, co-PI, or co-I on over a dozen NIH or CDC research grants.

**Samples of Background Reading Material**


