Development of cost-effective statistical sampling strategies and optimal design considerations for exposure assessment as part of the National Children's Study

Warren Strauss[1], Louise Ryan[2], Michele Morara[1], Haluk Ozkaynak[3] and Jim Quackenboss[3]

[1]Battelle Memorial Institute – Centers for Public Health Research and Evaluation
[2]Commonwealth Scientific and Industrial Research Organization (CSIRO)
[3]U.S. Environmental Protection Agency – National Exposure Research Laboratory

Key Words: Optimal Design, Measurement Error, Designed Missingness, Validation Sampling, Longitudinal, Environmental Exposure, Replicate Sampling

This project was performed to develop innovative statistical study design guidance for the acquisition of exposure data over time in the National Children's Study (NCS). The NCS is intended to investigate environmental influences on children's health and development–including understanding environmental exposures that may cause or exacerbate health impacts. Due to the large sample size and longitudinal nature of the Study, unique statistical issues arise that must be addressed before a cost-effective sampling design can be developed to gather the environmental and personal exposure data. A key issue for the NCS relates to obtaining enough samples to provide adequate statistical power to detect health effects attributable to environmental exposures, while being cost-effective, minimizing participant burden, and staying within the study's overall budget. Relevant specific issues include identification of potential sources of bias and/or uncertainty in the exposure measures (non-response, subject burden, attrition, and measurement error); and strategies to address these issues. The statistical techniques that have been developed concentrate on using a validation sampling approach—in which detailed exposure and/or health outcome assessment information is collected on a carefully selected subset of the study population, while cheaper (and presumably less accurate and less burdensome) surrogate measures are collected across the entire NCS cohort. Compared to classical design methods, these approaches reduce costs of exposure assessment by up to two thirds with no loss in statistical power while also reducing the burden of expensive and sometimes invasive measures on the majority of NCS study participants.

Please note: The validation sampling topic would also work well for a presentation at a breakout session, or a poster.