ATELIER « INFÉRENCE CAUSALE EN RECHERCHE SUR LA SANTÉ »
9 – 13 MAI 2011

WORKSHOP ON "CAUSAL INFERENCE IN HEALTH RESEARCH"
MAY 9 – 13, 2011

Bias analysis to guide new data collection : comprehensive *CYP2D6* genotyping in a study of tamoxifen resistance as an example

Timothy L. Lash *
tl@dce.au.dk

Bias analysis serves multiple objectives in epidemiologic data analysis. The objectives most often emphasized are quantification of uncertainty due to systematic errors and reduction in overconfidence by virtue of specifying hypotheses that compete with the causal hypothesis to explain non-null associations. A third objective, less often emphasized, is the utility of bias analysis to identify strategies for new data collection that will be most productive in evaluating the validity of an association. The author illustrates the value of this objective using the example of comprehensive *CYP2D6* genotyping in a study of tamoxifen resistance. Tamoxifen is an endocrine therapy that reduces the risk of breast cancer recurrence by about half. The parent drug is metabolized primarily by *CYP2D6* to more active forms. More than thirty polymorphisms in the *CYP2D6* gene reduce or eliminate its function, so may reduce the drug's effectiveness. We genotyped the most prevalent *CYP2D6* polymorphism and found no association between genotype and breast cancer recurrence. One possibility is that *CYP2D6* function is unrelated to breast cancer recurrence risk in tamoxifen-treated patients, and there is a biologic rationale to support this hypothesis. A second possibility is that incomplete genotyping of the multiple functional polymorphisms introduced non-differential misclassification and biased the association toward the null. We used bias analysis, relying on external data sources, to evaluate the plausibility of this second explanation and to guide a decision about devoting study resources toward additional data collection (*i.e.*, more comprehensive genotyping of other polymorphisms in the *CYP2D6* gene). The example illustrates the utility of bias analysis to guide new data collection within a study, or by initiation of new studies. This objective of bias analysis assists with productive expenditure of limited research resources to resolve competing explanations for observed associations.

*Department of Clinical Epidemiology, Aarhus University Hospital, Olof Palmes Alle 43–45, 8200 Aarhus N, Denmark.