

Title: Causal diagrams to guide the treatment of missing data in epidemiological studies with multiple incomplete variables

Author(s): Margarita Moreno-Betancur

Session: Biostatistics and Bioinformatics

Missing data are a common occurrence in epidemiological studies and may impact study conclusions due to potential bias and loss of statistical power. It is widely understood that if the data are “missing at random” (MAR) – an assumption allowing the probability of missing data to depend on observed values – then unbiased estimation is possible with appropriate methods. While the need to assess the plausibility of this assumption has been emphasised, the practical difficulty of these tasks and the stringency of MAR in the context of multiple incomplete variables are rarely acknowledged. Further, while MAR is sufficient, it is certainly not necessary: in a wide range of “missing not at random” (MNAR) scenarios unbiased estimation of certain parameters is possible.

Recent developments in the computer science literature suggest that directed acyclic graphs (DAGs) could be an intuitive tool for stating and assessing finer-grained assumptions, beyond the MAR-MNAR dichotomy. However, as we show, translating the assumptions in a given generic DAG to a decision about the missing data method is a surprisingly complex problem requiring a case-by-case treatment. Seeking a balance between detail and feasibility, we constructed eight “canonical” DAGs representing broad categories of missingness mechanisms that could be encountered in a typical point-exposure epidemiological study with incomplete exposure, outcome and confounders. For each DAG, we derived mathematically whether unbiased estimation of some common target parameters is possible using common procedures, or if sensitivity analyses are necessary. These DAGs and findings can be readily used by epidemiologists to articulate their assumptions, and choose a strategy to handle missing data depending on their target parameter. We use numerical simulations and the Longitudinal Study of Australian Children for illustration.