

## Design-based strategies for bias in observational studies

Sebastien Haneuse \*

shaneuse@hsph.harvard.edu

---

When interpreting results from an observational study, potential bias due to confounding is a key scientific challenge. As such considerable effort is often devoted at the design stage to the identification of potential confounders, as well as the description of analysis strategies that remove the bias. Numerous other biases, often encountered in observational studies, also lend themselves to consideration at the design stage. For example, selection bias can result in retrospective studies if there are systematic differences between participation rates among cases and controls and we present two recently proposed approaches to address this problem. Placing these approaches in the broader context of a design-based philosophy for overcoming bias, we note that an emerging area of research that may benefit from such a philosophy is that based on the use of electronic medical record (EMR) data. Such data provide a growing opportunity for cost-effective public health research since they typically consist of huge quantities of health-related data on large, diverse populations. A key challenge of using EMR data, however, is that mechanisms underlying data collection do not correspond to any particular scientific question. To illustrate this we present a recently funded grant that seeks to exploit EMR data to understand the relationship between body weight change and choice of treatment, among individuals with a new diagnosis of depression. Unfortunately, the measurement of weight in the EMR, at baseline and/or during follow-up, may be driven by a number of patient-, physician- and facility-level characteristics. While being a nuisance with respect to the main scientific goals, insufficient understanding and inadequate adjustment of the mechanisms underlying the measurement processes may result in bias. However, such mechanisms do lend themselves to design-based strategies for overcoming potential bias by proactively collecting additional data. We present the proposed data collection scheme for the body weight study and describe a number of outstanding statistical challenges.

---

\*Department of Biostatistics, Harvard School of Public Health, 655 Huntington Avenue, Boston, MA 02115, USA.