Causal diagrams: back to the future for Brazilian epidemiology

The search for causal links lies at the core of epidemiology as a scientific field oriented to the study of health-related events in human populations. However, the challenge of causal inference has intrigued philosophers and scientists alike for centuries.

Mainly beginning in the late 18th century, various referential contributions emerged for causal studies in epidemiology. These include the work by Lind and Snow on scurvy and cholera, respectively, the Henle-Koch postulates, within the germ theory, Hill's criteria of causation, Rothman's sufficient-component cause model, Rubin's potential outcomes model, and the vast writings by Miettinen, Robins, and Greenland and their school since the 1970s.

The approaches have differed substantially, but one common thread connects all the contributions closest to contemporary epidemiology: the assumption that the possibility of causal inference requires adherence to the principles of validity and precision and the existence of theoretical-operational models that sustain the causal hypotheses in question.

However, the good practice of drawing explicit causal models or graphs has not spread as widely as necessary, perhaps because these models tend to be highly difficult to operationalize, given the complexity involved in the determination of the health-disease process at the population level. We also appear to lack a better strategy for applying these models to study design and analysis of epidemiological data.

Among Brazilian epidemiology's relevant contributions to causality issues and causal inference, two have sought to lend greater cognizance and practicality to causal models. Almeida Filho [1] has painstakingly tackled the epidemiologist's daunting task of transposing concepts from the theoretical to the empirical level. Meanwhile, Victora et al. [2] emphasizes the role of these conceptual models in the analysis of epidemiological data, introducing the notion of modeling with hierarchical organization of variables.

Despite its relevant background and scientific output in causal inference, Brazilian epidemiology appears to have missed the huge methodological advances in this area, in both graduate training and research practice itself. At the international level, causal diagrams and new modeling strategies in causal inference in observational studies have been an area of prolific research at least since the 1980s, with a strong presence in graduate courses and the leading epidemiology journals. For example, in 1986 Robins [3] presented a graphical approach to the identification and estimation of causal parameters in occupational cohort studies potentially subject to the healthy worker bias. Later, a method known as G-estimation was developed to control bias in epidemiological studies when a time-dependent risk factor acts simultaneously as a confounding factor and intermediate variable [4]. The 1990s witnessed the first formal theoretical and conceptual elaboration of causal diagrams, particularly "Directed Acyclic Graphs" (DAGs), thereby unveiling an area of methodological development with a world of potential applications to epidemiological research [5].

The paper by Cortes et al. in this issue of CSP should be read against this backdrop. The article is timely for Brazilian epidemiology, which has suffered a mismatch between the importance of causal diagrams in the international epidemiological literature and their limited repercussions in the Brazilian academic and scientific community. One finds occasional applications of causal diagrams and related methods by Brazilian researchers, but at levels far short of desirable.

Readers of Cortes et al. will enjoy a comprehensive review of DAGs and their use in epidemiology, with an emphasis on their application to confounding. The authors present a topology of DAGs and define essential terms for communication via such diagrams, such as "d-separation criterion", "back-door path", and "collider variable". Cortes et al. illustrate

all this formal elaboration with a real-life research problem in which multiple variables are interconnected in a theoretical-operational model expressed by a DAG.

Readers will appreciate that DAGs are simple tools that allow researchers to identify, among a wide range of variables, a minimum set of potential confounders that need to be controlled to obtain valid results. DAGs are also useful for identifying variables that may appear to be eligible as confounders, but which, if controlled, may actually introduce confounding. In other words, the common practice of "controlling for everything" can have harmful effects.

Although causal diagrams possess a formal algebraic structure based on conditional probabilities that allows unbiased estimation of effect measures, Cortes et al. discuss their use heuristically in order to assist the selection of confounding variables for subsequent control using traditional analytical methods. However, the authors provide the necessary references for readers interested in greater depth, in order to extend the applications to problems beyond confounding.

In addition to assisting the identification and selection of variables to be used in the control of confounding, the adoption of DAGs helps reclaim the good practice in epidemiological research (and in scientific research as a whole) of explicitly stating the hypotheses in advance concerning the web of causal relations among the phenomena in question.

Cortes et al. also renew the hope that a new generation of epidemiologists will focus on the theme of causal inference in epidemiology and succeed in qualitatively expanding Brazilian research output on the determinants of population health.

Guilherme L. Werneck
Associate Editor

1. Almeida-Filho N. Epidemiologia sem números. Rio de Janeiro: Campus; 1989.
2. Victora CG, Huttly SR, Fuchs SC, Olinto MTA. The role of conceptual frameworks in epidemiological analysis: a hierarchical approach. Int J Epidemiol 1997; 26:224-7.
3. Robins J. A graphical approach to the identification and estimation of causal parameters in mortality studies with sustained exposure periods. J Chronic Dis 1987; 40 Suppl 2:139S-61.
4. Robins JM, Blevins D, Ritter G, Wulfsohn M. G-estimation of the effect of prophylaxis therapy for *Pneumocystis carinii* pneumonia on the survival of AIDS patients. Epidemiology 1992; 3:319-36.
5. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. Epidemiology 1999; 10:37-48.