

Diagramas causais: a epidemiologia brasileira de volta para o futuro

A intenção de estabelecer nexos causais está no cerne da Epidemiologia como campo científico orientado para o estudo de eventos relacionados à saúde em populações humanas. No entanto, inferir causalidade é um desafio que intriga filósofos e cientistas há vários séculos.

Principalmente a partir do final do século XVIII, surgiram diversas contribuições referenciais para os estudos causais no âmbito da pesquisa epidemiológica. Essas incluem os trabalhos de Lind e Snow sobre escorbuto e cólera, respectivamente; os postulados de Henle-Koch, fruto da revolução microbiológica; os critérios de causalidade de Hill; o modelo de causas suficientes e componentes de Rothman; o modelo de respostas potenciais de Rubin e os vastos escritos da escola de Miettinen, Robins e Greenland, a partir da década de 1970.

Que pesem as substanciais diferenças de abordagens, um aspecto comum àquelas contribuições mais próximas da epidemiologia contemporânea é o pressuposto de que a possibilidade de inferência causal requer aderência aos princípios de validade e precisão em estudos epidemiológicos, e a existência de modelos teórico-operacionais que sustentem as hipóteses causais em questão.

Entretanto, a boa prática de explicitar modelos ou gráficos causais não foi disseminada de forma tão abrangente como seria necessário, talvez porque estes modelos tendam a ser muito difíceis de operacionalizar, dada a complexidade envolvida na determinação do processo saúde-doença em âmbito populacional. Também, parece faltar uma melhor estratégia de apropriação desses modelos para fins de guiar o delineamento dos estudos e a análise dos dados epidemiológicos.

Dentre algumas relevantes contribuições da epidemiologia brasileira aos temas da causalidade e da inferência causal, duas delas procuraram dar maior cognição e praticidade aos modelos causais. Por um lado, Almeida Filho ¹, esmiuçando o difícil trabalho do epidemiologista em promover a transposição de conceitos de um plano teórico para o nível empírico. Por outro, Victora et al. ², enfatizando o papel desses modelos conceituais na análise de dados epidemiológicos e introduzindo a noção de modelagem com hierarquização de variáveis.

Que pese este histórico de interesse e produção científica no campo da inferência causal, a epidemiologia brasileira parece ter caminhado, tanto no que tange à formação pós-graduada quanto em relação à própria prática de pesquisa, ao largo dos imensos desenvolvimentos metodológicos nesta área. No âmbito internacional, ao contrário, o uso de gráficos causais e de novas estratégias de modelagem no contexto da inferência causal em estudos observacionais é uma área de estudos prolífica desde pelo menos a década de 1980, com forte penetração nos cursos de pós-graduação e principais periódicos de epidemiologia. Robins ³, já em 1986, por exemplo, apresentou uma abordagem gráfica para identificação e estimação de parâmetros causais em estudos de coorte ocupacional potencialmente sujeitos ao viés do trabalhador saudável. Posteriormente foi desenvolvido o método conhecido com *G-estimation*, uma ferramenta para o controle de vieses em estudos epidemiológicos quando um fator de risco tempo-dependente atua simultaneamente como um fator de confusão e uma variável intermediária ⁴. A partir da década de 1990, houve a formalização teórico-conceitual dos diagramas causais, particularmente dos gráficos acíclicos direcionados (*Directed Acyclic Graphs* – DAGs), descortinando um campo de desenvolvimento metodológico com um universo de aplicações potenciais na pesquisa epidemiológica ⁵.

É a partir desse pano de fundo que o artigo de Cortes et al., que integra este fascículo de CSP, deve ser apreciado. O artigo é muito oportuno para a epidemiologia brasileira na medida em que percebe-se um descompasso entre a importância do tema dos diagramas causais no contexto da literatura epidemiológica internacional e a sua pouca repercussão no ambiente acadêmico-científico nacional. Há, aqui ou ali, aplicações desses e outros métodos correlatos por pesquisadores brasileiros, mas em níveis muito aquém do desejável.

No artigo de Cortes et al., os leitores terão uma revisão compreensiva do uso dos DAGs em epidemiologia, com ênfase na abordagem da situação de confusão. Para tal, são apresentadas a topologia dos DAGs e os significados dos termos essenciais para a comunicação por meio desses gráficos, como “critério de *d*-separação”, “caminho pela porta de trás” e “variável colisora”. Toda essa formalização é dissecada mediante a apresentação de um problema concreto de pesquisa em que múltiplas variáveis são interconectadas em um modelo teórico-operacional expresso por meio de um DAG.

Apreende-se do artigo que DAGs são ferramentas simples, que permitem o pesquisador identificar dentre uma gama enorme de variáveis um conjunto mínimo de potenciais fatores de confusão que precisam ser controlados para a obtenção de resultados válidos. DAGs são úteis também para identificar variáveis que, aparentemente poderiam ser elegíveis como fatores de confusão, mas que se controladas na análise poderão provocar situação de confusão. Ou seja, a prática usual de “controlar para tudo” pode ter efeitos deletérios.

Ainda que esses gráficos tenham uma estrutura algébrica formal baseada em probabilidades condicionais que permite a estimativa não enviesada de medidas de efeito, o artigo de Cortes et al. dedica-se a apresentar de forma heurística a utilização destes gráficos para auxiliar a escolha de quais variáveis de confusão devem ser selecionadas para posterior controle usando-se métodos analíticos tradicionais. Entretanto, os autores proveem as fontes necessárias para que o leitor interessado se aprofunde no tema, buscando expandir sua aplicabilidade para problemas além da situação de confusão.

Além de ajudar na identificação e seleção de variáveis a serem utilizadas para fins de controle de confundimento, o uso dos DAGs contribui para retomar a boa prática da pesquisa epidemiológica e científica em geral, que é explicitar as hipóteses *a priori* sobre a rede de relações causais entre os fenômenos que estão sendo estudados.

Ademais, o artigo de Cortes et al. renova as esperanças de que uma nova geração de epidemiologistas voltará a se debruçar sobre o tema da inferência causal em epidemiologia, e conseguirá incrementar qualitativamente a produção científica brasileira no campo da investigação dos determinantes da saúde nas populações.

Guilherme L. Werneck

Editor Associado

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.