



# Lidando com fatores de confusão em estudos observacionais

Cristiane Fumo-dos-Santos<sup>1</sup>, Juliana Carvalho Ferreira<sup>1,2</sup>

## CENÁRIO PRÁTICO

Investigadores de um grande centro acadêmico em São Paulo examinaram a associação do uso de ventilação protetora, definida como volume corrente  $< 8$  mL/kg de peso corporal previsto e pressão de platô  $< 30$  cmH<sub>2</sub>O, com sobrevida em pacientes com COVID-19 grave. Eles também coletaram dados sobre a gravidade da doença na admissão em UTI, necessidade de terapia renal substitutiva e diversos parâmetros ventilatórios. Eles descobriram que o uso de ventilação protetora estava associado à melhora da sobrevida, com *hazard ratio* ajustada de 0,73 (IC95%: 0,57-0,94;  $p = 0,013$ ).

## INFERÊNCIA CAUSAL EM ESTUDOS OBSERVACIONAIS

Em estudos epidemiológicos, os investigadores não atribuem intervenções, mas sim classificam os indivíduos como expostos ou não expostos a fatores de risco para o desenvolvimento de um desfecho. Quando uma associação estatisticamente significativa é encontrada, diversas explicações possíveis precisam ser consideradas:

1. A associação é real e o preditor (ventilação protetora, no nosso exemplo) é verdadeiramente uma causa do desfecho (sobrevida, no nosso exemplo).
2. A associação é real, mas é uma relação de efeito-causa: o desfecho (sobrevida) causa o preditor (ventilação protetora). Em nosso exemplo não seria plausível considerar esta possibilidade, mas isso faz sentido em muitos casos.
3. A associação se deve ao acaso — erro aleatório. Como geralmente consideramos um valor de  $p < 0,05$  como significativo, e o valor de  $p$  em nosso exemplo foi 0,013, há uma probabilidade de 1,3% de que o acaso seja a explicação para essa associação.
4. A associação não é real, é resultado de um erro sistemático (viés), decorrente de aspectos metodológicos do estudo, como subestimar sistematicamente o peso corporal previsto dos pacientes.
5. A associação é real, mas é confundida pelo efeito de outra(s) variável(eis) associada(s) ao desfecho e ao preditor.

## O QUE É CONFUNDIMENTO?

Confundir deriva do latim *confundere*, misturar. A definição clássica de fator de confusão é qualquer terceira variável associada à exposição de interesse, que seja uma causa do desfecho de interesse e que não resida no caminho causal entre a exposição e o desfecho (Figura 1A). Por exemplo, no nosso cenário prático, os

investigadores consideraram que a complacência pulmonar, entre outras variáveis, era um potencial fator de confusão, porque a baixa complacência pulmonar é uma causa de morte (portanto, reduzindo a sobrevida), e também está associada ao preditor — quando a complacência é muito baixa, pode ser mais difícil aplicar ventilação protetora. Já a gravidade da doença na admissão não foi tratada como fator de confusão pelos investigadores, pois embora esteja altamente associada ao desfecho (óbito), não possui relação causal com o preditor de interesse (ventilação protetora).<sup>(1)</sup> Embora possamos ter fatores de confusão na pesquisa experimental, essa é uma questão mais importante a ser considerada em estudos observacionais.<sup>(2)</sup>

## POR QUE DEVEMOS NOS PREOCUPAR EM IDENTIFICAR FATORES DE CONFUSÃO?

Fatores de confusão podem levar à superestimação ou subestimação do efeito do principal preditor sobre o desfecho de interesse, tornando o efeito não confiável e interferindo em nossa capacidade de fazer inferências causais em estudos observacionais.<sup>(2)</sup> Portanto, estratégias estatísticas são recomendadas para controlar ou ajustar a análise de fatores de confusão a fim de observar o efeito verdadeiro e isolado do preditor de interesse no desfecho.

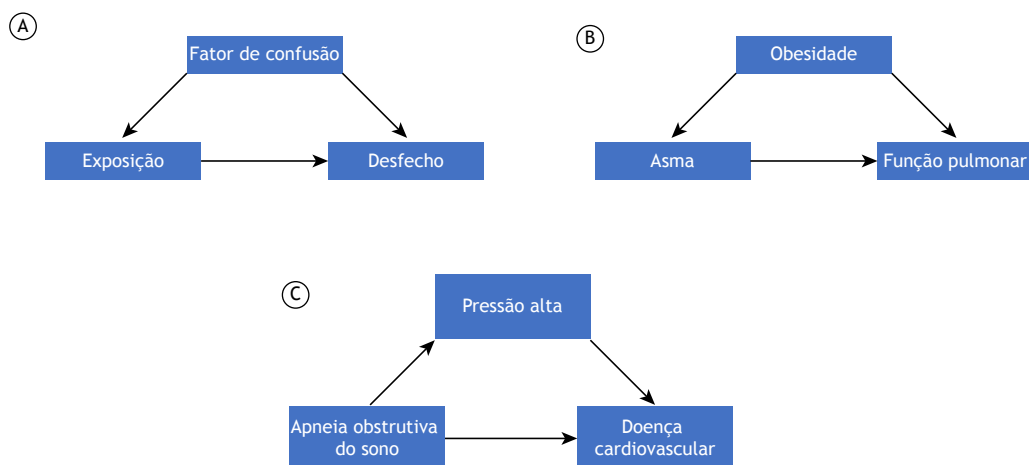
Não devemos identificar um fator de confusão com base em testes estatísticos, mas sim no conhecimento clínico prévio ou na fisiopatologia do processo que estamos estudando.<sup>(1)</sup> Uma das estratégias mais aceitas para identificar um fator de confusão é utilizar o conhecimento prévio sobre o desfecho de interesse e modelos causais, especialmente critérios gráficos.<sup>(2)</sup> Essa abordagem é importante porque a forma tradicional de identificar o fator de confusão, conforme descrito anteriormente, é muitas vezes inadequada em estruturas mais complexas.<sup>(3)</sup>

## COMO PODEMOS LIDAR COM FATORES DE CONFUSÃO?

A melhor maneira de lidar com fatores de confusão é planejar com antecedência. Um ensaio controlado randomizado atribui indivíduos aleatoriamente aos braços de intervenção e controle do estudo, dispersando os fatores de confusão conhecidos e desconhecidos em cada braço. No entanto, esse desenho não é adequado para responder a muitas perguntas de estudo importantes.<sup>(1)</sup>

1. Methods in Epidemiologic, Clinical, and Operations Research-MECOR-program, American Thoracic Society/Asociación Latinoamericana del Tórax, Montevideo, Uruguay.

2. Divisão de Pneumologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo (SP) Brasil.



**Figura 1.** Em A, representação do padrão de confundimento: a variável é relacionada à exposição, é causa do desfecho e não está na via causal entre a exposição principal e o desfecho de interesse. Em B, a obesidade é um fator de confusão na relação entre asma e função pulmonar, uma vez que a obesidade pode piorar a asma e causar redução da função pulmonar. O ajuste para obesidade é aconselhável neste cenário. Em C, o modelo representa um efeito de mediação — a apneia obstrutiva do sono pode levar a doenças cardiovasculares (efeito direto), mas a apneia obstrutiva do sono também pode levar à hipertensão arterial, que causa doenças cardiovasculares (efeito indireto). Neste caso, o ajuste para pressão arterial elevada não é apropriado.

Selecionar indivíduos com a mesma característica também é uma estratégia: para estudar a redução da função pulmonar na asma, pesquisadores podem excluir pessoas com obesidade. O problema dessa estratégia é que os resultados não se aplicam a todos os indivíduos com asma, mas apenas aos pacientes asmáticos não obesos. Outra forma de lidar com o confundimento é parear indivíduos: o pesquisador seleciona o mesmo número de participantes com e sem obesidade tanto no grupo exposto quanto no grupo não exposto.<sup>(1)</sup> Novamente, porém, a manipulação resulta na redução da generalização dos resultados.

A estratégia mais comumente usada para lidar com fatores de confusão é controlar ou ajustar durante a

análise estatística uma vez que os modelos de regressão podem abordar vários preditores ao mesmo tempo.<sup>(3)</sup> Neste caso, é realmente importante construir um modelo causal e ajustá-lo apenas para fatores de confusão ao invés de ajustá-lo para todas as variáveis com base em valores de *p*, por exemplo.

A mensagem principal é que os fatores de confusão podem interferir na inferência causal em estudos observacionais, e precisamos planejar com antecedência para identificar, medir, minimizar e ajustar os fatores de confusão a fim de usar os resultados de estudos observacionais para orientar futuras pesquisas e tomadas de decisões clínicas.

## REFERÊNCIAS

1. van Stralen KJ, Dekker FW, Zoccali C, Jager KJ. Confounding. *Nephron Clin Pract.* 2010;116(2):c143-c147. <https://doi.org/10.1159/000315883>
2. Santos S, Zugna D, Pizzi C, Richiardi L. Sources of confounding in life course epidemiology. *J Dev Orig Health Dis.* 2019;10(3):299-305. <https://doi.org/10.1017/S2040174418000582>
3. Lederer DJ, Bell SC, Branson RD, Chalmers JD, Marshall R, Maslove DM, et al. Control of Confounding and Reporting of Results in Causal Inference Studies. Guidance for Authors from Editors of Respiratory, Sleep, and Critical Care Journals [published correction appears in *Ann Am Thorac Soc.* 2019 Feb;16(2):283]. *Ann Am Thorac Soc.* 2019;16(1):22-28. <https://doi.org/10.1513/AnnalsATS.201808-564PS>