



Perda de seguimento e dados faltantes: questões importantes que podem afetar os resultados do seu estudo

Juliana Carvalho Ferreira^{1,2,a}, Cecília Maria Patino^{1,3,b}

Pesquisadores utilizaram dados provenientes do banco de dados do Programa de Controle da Tuberculose do Estado de São Paulo para avaliar a relação entre o local do diagnóstico e os desfechos do tratamento. Eles relataram que 25% dos casos novos de tuberculose foram diagnosticados no pronto-socorro e que, em comparação aos pacientes diagnosticados no ambulatório, tinham maior probabilidade de apresentar desfechos desfavoráveis do tratamento, incluindo perda de seguimento (em 12%) e óbito (em 10%).⁽¹⁾

PERDA DE SEGUIMENTO E TIPOS DE DADOS FALTANTES

Para estudar as relações de uma exposição (em estudos de coorte) ou de uma intervenção (em ensaios clínicos randomizados) com os desfechos clínicos de forma prospectiva, os investigadores coletam informações de exposição/intervenção no início do estudo e então fazem o seguimento dos participantes do estudo ao longo do tempo e coletam dados de desfecho. Se o seguimento for incompleto ou for interrompido, levando a dados faltantes no final do estudo, isso pode afetar a validade interna. Os participantes com dados faltantes, comparados aos com dados completos, podem diferir sistematicamente, por exemplo, quando a perda de seguimento está relacionada à morte dos participantes.

Dados faltantes ocorrem por vários motivos: 1. a variável de interesse não é medida pela equipe de pesquisa (por exemplo, esquecer-se de medir o peso basal); 2. o participante do estudo não comparece a uma visita ou teste agendado do estudo; ou 3. a variável é medida, mas a equipe não registra o valor da variável na ficha de coleta de dados. No entanto, a perda de seguimento, situação em que os dados dos pacientes não estão disponíveis até o final do período de seguimento, é o mecanismo mais crítico de dados faltantes, pois pode incluir dados faltantes para o desfecho, e dados de desfecho são cruciais para responder a pergunta do estudo.

Dados faltantes são classificados como *missing completely at random* (MCAR, faltantes de forma totalmente aleatória), quando a incompletude não está relacionada à exposição, a covariáveis ou ao desfecho; *missing at random* (MAR, faltantes de forma aleatória), quando a incompletude está relacionada à exposição ou a fatores de confusão, mas não ao desfecho; e *missing not at random* (MNAR, faltantes de forma não aleatória), quando a incompletude pode estar relacionada ao desfecho (Quadro 1).⁽²⁾ A perda de seguimento e dados faltantes podem ameaçar a validade interna de um estudo mesmo se o mecanismo for MCAR, em que consideramos que os demais participantes são uma amostra aleatória da população inicial do estudo, pois o poder do estudo será reduzido. Se o mecanismo for MAR, ajustes e métodos de imputação podem ser utilizados, mas isso pode introduzir vieses no estudo. Se o mecanismo for MNAR, há um sério risco de resultados enviesados.⁽³⁾

COMO LIDAR COM A PERDA DE SEGUIMENTO E DADOS FALTANTES

A melhor estratégia para evitar dados faltantes é evitar a perda de seguimento. São fundamentais o planejamento cuidadoso do estudo, o treinamento da equipe, a implementação de procedimentos de qualidade de dados e o desenvolvimento de mecanismos para reter e contatar os participantes. Além disso, existem métodos estatísticos disponíveis para lidar com dados faltantes, mas esses procedimentos devem ser planejados a priori e com o auxílio de um bioestatístico.⁽³⁾ No entanto, existem situações em que os investigadores podem não superar os problemas relacionados aos dados faltantes em virtude de o mecanismo de incompletude ser MNAR. Nesse caso, perdas de seguimento de 20%, por exemplo, podem resultar em sérios vieses e, portanto, não devem ser consideradas "aceitáveis". Lembre-se, dados faltantes são comuns, e as melhores práticas incluem pensar nisso logo no início, ao definir a pergunta do estudo e escrever o protocolo.

Quadro 1. Tipos de dados faltantes e estratégias para sua minimização.

Tipos de dados faltantes	Exemplos	Estratégias para minimização de dados faltantes
MCAR	Participante se muda para outro estado e abandona o estudo; um resultado de teste é perdido no laboratório	Desenvolver fichas de coleta padronizadas; monitorar a qualidade dos dados; manter as informações de contato do participante atualizadas
MAR	Em uma coorte de pacientes com DPOC, participantes com doença leve têm maior probabilidade de abandonar o estudo porque são assintomáticos	Oferecer benefícios e incentivos para reter os participantes; contatar os participantes com regularidade; realizar um estudo piloto para a identificação de fatores de risco para a perda de seguimento; e desenvolver estratégias para superá-los
MNAR	A perda de seguimento é maior entre pacientes com tuberculose que apresentam efeitos adversos graves aos fármacos antituberculose do que entre pacientes que toleram o tratamento, e a não adesão ao tratamento está relacionada a óbito	Oferecer apoio adequado aos participantes do estudo; desenvolver estratégias para reter os participantes com alto risco de perda de seguimento; e desenvolver métodos alternativos para mensuração do desfecho mesmo para participantes com perda de seguimento

MCAR: *missing completely at random* (faltantes de forma totalmente aleatória); MAR: *missing at random* (faltantes de forma aleatória); e MNAR: *missing not at random* (faltantes de forma não aleatória).

REFERÊNCIAS

1. Ranzani OT, Rodrigues LC, Waldman EA, Prina E, Carvalho CRR. Who are the patients with tuberculosis who are diagnosed in emergency facilities? An analysis of treatment outcomes in the state of São Paulo, Brazil. *J Bras Pneumol.* 2018;44(2):125-133. <https://doi.org/10.1590/s1806-37562017000000384>
2. Pedersen AB, Mikkelsen EM, Cronin-Fenton D, Kristensen NR, Pham TM, Pedersen L, et al. Missing data and multiple imputation in clinical epidemiological research. *Clin Epidemiol.* 2017;9:157-166. <https://doi.org/10.2147/CLEP.S129785>
3. Kristman V, Manno M, Côté P. Loss to follow-up in cohort studies: how much is too much? *Eur J Epidemiol.* 2004;19(8):751-60

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.

3. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.

a. <http://orcid.org/0000-0001-5742-2157>; b. <http://orcid.org/0000-0001-6548-1384>