

Dados Longitudinais e o Viés da Correlação de Medidas: A Alternativa dos Modelos Mistos

Longitudinal Data and Correlated Measures Bias: The Alternative of Mixed Models

Johnnatas Mikael Lopes,¹ Marcello Barbosa O.G. Guedes,² Rafael Limeira Cavalcanti,³ Clecio Gabriel de Souza²

Universidade Federal do Vale do São Francisco (UNIVASF) - Colegiado de Medicina,¹ Paulo Afonso, BA – Brasil

Universidade Federal do Rio Grande do Norte (UFRN) – Fisioterapia,² Natal, RN – Brasil

Centro Universitário Maurício de Nassau (UNINASSAU) – Fisioterapia,³ Natal, RN – Brasil

Estudos longitudinais têm duas tipologias importantes de dados: desfechos únicos ou medidas repetidas.¹ Coortes de desfecho único, como a morte ou surgimento de doença, devem ter um tratamento de dados diferente daqueles estudos com desfecho de medidas repetidas. Mas todos eles têm em comum a detecção de mudanças ao longo do tempo e os fatores contribuintes para esta alteração. Isto difere dos estudos transversais que buscam apenas relação de variáveis, sem o necessário efeito de causalidade.

O estudo de Fernandes et al.,² intitulado de *Relação entre Estilo de Vida e Custos Relacionados ao Uso de Medicamentos em Adultos*, publicado neste periódico no volume 112, número 6 em 2019, utilizou variáveis independentes de comportamento para estimar os efeitos sobre o desfecho de custos com medicamentos, coletado na forma de medidas repetidas em um desenho de coorte prospectiva.

O que se pretende com esta exposição é evidenciar que, provavelmente, ocorreu um equívoco na análise de dados do trabalho de Fernandes et al.,² que compromete as inferências de causalidade devido à grande possibilidade de precisão das estimativas apresentadas estarem equivocadas.

Vamos aos fatos. Considerando o desenho de coorte prospectiva com medidas repetidas constata-se a existência de estrutura hierárquica nos dados do desfecho devido ao aninhamento deles nas diversas medidas do mesmo participante. O aninhamento de dados faz com que o erro, isto é, o que foi predito pelo modelo e a real aferição das medidas, do mesmo participante em momentos distintos sejam correlacionados.³ Isto é uma condição para não uso da regressão linear múltipla (RLM) a qual tem como pressuposto a independência do erro dada pela assunção de que a distribuição de cada participante é igual. A RLM não extrai

dos dados aquilo que é variabilidade dentro do indivíduo da variabilidade entre indivíduos (população).³

Usar RLM em medidas repetidas gera erros-padrão dos coeficientes de regressão viciados. Exige-se nesta situação a aplicação de uma matriz de covariância que produzirá estimativas mais confiáveis, ou seja, intervalos de confiança mais estreitos a partir de modelos de efeitos mistos.⁴ Esta é a melhor alternativa para se verificar mudanças ao longo do tempo ou o efeito de condicionantes no desfecho em estudos longitudinais, controlando os efeitos individuais.

Como existe uma variabilidade maior entre indivíduos que intra-indivíduo, devido principalmente às diferenças biológicas e de condicionantes sociais, observa-se que os custos com medicamentos irão ser mais correlacionados ao longo do tempo no mesmo indivíduo que entre os participantes. Pensar que essa distribuição é igual nos participantes do estudo, ignora o pressuposto teórico da determinação social em saúde no comportamento das pessoas.⁵

A criação de RLM distintas (A, B, C e D), ver Fernandes et al.,² não controla esse efeito de covariância e, portanto, podem estar produzindo coeficientes com intervalos de confiança enviesados para as variáveis independentes e não consegue também detectar a taxa de mudança em relação ao basal.³ Além disso, com os modelos mistos seria possível também aproveitar medidas que foram aferidas em participantes perdidos, aumentando a sensibilidade da modelagem.⁴

Em outra perspectiva, sendo o objetivo da pesquisa estimar a inter-relação do custo com medicamentos e os hábitos comportamentais, sem o estabelecimento de causalidade, seria necessário apenas um desenho transversal dos participantes com a coleta de dados do desfecho e variáveis independentes em um único momento. Assim, o modelo de regressão basal seria suficiente para estimar associações brutas e ajustadas.¹

Dessa forma, o uso de RLM deve ficar restrito aos desenhos transversais de pesquisa e que estudos longitudinais necessitam diferenciar o efeito individual do efeito populacional na identificação das mudanças temporais e os condicionantes da mesma. Possivelmente, os achados de Fernandes et al.,² devem estar enviesados quanto às conclusões sobre a relação inversa do uso de álcool com os custos dos medicamentos ou as relações não-significativas do ponto de vista estatístico com a gordura corporal, gênero e tabagismo que têm grande impacto em outras situações de saúde, principalmente crônicas.

Palavras-chave

Estudos de Coortes; Estudos Longitudinais; Estudos Transversais; Epidemiologia; Bioestatística.

Correspondência: Johnnatas Mikael Lopes •

UNIVASF - Colegiado de Medicina - Centro de Formação Profissional de Paulo Afonso (CFPPA) - Rua da Aurora, S/N, Quadra 27, Lote 3. CEP 48607-190, Bairro General Dutra Paulo Afonso, BA – Brasil

E-mail: johnnataslopes2@gmail.com

Artigo recebido em 02/09/2019, revisado em 04/09/2019, aceito em 04/09/2019

DOI: <https://doi.org/10.36660/abc.20190601>

Referências

1. Medronho RA, Bloch KV, Raggio LR, Werneck GL. Epidemiologia. 2ª ed. São Paulo: Atheneu; 2009. 685 p.
2. Fernandes RA, Mantovani AM, Sanches Codogno J, Camilo Turi-Lynch B, Pokhrel S, Anokye N. Relação entre Estilo de Vida e Custos Relacionados ao Uso de Medicamentos em Adultos. Arq Bras Cardiol. 2018;11(2):749–55.
3. Fausto MA, Carneiro M, Antunes CMD, Pinto JA, Colosimo EA. O modelo de regressão linear misto para dados longitudinais: Uma aplicação na análise de dados antropométricos desbalanceados. Cad Saude Publica. 2008;24(3):513–24.
4. Helena Constantino Spyrides M, José Struchiner C, Tereza Serrano Barbosa Gilberto Kac M. Análise de Dados com Medidas Repetidas. In: Kac C, Sichieri R, Gigante D Epidemiologia nutricional. Rio de Janeiro: Editora FIOCRUZ/Atheneu; 2007.p. 245–60.
5. Garbois JA, Sodre F, Dalbello-Araujo M. Da noção de determinação social à de determinantes sociais da saúde. Saúde Debate. 2017;41(112):63–76.

Carta-resposta

Agradeço a oportunidade de resposta aos questionamentos relacionados ao nosso manuscrito recentemente publicado no Arquivos Brasileiros de Cardiologia.¹ A discussão acadêmica sempre é saudável e bem-vinda.

Inicialmente, agradeço o interesse em nosso trabalho. O questionamento levantado refere-se ao uso da regressão linear no tratamento de dados provenientes de uma coorte prospectiva com medidas repetidas, fato que teria levado a geração de estimativas equivocadas (faz-se a sugestão do uso da regressão linear mista). A crítica sustenta-se na fragilidade da regressão linear em detectar a variabilidade intra indivíduo, ao passo que o modelo centra na variabilidade entre os indivíduos. Do ponto de vista teórico, o comentário está correto, porém não reflete a forma como os dados foram analisados neste estudo.

A variável dependente deste estudo foi definida como “custos com medicamentos acumulados ao longo de 12 meses”. Neste estudo, não se tentou analisar a trajetória dos custos com medicamento ao longo do ano² (e como estas trajetórias seriam afetadas pelas variáveis comportamentais) e, também, não se buscou identificar a relação entre mudanças frente à linha de base (para variáveis dependentes e independentes). Houve sim, a tentativa de analisar a relação das variáveis comportamentais com o montante final acumulado ao longo do ano.

De fato, esta variável dependente é inusitada em sua forma de construção, pois foi elaborada longitudinalmente (gastos com medicamentos computados ao longo de 12 meses), mas tratada transversalmente (total acumulado ao longo dos 12 meses). O montante total de gastos com medicamento reflete um construto transversal, embora sua construção considere os 12 meses de seguimento. Esta particularidade da variável dependente, somada ao fato das variáveis comportamentais terem sido coletadas em apenas dois momentos do estudo (linha de base e ao final de 12 meses), nos levou a criar os quatro modelos propostos no artigo, os quais caracterizam uma leitura transversal do problema (principalmente os modelos A [dados de linha de base] e B [ao final de 12 meses]). Infelizmente, a avaliação mensal das variáveis comportamentais não foi uma opção metodológica disponível.

Em um modelo ideal, a variável dependente e as variáveis independentes deveriam ser coletadas mensalmente, possibilitando identificar o impacto de mudanças nas variáveis comportamentais sobre mudanças na trajetória dos custos com medicamento ao longo do ano. Porém, repito, este não foi o objetivo do referido artigo.¹ Para este tipo de análise, modelos de equação estrutural com modelagem específica (*Latent Growth Curve Analysis*) seriam mais interessantes (até mais que a regressão linear mista), pois possibilitariam analisar o impacto direto da taxa de mudança ao longo do seguimento na variável independente sobre a taxa de mudança da variável dependente.³ As medidas de “impacto” geradas pelo modelo são facilmente interpretadas, pois podem ser expressas em medidas de correlação, as quais adicionalmente fornecem medidas de tamanho de efeito.⁴

Adicionalmente, a apresentação da variável dependente da forma como foi apresentada (transversalmente, os gastos acumulados ao longo do seguimento) se fez necessária devido a particularidades observadas na estrutura da mesma. Diferentemente de outras variáveis usualmente mensuradas nas diferentes áreas das ciências da saúde (estatura, pressão arterial, componentes do perfil lipídico), as quais não apresentam o valor zero, os gastos com medicamentos ocorrem de maneira irregular, refletindo-se na elevada ocorrência de valores zero, ou seja, é possível relatar gastos no primeiro mês de coleta e, ao longo dos demais meses, relatar nenhum gasto. Nesse pano de fundo, análises considerando a variável mês a mês seriam problemáticas. Da mesma forma, a questão da variabilidade intra indivíduo precisa ser considerada com cautela neste estudo devido ao fato do gasto com medicamento do mês anterior não se repetir no mês seguinte, diferentemente do observado para variáveis como a estatura⁵ que, mesmo sem ganho, o valor do mês anterior se repetirá no mês seguinte.

Por fim, a ausência de relações significativas para obesidade e fumo não chega a ser uma surpresa neste estudo, devido ao fato da amostra ser relativamente jovem, sem a presença de doenças crônicas e com baixa ocorrência de tabagismo.

Rômulo Araújo Fernandes

Referências

1. Fernandes RA, Mantovani AM, Sanches Codogno J, Camilo Turi-Lynch B, Pokhrel S, Anokye N. Relação entre Estilo de Vida e Custos Relacionados ao Uso de Medicamentos em Adultos. *Arq Bras Cardiol.* 2018;11(2):749-55.
2. Lemes ÍR, Fernandes RA, Turi-Lynch BC, Codogno JS, de Moraes LC, Koyama KAK, Monteiro HL. Metabolic Syndrome, Physical Activity, and Medication-Related Expenditures: A Longitudinal Analysis. *J Phys Act Health.* 2019;16(10):830-5.
3. Brailean A, Aartsen MJ, Muniz-Terrera G, Prince M, Prina AM, Comijs HC, Huisman M, Beekman A. Longitudinal associations between late-life depression dimensions and cognitive functioning: a cross-domain latent growth curve analysis. *Psychol Med.* 2017;47(4):690-702.
4. Maher JM, Markey JC, Ebert-May D. The other half of the story: effect size analysis in quantitative research. *CBE Life Sci Educ.* 2013;12(3):345-51.
5. Fausto MA, Carneiro M, Antunes CM, Pinto JA, Colosimo EA. Mixed linear regression model for longitudinal data: application to an unbalanced anthropometric data set. *Cad Saude Publica.* 2008;24(3):513-24.



Este é um artigo de acesso aberto distribuído sob os termos da licença de atribuição pelo Creative Commons