

Explorando os efeitos da recusa de participação na linha de base ao seguimento sobre a validade em estudos de coorte

Exploring the effects of non-response to the baseline and follow-up on the validity in cohort studies

Cláudia Medina Coeli¹, Marília Sá Carvalho², Dora Chor³,
Rejane Sobrino Pinheiro¹, Patrícia Viana Guimarães⁴, Daniel Skaba⁵

Resumo

Poucos artigos avaliaram os efeitos da não resposta em estudos de coorte. O objetivo desse estudo foi avaliar se a não resposta, tanto na linha de base como ao seguimento, pode introduzir viés de seleção na associação entre nível ocupacional e mortalidade. Foram aplicadas as estimativas de não resposta e de mortalidade do estudo Whitehall II a uma coorte hipotética que abrangeu 110.000 pessoas-ano (nível ocupacional: baixo – 50.000; elevado – 60.000). Foram considerados quatro cenários: (1) participação integral; (2) não resposta na linha de base; (3) não resposta ao seguimento; (4) não resposta na linha de base e ao seguimento. Foram calculadas as razões de taxas (RR) de mortalidade considerando o nível ocupacional elevado como referência e a ausência de confundimento e viés de informação. A RR do primeiro cenário foi 1,79, com somente o último cenário mostrando um impacto importante da não resposta na estimativa da razão de taxas (RR=1,2). Os resultados sugerem que a combinação da não resposta em diferentes estágios de um estudo de coorte pode causar viés de seleção.

Palavras-chave: estudos de coortes; recusa de participação; viés de seleção.

Abstract

Few articles have evaluated the effects of non-response in cohort studies. The aim of this study was to assess if non-response, both to the baseline and to follow-up, may introduce selection bias on the association between occupational level and mortality. There were applied the non-response and mortality estimates of the Whitehall II study to a hypothetical cohort encompassing 110.000 person-years (occupational level: low – 50,000; high – 60,000). There were considered four scenarios: (1) full participation; (2) non-response to baseline; (3) non-response to follow-up; (4) non-response to baseline and follow-up. There were calculated mortality rate ratios (RR) considering the high occupation level as reference and absence of confounding or information bias. The first scenario RR was 1.79, with only the last scenario showing an important impact of the non-response on the rate ratio estimate (RR=1.2). The results suggest that combining non-response in different stages of cohort studies may cause selection bias.

Keywords: cohort studies; refusal to participate; selection bias.

Trabalho realizado na Universidade Federal do Rio de Janeiro (UFRJ) – Rio de Janeiro (RJ), Brasil.

¹Professora Adjunta do Departamento de Medicina Preventiva da Faculdade de Medicina e Instituto de Estudos em Saúde Coletiva da UFRJ – Rio de Janeiro (RJ), Brasil.

²Pesquisadora titular da Fundação Oswaldo Cruz (FIOCRUZ) – Rio de Janeiro (RJ), Brasil.

³Professora Adjunta da Escola Nacional de Saúde Pública da FIOCRUZ – Rio de Janeiro (RJ), Brasil.

⁴Doutoranda do Programa de Pós-graduação em Saúde Coletiva do Instituto de Estudos em Saúde Coletiva da UFRJ – Rio de Janeiro (RJ), Brasil.

⁵Tecnologista sênior da Fundação Instituto Brasileiro de Geografia e Estatística (IBGE) – Rio de Janeiro (RJ), Brasil.

Endereço para correspondência: Cláudia Medina Coeli – Avenida Horácio Macedo, s/n – Ilha do Fundão – Cidade Universitária – CEP: 21941-598 – Rio de Janeiro (RJ), Brasil – E-mail: coeli@iesc.ufrj.br

Fonte de financiamento: Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Conflito de interesse: nada a declarar.

Estudos de coorte são sujeitos a viés de seleção quando a inclusão e permanência de participantes expostos e não expostos depende da probabilidade de adoecer. A recusa de participação na linha de base, a não resposta ao longo do seguimento e a recusa em responder alguns itens do questionário podem introduzir viés de seleção se esses eventos estiverem associados à probabilidade futura de desenvolvimento do desfecho de interesse¹. Embora efeito importante sobre as medidas de associação seja esperado apenas quando a probabilidade de ocorrência de desfecho for elevada¹, esse pode se tornar relevante se a não participação for simultaneamente associadas à exposição e ao desfecho, representando, portanto, a situação com maior impacto potencial sobre a validade das medidas de associação estimadas^{1,2}.

Fatores como idade, sexo, renda, nível educacional, raça, estado conjugal, tabagismo e estado de saúde podem influenciar a participação em estudos longitudinais³⁻⁶. Por outro lado, maior risco de mortalidade foi observado em não participantes quando comparados aos participantes⁵⁻⁷. O viés de seleção pode se originar se a exposição (ou uma causa da mesma) e o desfecho (ou uma causa do mesmo) influenciarem a participação no estudo, já que a análise é realizada condicionada ao efeito comum (participação) da exposição e do desfecho⁸. Entretanto, poucos estudos buscaram explorar os efeitos da não resposta, quer na linha de base ou ao seguimento, sobre a validade de medidas de associação em estudos de coorte⁹.

Contrariando a ideia de que a não participação ao longo do seguimento poderia ser mais influenciada pela probabilidade futura de desenvolvimento do desfecho do que a recusa de participação inicial, Ferrie et al.⁶, avaliando a experiência do estudo Whitehall II, observaram que não participantes, quer por recusa inicial ou durante o seguimento, apresentaram risco semelhante de mortalidade, sendo esse o dobro do observado entre os participantes que completaram o estudo, não havendo modificação desse padrão de acordo com o nível ocupacional. Esses autores não exploraram, entretanto, os efeitos dessas perdas sobre a associação entre nível ocupacional e a mortalidade. Nesse estudo, buscou-se avaliar essa questão aplicando a uma coorte hipotética as estimativas de mortalidade e de não resposta observadas no estudo Whitehall II.

METODOLOGIA

Foi construída uma coorte hipotética formada pela experiência de seguimento completo de 110.000 participantes-ano, sendo 50.000 com nível ocupacional baixo e 60.000 com nível ocupacional elevado (categoria de referência). Essa experiência de pessoa-tempo buscou reproduzir a encontrada no

estudo Whitehall II, embora nesse trabalho tenha sido ignorada a experiência de participantes com nível ocupacional intermediário. Como medida de associação, empregou-se a razão de taxas, assumindo a ausência de outros erros sistemáticos (viés de informação ou confundimento).

Foram elaborados quatro cenários, a saber: (1) participação integral (na linha de base e ao seguimento), sendo a medida de associação estimada considerada válida, i.e., considerou-se esse cenário como de referência para a avaliação de viés; (2) não resposta na linha de base; (3) não resposta ao seguimento; (4) não resposta na linha de base e ao seguimento, i.e., a combinação dos que recusaram a participação no estudo com as perdas ao seguimento. Para o presente estudo, considerou-se como perda de seguimento a categoria de participantes do estudo Whitehall II que faltaram a uma ou mais fases de coleta de dados, incluindo a última. As informações sobre óbitos e taxas de mortalidade brutas apresentadas por Ferrie et al.⁶ foram empregadas para a obtenção da experiência de pessoa-tempo em cada estrato formado pela combinação do nível ocupacional e situação de resposta na linha de base e ao seguimento no artigo original. A seguir, os dados relativos aos óbitos foram somados considerando os cenários acima descritos.

Para o cálculo da experiência de pessoa-tempo nos quatro cenários, foram utilizadas estratégias diferentes. No primeiro cenário (participação integral), agregou-se a experiência de pessoa-tempo em cada nível ocupacional. No segundo (não resposta na linha de base), subtraiu-se a experiência dos que recusaram participar do estudo do total de pessoa-tempo do primeiro cenário. Com relação às perdas de seguimento, na ausência de informação sobre o momento em que as mesmas ocorreram, optou-se por considerar que a perda se deu no último dia de seguimento para todos os perdidos de observação. Dessa forma, o total de pessoa-tempo do terceiro cenário foi igual ao do primeiro, enquanto o do quarto cenário ficou igual ao do segundo, havendo apenas a perda da informação relativa aos óbitos.

Por fim, foram calculados os indicadores taxa de mortalidade e proporção de experiência populacional seguida em cada nível ocupacional. Esses indicadores foram, então, aplicados à coorte hipotética para estimativa de óbitos, experiência populacional e cálculo das razões de taxas nos diferentes cenários avaliados.

RESULTADOS

Na Tabela 1 são apresentados os resultados obtidos em cada cenário hipotético. Quando comparado ao grupo com nível ocupacional elevado, o grupo de nível baixo apresentou

Tabela 1. Associação entre nível ocupacional e mortalidade em uma coorte hipotética considerando diferentes cenários de acordo com a não resposta inicial e ao longo do seguimento

Cenário	Óbitos	Pessoas-ano	Taxa de mortalidade*	Razão das taxas	IC95%
(1) Participação integral (na linha de base e ao seguimento)					
Nível ocupacional baixo	305	50.000	6,1	1,79	1,6–2,01
Nível ocupacional elevado	204	60.000	3,4	1	
(2) Não resposta na linha de base					
Nível ocupacional baixo	219	40.969	5,3	1,66	1,45–1,90
Nível ocupacional elevado	177	55.127	3,2	1	
(3) Não resposta ao seguimento					
Nível ocupacional baixo	195	50.000	3,9	1,58	1,27–1,97
Nível ocupacional elevado	148	60.000	2,5	1	
(4) Não resposta na linha de base e ao seguimento					
Nível ocupacional baixo	109	40.969	2,6	1,2	0,91–1,57
Nível ocupacional elevado	122	55.127	2,2	1	

*Taxa de mortalidade por 1.000; IC95%: intervalo de confiança de 95%

aproximadamente o dobro da perda de experiência populacional devido a não participação na linha de base. Considerando o cenário ideal com ausência de não resposta (cenário 1), verificou-se uma mortalidade maior entre os participantes de menor nível ocupacional quando comparados aos de maior nível, sendo a razão de taxas igual a 1,79. Nos demais cenários pode ser observado um enviesamento das estimativas pontuais em direção à ausência de associação, além da perda da precisão. Entretanto, apenas quando houve combinação da não resposta em ambas as fases (cenário 4) observou-se perda da significância estatística e uma diferença relevante da medida de associação estimada (cerca de 30% menor).

DISCUSSÃO

No presente estudo, aplicando-se a experiência de seguimento do estudo Whitehall II a uma coorte hipotética, observou-se que tanto a recusa inicial de participação, quanto a não resposta ao seguimento não apresentaram impacto importante sobre a razão de taxas de mortalidade entre os níveis ocupacionais. Esses resultados estão de acordo com o observado por estudos que avaliaram a não resposta inicial⁵ e ao longo do seguimento^{7,10}. Foi observado, entretanto, que a combinação das duas perdas levou ao enviesamento importante da medida de associação. Coerente com o verificado no estudo Whitehall II⁶, no estudo GAZEL¹¹ foi encontrado que os fatores que influenciam a participação inicial são semelhantes aos que influenciam a permanência no longo prazo. Dessa forma, os resultados poderiam ser explicados por uma amplificação, ao longo do seguimento, da indução de viés de seleção pelos fatores associados à participação inicial¹¹.

As limitações desse estudo foram avaliar a experiência de um único fator de exposição (nível ocupacional) sobre

um único desfecho (mortalidade), partindo-se da premissa da inexistência de outras ameaças à validade (viés de informação ou confundimento). Entretanto, o nível ocupacional é um dos fatores que mais consistentemente vem sendo apontado como associado à recusa de participação em estudos de coorte e a mortalidade o desfecho, em geral, mais estudado em função da disponibilidade ampla de dados de mortalidade para a realização do seguimento passivo^{4,6,12}. Quanto a outros fatores que poderiam influenciar a validade do estudo, os resultados podem ser interpretados como o que ocorreria nesse cenário particular. Cabe ressaltar, que na análise conduzida por Ferrei et al.⁶ o dobro de mortalidade observada nos perdidos de participação em relação aos que permaneceram foi estimada após ajuste para gênero e idade. Outra limitação foi não poder estimar de forma mais adequada a perda de experiência populacional devida a não resposta ao seguimento. Entretanto, uma análise (não apresentada) considerando o cenário oposto (apenas um dia de seguimento), embora levando ao menor enviesamento das estimativas, não mudou o padrão observado.

Poucos estudos buscaram avaliar o impacto da não resposta sobre a validade em estudos de coorte⁹ e, até onde foi possível saber, esse estudo foi o primeiro a explorar o impacto da combinação da não resposta inicial e ao longo do seguimento. Em um recente editorial, Stang⁹ chamou a atenção para a necessidade de desenvolvimento de estudos metodológicos que permitam melhor avaliar os impactos da não participação sobre as medidas de efeito em estudos de coorte. As estratégias metodológicas para a avaliação dessa questão incluem a comparação entre participantes e não participantes em relação a dados disponíveis em bases de dados administrativos (sexo, idade, local de residência), o seguimento passivo de todos os elegíveis para a avaliação da

mortalidade, a aplicação de questionários resumidos a uma amostra de não participantes, a comparação entre os participantes que respondem precocemente ao recrutamento com aqueles que respondem tardiamente, análises de sensibilidade e estudos de simulação⁹. Todas essas estratégias quando aplicadas isoladamente apresentam limitações, mas seu uso combinado pode permitir alcançar uma melhor compreensão do impacto da participação e dos fatores associados sobre a validade interna de estudos de coorte. No Brasil, em particular, a despeito do crescente interesse no desenvolvimento de estudos de coorte¹², pouco se conhece sobre os fatores

que determinam a participação e permanência em estudos longitudinais. A crescente disponibilidade de bases de dados epidemiológicas e administrativas, aliada ao uso de técnicas de *linkage* de bases de dados, representa uma alternativa factível para a avaliação de impactos de perdas em estudos de coorte no Brasil¹³.

Concluindo, o presente estudo reforça a necessidade da melhor avaliação dos impactos da não participação sobre a validade em estudos de coorte. Os resultados sugerem que a combinação da não resposta em diferentes estágios de um estudo de coorte pode causar viés de seleção.

REFERÊNCIAS

1. Kelsey J, Whittemore A, Evans A, Thompson W. *Methods in Observational Epidemiology* 1996. 2nd ed. New York: Editora Oxford, 1996.
2. Greenland S. Response and follow-up bias in cohort studies. *Am J Epidemiol*. 1977;106(3):184-7.
3. Jackson R, Chambless LE, Yang K, Byrne T, Watson R, Folsom A, et al. Differences between respondents and nonrespondents in a multicenter community-based study vary by gender ethnicity. The Atherosclerosis Risk in Communities (ARIC) Study Investigators. *J Clin Epidemiol*. 1996;49(12):1441-6.
4. Goldberg M, Chastang JF, Leclerc A, Zins M, Bonenfant S, Bugel I, et al. Socioeconomic, demographic, occupational, and health factors associated with participation in a long-term epidemiologic survey: a prospective study of the French GAZEL cohort and its target population. *Am J Epidemiol*. 2001;154(4):373-84.
5. Harald K, Salomaa V, Jousilahti P, Koskinen S, Vartiainen E. Non-participation and mortality in different socioeconomic groups: the FINRISK population surveys in 1972-92. *J Epidemiol Community Health*. 2007;61(5):449-54.
6. Ferrie JE, Kivimaki M, Singh-Manoux A, Shortt A, Martikainen P, Head J, et al. Non-response to baseline, non-response to follow-up and mortality in the Whitehall II cohort. *Int J Epidemiol*. 2009;38(3):831-7.
7. Heilbrun LK, Nomura A, Stemmermann GN. The effects of nonresponse in a prospective study of cancer. *Am J Epidemiol*. 1982;116(2):353-63.
8. Hernan MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. *Epidemiology*. 2004;15(5):615-25.
9. Stang A. Nonresponse research--an underdeveloped field in epidemiology. *Eur J Epidemiol*. 2003;18(10):929-31.
10. Alonso A, Seguí-Gómez M, de Irala J, Sánchez-Villegas A, Beunza J, Martínez-Gonzalez M. Predictors of follow-up and assessment of selection bias from dropouts using inverse probability weighting in a cohort of university graduates. *Eur J Epidemiol*. 2006;21(5):351-8.
11. Goldberg M, Chastang JF, Zins M, Niedhammer I, Leclerc A. Health problems were the strongest predictors of attrition during follow-up of the GAZEL cohort. *J Clin Epidemiol*. 2006;59(11):1213-21.
12. Carvalho MS, Lopes C. Métodos em estudos de coorte. *Rev Bras Epidemiol*. 2005;8(3):234-5.
13. Coutinho ESF, Coeli CM. Acurácia da metodologia de relacionamento probabilístico de registros para identificação de óbitos em estudos de sobrevida. *Cad Saude Publica*. 2006;22(10):2249-52.

Recebido em: 11/03/2011
Aprovado em: 12/03/2012