

Revista de Saúde Pública

Journal of Public Health

Validade de histórias ocupacionais de informantes secundários

Validity of occupational histories from proxy respondents

Ricardo Cordeiro

Departamento de Saúde Pública da Faculdade de Medicina de Botucatu da Universidade Estadual Paulista. Botucatu, SP, Brasil

Validade de histórias ocupacionais de informantes secundários*

Validity of occupational histories from proxy respondents

Ricardo Cordeiro

Departamento de Saúde Pública da Faculdade de Medicina de Botucatu da Universidade Estadual Paulista. Botucatu, SP, Brasil

Descritores

Saúde ocupacional#. Validade#. Trabalho, história#. Entrevistas, utilização#. Sensibilidade e especificidade#. Família. Estudos casos-controles. Ocupações. – História ocupacional.

Resumo

Objetivo

Embora sejam freqüentes os estudos epidemiológicos trazendo informações ocupacionais obtidas em entrevistas com familiares de trabalhadores, há poucos trabalhos analisando a validade dessas informações, sendo nenhum deles realizados em países em desenvolvimento. Objetivou-se estudar a validade do uso de informações ocupacionais obtidas de informantes secundários em estudos epidemiológicos, pela análise de concordância entre histórias ocupacionais obtidas em entrevistas independentes com os próprios trabalhadores e seus familiares.

Métodos

A história ocupacional de trabalhadores, residentes no Município de Botucatu, SP, em 1998, foi obtida por entrevistas com os próprios trabalhadores e com os familiares próximos. Calcularam-se a sensibilidade e a especificidade das informações ocupacionais dos familiares, considerando-se “padrão-ouro” a informação do trabalhador. Avaliou-se a concordância dos pares de informação pelo coeficiente Kappa.

Resultados

Foram entrevistados 2.163 pares trabalhador/familiar. A sensibilidade da informação ocupacional dos familiares variou entre 77,5% (IC95%; 64,6% – 90,4%) e 98,9% (97,3% – 100,0%), enquanto a especificidade variou entre 96,9% (96,0% – 97,7%) e 99,9% (99,7% – 100,0%). O coeficiente Kappa para a concordância da informação ocupação principal, segundo as duas fontes, foi de 0,86 (0,85 – 0,88).

Conclusões

A utilização de informações ocupacionais obtidas de informantes secundários, se tratadas como variáveis categóricas, tem validade. Utilizando-se informações relativas a tempos acumulados de trabalho, concluiu-se que informantes secundários subestimam os tempos de exposição quando comparados com os próprios trabalhadores.

Keywords

Occupational health#. Validity#. Work, history#. Interviews, utilization#. Sensibility and specificity#. Family. Workers. – Occupational history.

Abstract

Objective

Occupational studies often use data obtained from proxy respondents. However, few investigations have been conducted on the validity of occupational data provided

Correspondência para/Correspondence to:

Ricardo Cordeiro
Caixa Postal 549
18618-970 Botucatu, SP, Brasil
E-mail: cordeiro@fmb.unesp.br

*Financiado pela Fundação de Amparo à Pesquisa do Estado de São Paulo (Processos 99/03437, 97/12782-9, 97/12783-5, 97/12784-1) e pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (Processo 520605/96-9).

Edição subvencionada pela Fapesp (Processo nº 00/01601-8).

Recebido em 10/2/2000. Reapresentado em 27/6/2000. Aprovado em 31/7/2000.

by workers' surrogates. This study aimed to compare self-reported occupational data to information obtained from next-of-kin, as a contribution to assess the validity of using proxy respondents to obtain information about workers.

Methods

Worker/next-of-kin pairs, residents in Southeastern Brazil, were interviewed separately in 1998 about worker's occupational past history. The concordance, sensitivity, and specificity of proxy reports about workers' occupations were examined comparing to self-reports.

Results

A total of 2.163 worker/next-of-kin pairs were interviewed. The Kappa statistic for the agreement between worker's and next-of-kin's report about the worker's main occupation was 0.86 (CI95%; 0.85 – 0.88). The sensitivity of proxy reports on occupations ranged from 77.5% (64.6% – 90.4%) to 98.9% (97.3% – 100.0%), and specificity ranged from 96.9% (96.0% – 97.7%) to 99.9% (99.7% – 100.0%).

Conclusions

These are encouraging findings for the use of proxy respondents in occupational studies when occupations are considered as categorical variables. However, caution is required when assessing occupational exposition by means of cumulative work time because next-of-kin may underestimate this information.

INTRODUÇÃO

São freqüentes os estudos epidemiológicos que analisam informações obtidas em inquéritos de amostras feitos por intermédio de entrevistas domiciliares, geralmente realizadas durante o dia. Adultos jovens, principalmente do sexo masculino, costumam ser sub-representados em tais inquéritos, uma vez que nesse horário predominam donas de casa, crianças e idosos nos domicílios residenciais.⁹ A superação dessa dificuldade pode ser feita com a realização de visitas domiciliares repetidas, em horário noturno, ou aos finais de semana. Entretanto, esse procedimento aumenta consideravelmente o custo, a necessidade de supervisão de campo e o tempo gasto na pesquisa. Por isso, é comum, em situações como essa, obter-se de familiares informações sobre moradores ausentes na ocasião da visita domiciliar. Situação semelhante ocorre em estudos epidemiológicos que analisam informações relativas a indivíduos já falecidos ou mentalmente desabilitados, recorrendo-se freqüentemente a entrevistas com seus familiares para a obtenção de informações.¹⁵

Entretanto, ao se obterem informações sobre um indivíduo, entrevistando um terceiro, ainda que esse seja um familiar bastante próximo, introduz-se mais uma fonte de erro no método, implicando aumento da variância dos estimadores utilizados e conseqüente diminuição do poder do estudo. Há ainda a possibilidade de introduzir-se um viés na análise, se houver erro sistemático nas respostas dos familiares entrevistados.

As dificuldades apontadas são particularmente sen-

tidas em estudos de saúde dos trabalhadores, nos quais informações sobre exposições progressas a riscos ocupacionais são fundamentais. Por razões óbvias, a população empregada geralmente encontra-se fora de casa durante o dia, e o informante secundário, muitas vezes, tem dificuldade em referir a história ocupacional de seu familiar, que pode ter se iniciado há 10, 20, 30 ou mais anos.

Há poucos trabalhos na literatura estudando a validade de dados ocupacionais obtidos de informantes secundários;² nenhum deles realizados em países em desenvolvimento. A maioria desses estudos relata ser apenas razoável a concordância entre informante primário e secundário, recomendando bastante cautela nas análises. Com o objetivo de contribuir para o estudo da validade do uso de informações ocupacionais obtidas de informantes secundários em estudos epidemiológicos, o presente trabalho analisa a concordância entre histórias ocupacionais obtidas em entrevistas independentes com os próprios trabalhadores e com seus familiares.

MÉTODOS

O presente trabalho insere-se em um amplo estudo caso-controle, realizado, em 1997/1998, na cidade de Botucatu, região Centro-oeste do Estado de São Paulo, visando estimar riscos de óbito associados a exposições ocupacionais. Os critérios de eleição para os casos foram: (1) ser morador da cidade e (2) ter falecido em 1997 com idade maior ou igual a 10 anos. Foram estudados 615 casos. A cada caso foram emparelhados 4 moradores vivos, segundo sexo e faixa etária (idade do caso \pm 1 ano). Esses foram escolhi-

dos aleatoriamente a partir de um cadastro anteriormente construído em um estudo que estimou indicadores de mortalidade para a população economicamente ativa do município.⁶ Foram estudados 2.460 controles. No presente estudo, é analisada a concordância das histórias ocupacionais desses indivíduos obtidas dos próprios controles e de seus familiares.

Após a alocação dos potenciais controles, esses foram visitados por entrevistadores, em seus domicílios, no segundo semestre de 1998. Aqueles não localizados em três visitas, em dias e horários diferentes, foram excluídos do estudo, sendo substituídos por outros alocados de modo análogo. Durante a visita, apresentavam-se os objetivos da pesquisa e solicitava-se o consentimento do morador para a sua inclusão no estudo como controle. Após seu consentimento, em entrevista em ambiente isolado sem a presença de nenhum outro morador, a história ocupacional do controle – desde os 10 anos de idade até a última ocupação exercida – era recuperada, anotando-se pormenorizadamente a descrição das ocupações e seus respectivos tempos de exercício.

Ao final da entrevista, solicitava-se ao controle que indicasse o familiar, residente no mesmo domicílio, que melhor soubesse relatar sua história ocu-

pacional. Após ser informado dos objetivos do estudo e consentir em participar do mesmo, o familiar indicado era entrevistado em ambiente isolado e sem a presença do controle. Nessa segunda entrevista, era pesquisada a história ocupacional do controle de forma idêntica à primeira entrevista. Controles que moravam sós ou com parentes que de antemão referiam não conhecer a sua história ocupacional foram excluídos do estudo.

Definiu-se a ocupação principal do controle como aquela exercida por mais tempo durante sua vida ocupacional. Cada uma das ocupações identificadas nas histórias ocupacionais de ambas as fontes foi codificada segundo a classificação de ocupações de Rumel,¹³ que as classifica em 25 subgrupos, que podem ser conferidos na Tabela 1. Para categorizar controles que nunca trabalharam, acrescentou-se a essa classificação o *Subgrupo 601*: indivíduos que nunca trabalharam.

Após a codificação das ocupações, calcularam-se a especificidade e a sensibilidade da informação ocupação principal fornecida pelo familiar (assumindo-se a informação do controle como “padrão-ouro”), bem como o coeficiente Kappa⁴ para os pares de informação ocupação habitual.

Tabela 1 – Distribuição conjunta do subgrupo de ocupação principal dos controles, segundo controles (vertical) e familiares (horizontal). Botucatu, SP, Brasil, 1998.

	101	102	103	201	301	302	303	304	305	306	307	401	402	403	404	405	406	407	408	409	410	411	412	413	501	601	n	
101	65	0	1	7	0	6	0	0	0	1	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	82	
102	1	84	2	0	0	0	0	0	0	1	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	91
103	0	0	80	0	0	0	0	0	2	4	0	0	1	0	0	2	0	0	0	0	0	0	0	0	0	0	1	90
201	3	0	1	465	0	6	4	0	24	0	0	0	1	6	0	1	0	0	0	0	0	0	0	0	6	15	532	
301	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	
302	1	1	0	3	0	145	0	0	6	1	0	0	7	0	0	0	4	0	0	0	0	0	0	0	1	7	176	
303	0	0	0	5	0	3	96	0	2	0	0	0	0	1	1	0	1	0	0	0	0	0	0	1	0	0	110	
304	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	
305	0	0	4	18	0	6	3	0	240	1	0	0	1	2	0	0	2	0	0	0	0	0	0	0	0	0	7	284
306	2	2	2	0	0	2	1	2	1	104	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	8	127
307	0	0	0	0	0	0	0	0	0	0	13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	13	
401	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
402	0	0	1	3	0	0	1	0	0	1	0	0	125	2	0	1	1	0	0	0	0	0	0	0	0	0	0	135
403	0	0	0	3	0	0	0	0	3	0	0	0	3	109	0	0	0	0	0	0	0	0	0	0	0	0	0	118
404	0	0	0	0	0	0	0	0	0	0	0	0	0	1	9	0	0	0	0	0	0	0	0	0	0	0	1	11
405	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	15	0	0	0	0	0	0	0	0	0	0	2	17
406	0	0	0	4	0	0	0	0	1	2	0	0	0	0	0	0	82	0	0	0	0	0	0	0	0	0	1	90
407	0	0	0	0	0	0	0	0	0	0	0	0	3	0	1	0	0	30	0	0	0	0	0	0	0	0	0	34
408	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
409	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	
410	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	4	0	0	0	0	0	6	
411	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0	0	0	7	
412	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	1	4	
413	0	0	0	2	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	4	0	0	8	
501	0	0	0	4	0	0	0	0	0	0	0	0	1	0	0	0	0	4	0	0	0	0	0	0	0	31	0	40
601	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	174	176
n	72	87	91	516	1	168	105	12	279	115	14	0	144	125	11	18	96	34	0	1	4	7	3	5	38	217	2.163	

Coeficiente Kappa: 0,86 (IC95% 0,85 – 0,88)

Concordância simples: 87,7%

101: administrativos, 102: cientistas e artistas, 103: técnicos, 201: agropecuários e produção extrativa vegetal e animal, 301: serviços de eletricidade, 302: comércio, 303: transporte, 304: comunicações, 305: prestação de serviços, 306: escritório, 307: segurança pública e defesa nacional, 401: extrativa mineral, 402: mecânica, metalúrgica e material elétrico, 403: construção civil, 404: alimentação, bebidas e cigarros, 405: têxtil e cortiça de couro, 406: vestuário e artefatos de couro, 407: madeira e móveis, 408: papel e papelão, 409: gráficos, 410: cerâmica e vidro, 411: jóias e instrumento de precisão, 412: artefatos de borracha e plástico, 413: movimentos de cargas, 501: trabalhadores braçais, 601: indivíduos que nunca trabalharam.

A concordância entre as duas fontes de informação em relação aos tempos acumulados de trabalho durante a vida ocupacional (variável contínua) foi avaliada por intermédio do coeficiente de correlação entre esses dois conjuntos de dados, bem como ajustando-se um modelo de regressão linear simples com a reta ajustada passando pela origem (sem intercepto), tendo como variável dependente o tempo acumulado de trabalho segundo o familiar e como variável independente o tempo acumulado de trabalho segundo o trabalhador. Complementarmente, entre os pares estudados, analisou-se a distribuição da medida:

$$\ln \left(\frac{\sum_{\text{todas ocupações}} \text{tempo de trabalho segundo o controle}}{\sum_{\text{todas ocupações}} \text{tempo de trabalho segundo o familiar}} \right)$$

Ela expressa o logaritmo natural da razão entre a somatória dos tempos de trabalho acumulados nos diferentes subgrupos de ocupação segundo os controles e seus familiares.

Compararam-se as médias dos tempos acumulados de trabalho dos trabalhadores, segundo as duas fontes de informação, realizando-se teste *t* de Student para médias de populações independentes. Testaram-se a independência entre o informante (trabalhador ou familiar) e as ocupações principais dos trabalhadores classificadas segundo Rumel,¹³ utilizando-se tabelas de dupla entrada e estatísticas de associação “odds ratio”.

RESULTADOS

Dentre os 2.460 controles alocados, 297 (12,1%) viviam sós ou com parentes que não foram capazes de informar a história ocupacional completa do controle, sendo excluídos da análise. Foram então comparadas as histórias ocupacionais obtidas dos restantes 2.163 controles e seus familiares. Dentre os con-

troles, 1.245 (57,6%) eram do sexo masculino e 918 (42,4%) do feminino. A idade média dos controles foi de 67,3 anos, sendo de 63,6 anos entre os homens e 71,9 anos entre as mulheres. A Figura 1 ilustra a distribuição de sexo e idade dos controles estudados. Dentre os familiares, 1.051 (48,6%) eram cônjuges dos controles, 590 (27,3%) eram filhos adultos e 522 (24,0%) outros parentes adultos.

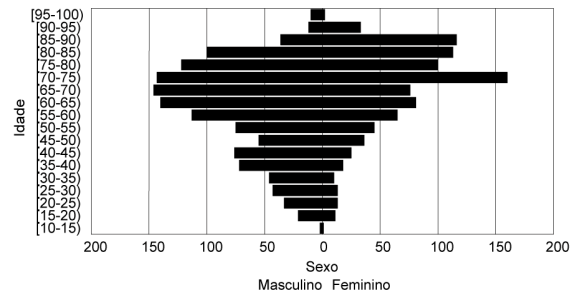


Figura 1 – Distribuição dos trabalhadores estudados segundo sexo e idade. Botucatu, SP, Brasil. 1998.

Codificando-se a ocupação principal dos trabalhadores, segundo subgrupos da classificação de Rumel,¹³ verificou-se que a concordância simples entre os pares controle-parente foi de 87,7%. A essa concordância correspondeu um coeficiente Kappa⁴ de 0,86 (IC95%; 0,85 – 0,88). A Tabela 1 apresenta a distribuição conjunta da ocupação principal do controle, classificada conforme as duas fontes de informação. O coeficiente Kappa⁴ não mostrou alteração importante segundo o grau de parentesco do familiar.

Na Tabela 2 são apresentadas, para os subgrupos ocupacionais mais prevalentes na população estudada, a especificidade e a sensibilidade da informação ocupação principal obtida do familiar e codificada no âmbito do subgrupo, segundo a classificação de Rumel.¹³ Observa-se que a sensibilidade variou entre 77,5% (IC95%; 64,6% – 90,4%) e 98,9% (97,3% –

Tabela 2 – Sensibilidade e especificidade das informações subgrupo de ocupação principal. Botucatu, SP, Brasil. 1998.

Subgrupo ocupacional*	Sensibilidade (IC95%) %	Especificidade (IC95%) %
Administrativos	79,3 (70,5-88,0)	99,7 (99,4-99,9)
Cientistas	92,3 (86,8-97,8)	99,9 (99,7-100,0)
Técnicos	88,9 (82,4-95,4)	99,5 (99,2-99,8)
Lavradores	87,4 (84,6-90,2)	96,9 (96,0-97,7)
Comerciários	82,4 (76,8-88,0)	98,8 (98,4-99,3)
Motoristas	87,3 (81,0-93,5)	99,6 (99,3-99,8)
Prestação de serviços	84,5 (80,3-88,7)	97,9 (97,3-98,6)
Escritório	81,9 (75,2-88,6)	99,5 (99,1-99,8)
Metalúrgicos	92,6 (88,2-97,0)	99,1 (98,6-99,5)
Construção civil	92,4 (87,6-97,2)	99,2 (98,8-99,6)
Têxteis	88,2 (72,9-100,0)	99,9 (99,7-100,0)
Vestuário	91,1 (85,2-97,0)	99,3 (99,0-99,7)
Madeira	88,2 (77,4-99,1)	99,8 (99,6-100,0)
Braçais	77,5 (64,6-90,4)	99,7 (99,4-99,9)
Sem ocupação	98,9 (97,3-100,0)	97,8 (97,2-98,5)

*Segundo Rumel¹³ (1985).

100,0%), ao passo que a especificidade variou entre 96,9% (96,0% – 97,7%) e 99,9% (99,7% – 100,0%).

O coeficiente de correlação entre os tempos acumulados de trabalho, segundo as duas fontes de informação, foi de 0,91. A Tabela 3 apresenta estatísticas obtidas pelo ajuste do modelo de regressão linear. Observou-se um bom ajuste ($r^2=0,95$, $F=41.124$), chamando a atenção a estimativa do coeficiente angular ($b=0,92$), indicando que os valores do tempo acumulado de trabalho que os familiares atribuíram aos controles foram, em média, 92% menores que os atribuídos pelos próprios controles. A média do tempo acumulado de trabalho dos controles foi 33,9 anos, segundo eles próprios, e 31,5 anos, segundo seus familiares, valores significativamente diferentes entre si ($p=0,0001$).

A Figura 2 apresenta um diagrama de barras em que estão dispostos os logaritmos naturais da razão entre os tempos acumulados de trabalho em todas as ocupações, segundo os controles e segundo seus familiares, mostrando um predomínio de valores positivos.

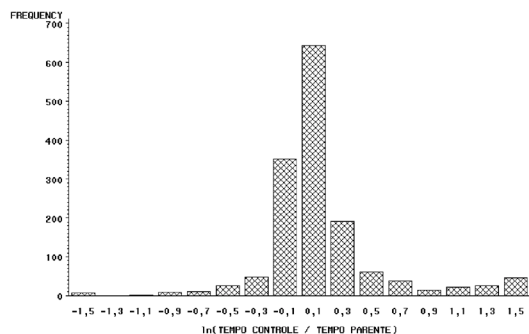


Figura 2 – Distribuição do log da razão* (soma dos tempos acumulados de trabalho segundo o controle)/(soma dos tempos acumulados de trabalho segundo o familiar).

*Foram excluídos da figura os valores $\log(\text{razão}) = 0$

Ao se testarem a independência entre o informante (controle ou familiar) e as ocupações dos controles classificadas segundo subgrupos, não se evidenciou, com nível de significância $=0,01$, nenhuma associação entre informante e qualquer ocupação.

DISCUSSÃO

A concordância simples não expressa a real concordância entre pares de informação, uma vez que respos-

tas ao acaso de indivíduos totalmente não relacionados podem coincidir, sem que isso signifique concordância de fato.¹ Por isso, para medir a concordância entre as informações dos pares controle-parente, quanto a ocupações e tempos de trabalho, utilizou-se o coeficiente Kappa,⁴ função da frequência de pares concordantes em classificações independentes, segundo variáveis categóricas exclusivas e exaustivas, sendo corrigidas as concordâncias ao acaso.⁷ Além do coeficiente Kappa,⁴ avaliou-se também a correspondência das informações dos familiares às dos controles, calculando-se a sensibilidade e a especificidade da informação subgrupo de ocupação principal segundo familiares. Ambas são importantes ao se interpretar o potencial impacto de erros de classificação sobre estimativas de risco relativo em estudos caso-controle.⁵

Codificando-se a ocupação principal em relação ao subgrupo da classificação de Rumel,¹³ tanto a concordância simples quanto o coeficiente Kappa⁴ assumiram valores superiores a 0,85. Os 10% a 15% de discordância entre as informações não parecem apresentar nenhuma tendência preferencial (como, por exemplo, a “promoção” da ocupação a um subgrupo ocupacional de maior *status* social ou financeiro, conforme outros relatos¹²), dada a não verificação de associação entre o tipo de informante (controle ou familiar) e as ocupações referidas.

Landis & Koch¹⁰ e Wang et al¹⁵ sugerem a seguinte classificação como referência para a descrição do grau de concordância entre informações, medido pelo coeficiente de Kappa:⁴

Coeficiente de Kappa	Força de concordância
< 0,00	pobre
0,00 – 0,20	pequena
0,21 – 0,40	razoável
0,41 – 0,60	moderada
0,61 – 0,80	substancial
0,81 – 1,00	quase perfeita

Assim, os resultados encontrados sugerem ser muito boa a concordância das informações ocupacionais dos familiares em relação aos controles, situando-se num patamar superior ao geralmente relatado em outros estudos.

Entretanto, ao se estudar a variável contínua *tempo acumulado de trabalho*, embora a concordância continue elevada (conforme sugerem o coeficiente de corre-

Tabela 3 – Estimativas e estatísticas obtidas no ajuste do modelo de regressão: Tempo de Trabalho segundo Familiar_i = b Tempo de Trabalho segundo Trabalhador. Botucatu, SP, Brasil. 1998.

Parâmetro	Estimativa	Desvio-padrão	t para H ₀ : parâmetro = 0	Valor p
Declividade	0,92	0,004	202,79	<0,0001

$r^2 = 0,95$
 $F = 41124$

lação entre as duas fontes de informação e as estatísticas r^2 e F do ajuste linear obtido), as estimativas feitas pelos familiares tenderam a ser menores que aquelas dos próprios controles. Desse modo, analisando-se as ocupações não como variáveis categóricas, mas como tempos acumulados de exercício profissional (variável relacionada aos níveis de exposição a riscos ocupacionais), os familiares tenderam a subestimar os tempos de trabalho dos controles. Tal tendência está expressa na Figura 2, com a distribuição do logaritmo natural da razão entre a somatória dos tempos de trabalho acumulados nos diferentes subgrupos de ocupação, segundo controles e segundo seus familiares. Observa-se a predominância de valores positivos dessa razão, que representa os pares em que o controle atribuiu maior tempo de trabalho que o familiar. Para explicar essa tendência, uma hipótese plausível seria o desconhecimento da história ocupacional remota do controle, principalmente por parte de seus cônjuges e filhos, que, em geral, o conheceram após esse já ter iniciado sua vida ocupacional há algum tempo. Resultado semelhante foi encontrado por Lerchen & Samet,¹¹ que estudaram em 1986 a concordância entre pares de informações ocupacionais de 80 pacientes que faleceram de câncer de pulmão entre 1980 e 1984. As informações foram fornecidas em vida, pelos próprios pacientes, e por suas esposas, após o óbito. A média do tempo acumulado de trabalho segundo as esposas foi de 33,5 anos, valor significativamente menor que o referido pelos próprios pacientes: 41,0 anos. Com relação às diferentes ocupações exercidas durante toda a vida, a concordância bruta foi de 41,0% nesse estudo, de acordo com uma classificação de 3 dígitos.¹⁴

O presente estudo não é uma análise de validade de dados, uma vez que não se está verificando o grau de correção das informações ocupacionais obtidas, ou seja, o quanto essas informações expressam a real história ocupacional dos trabalhadores em questão. Não apenas as informações obtidas dos familiares, mas também as referidas pelos próprios controles, podem não ser acuradas. Esse é basicamente um estudo de concordância entre pares de informações. Entretanto, conforme estudos clássicos sobre concordância,^{3,8} a correspondência entre diferentes observadores de um mesmo fenômeno sugere algum grau de

validade das medidas. Assim, o alto grau de concordância obtido no presente estudo sugere que as respostas de controles e familiares foram acuradas e não apenas consistentes. Se controles e familiares não estivessem de acordo em seus relatos ocupacionais, seria pouco provável que os erros se dessem na mesma extensão e direção para ambos.⁹ Para a maioria dos estudos epidemiológicos, o critério de exata concordância, além de ser inatingível, é também prescindível, uma vez que uma alta concordância associada à ausência de erro sistemático são suficientes para a demonstração de associações entre determinantes e fenômenos do processo saúde/doença.⁹ Em estudos epidemiológicos visando estimar riscos associados a exposições ocupacionais, a acurácia parece ser suficiente para evidenciar tais associações, se essas forem de magnitude suficientemente grande. Muito embora, se os erros forem aleatoriamente distribuídos, as associações observadas serão com maior probabilidade subestimadas.⁷

Entretanto, quando se analisaram as informações relativas aos tempos de exercício profissional em diferentes ocupações, as estimativas fornecidas pelos familiares foram sistematicamente menores – embora em pequena magnitude – do que aquelas fornecidas pelos próprios controles. Tal fato pode implicar viés, se esses tempos forem utilizados como quantificadores de exposição em estudos comparativos como, por exemplo, estudos de caso-controle em que os casos são identificados após o falecimento. Nessa situação, as informações sobre os casos originam-se de informantes secundários, enquanto as dos controles originam-se de informantes primários. Numa situação como essa, o estudo tenderia a subestimar riscos ocupacionais e superestimar fatores de proteção.

Concluindo, os resultados sugerem que a utilização de informações ocupacionais obtidas de familiares próximos como informantes secundários, se tratados como variáveis categóricas, têm validade. Já na utilização de informações relativas a tempos acumulados de trabalho, os dados sugerem que informantes secundários subestimam os tempos de exposição quando comparados com os próprios trabalhadores.

REFERÊNCIAS

1. Armitage P, Berry G. *Statistical methods in medical research*. Oxford: Blackwell; 1994.
2. Boyle CA, Brann EA. Proxy respondents and the validity of occupational and other exposure data. *Am J Epidemiol* 1992;136:712-21.
3. Campbell DT, Fiske DW. Convergent and discriminant validation by multitrait-multimethod matrix. *Psychol Bull* 1959;57:81-105.
4. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960;20:37-46.
5. Copeland KT, Checkoway H, McMichael AJ, Holbrook RH. Bias due to misclassification in the estimation of relative risk. *Am J Epidemiol* 1977;105:488-95.

6. Cordeiro R, Olivencia PER, Cardoso CF, Cortez DB, Kakinami E et al. Desigualdade de indicadores de mortalidade no Sudeste do Brasil. *Rev Saúde Pública* 1999;33:593-601.
7. Fleiss JL. *Statistical methods for rates and proportions*. New York: John Wiley; 1973.
8. Heise D, Bohrnsted G. Validity and reliability. In: Borgatta EF, Bohrnstedt G, editors. *Sociological methodology*. San Francisco: Jossey-Bass; 1970. p. 46-59.
9. Kolonel LN, Hirohata T, Nomura AMY. Adequacy of survey data collected from substitute respondents. *Am J Epidemiol* 1977;106:476-84.
10. Landis JR, Koch GG. The measurement of observer agreement for categorial data. *Biometrics* 1977;33:159-74.
11. Lerchen ML, Samet JM. An assessment of the validity of questionnaire responses provided by a surviving spouse. *Am J Epidemiol* 1986;123:481-9.
12. Rogot E, Reid DD. The validity of data from next-of-kin in studies of mortality among migrants. *Int J Epidemiol* 1975;4:51-4.
13. Rumel D. Indicadores de mortalidade por categoria ocupacional e nível social [Dissertação]. São Paulo: Faculdade de Saúde Pública da Universidade de São Paulo; 1985.
14. United States Department of Commerce, Office of Federal Statistical Policy and Standards. *Standard occupational classification manual*. Washington (DC): US Government Printing Office; 1972.
15. Wang F, Semchuk KM, Love EJ. Reliability of environmental and occupational exposure data provided by surrogate respondents in a case-control study of Parkinson's disease. *J Clin Epidemiol* 1994;47:797-807.