



Mortalidade adulta por nível de escolaridade em São Paulo: análise comparativa a partir de diferentes estratégias metodológicas*

Mirian Martins Ribeiro**
Cassio Maldonado Turra***
Cristine Campos de Xavier Pinto****

Neste artigo, são estimados os diferenciais educacionais de mortalidade de adultos residentes em São Paulo. É realizada uma análise comparativa de estimativas a partir de dados do Censo 2010 e do Sistema de Informação de Mortalidade (SIM) – Datasus e de três formas distintas de mensuração da escolaridade: registrada no SIM; declarada no Censo para o responsável pelo domicílio; e imputada estatisticamente no Censo para indivíduos que morreram. Para as imputações da escolaridade, utilizou-se o método de Dempester (1977), que propõe o uso do algoritmo esperança-maximização (algoritmo E-M) para lidar com dados faltantes. Foram considerados três níveis de escolaridade (baixo, médio e alto) e estimadas as taxas de mortalidade com base em modelos Poisson. Os resultados indicam que a obtenção de escolaridade pode reduzir em até 77% as taxas de mortalidade entre 25 e 59 anos de idade. Além disso, em um país em que a população tem baixa escolaridade, obter ensino médio representa um ganho significativo do ponto de vista da sobrevivência adulta (cerca de 50%). Encontraram-se padrões de mortalidade por escolaridade semelhantes para as estimativas obtidas com dados registrados no SIM e aqueles imputados no Censo 2010. Além disso, a análise sugere que estimativas assumindo a escolaridade do responsável pelo domicílio resultam em diferenciais de mortalidade atípicos, provavelmente distorcidos pela transição de educação no Brasil. Espera-se que o modelo de imputação proposto aqui possa ser utilizado em futuras análises dos dados de mortalidade a partir do Censo 2010.

Palavras-chave: Mortalidade adulta. Dados faltantes. Imputação. Diferenciais socioeconômicos em mortalidade. Brasil. Escolaridade.

* O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), código de financiamento 001. Cassio M. Turra agradece, também, o apoio recebido do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). Agradecemos à professora Simone Wajnman, além de dois pareceristas anônimos, pelos comentários recebidos em versões anteriores desse trabalho.

** Instituto de Ciências Sociais Aplicadas, Universidade Federal de Ouro Preto (Ufop), Ouro Preto-MG, Brasil (miriannina@gmail.com; <https://orcid.org/0000-0002-4256-220X>).

*** Centro de Desenvolvimento e Planejamento Regional (Cedeplar), Universidade Federal de Minas Gerais (UFMG), Belo Horizonte-MG, Brasil (turra@cedeplar.ufmg.br; <https://orcid.org/0000-0003-4051-3567>).

**** Escola de Economia de São Paulo (EESP), Fundação Getúlio Vargas (FGV), São Paulo-SP, Brasil (cristine.pinto@fgv.br; <https://orcid.org/0000-0003-2955-9759>).

Introdução

O objetivo deste artigo é estimar e comparar os diferenciais educacionais de mortalidade de adultos residentes em São Paulo, a partir de diferentes bases de dados e métodos. Em razão da inexistência de uma base de dados com alta qualidade de declaração da escolaridade para todas as unidades da federação ou municípios, que permita conhecer os diferenciais socioeconômicos de mortalidade no Brasil, no nível individual, oferecemos uma análise comparativa de estimativas de mortalidade geradas a partir de três diferentes metodologias: dados de óbitos do censo demográfico assumindo que a escolaridade dos mortos é igual àquela declarada para o responsável pelo domicílio; dados de óbitos do censo demográfico com escolaridade dos mortos imputada por métodos estatísticos; e dados de óbitos por escolaridade extraídos do Sistema de Informação de Mortalidade (SIM)/Datasus. Em um país marcado por inúmeras formas de desigualdade que se expressam nas condições de saúde e de mortalidade de sua população, bem como pela necessidade de expansão e aprimoramento de suas bases de dados, a busca pelos verdadeiros níveis e padrões de mortalidade adulta por escolaridade é fundamental para orientar políticas públicas.

Apesar de suas limitações, a questão sobre óbitos ocorridos nos domicílios, incluída no Censo de 2010, gerou novas possibilidades de investigação sobre os padrões de mortalidade no Brasil. Trata-se de uma alternativa às estimativas calculadas com os óbitos registrados no Sistema de Informações em Mortalidade (SIM), que tem sido a principal fonte de dados para os estudos de mortalidade no país, mas seus registros não estão livres de erros de cobertura e, principalmente, erros de conteúdo. Portanto, tais informações merecem ser confrontadas com fontes alternativas. Os óbitos declarados no Censo 2010 têm sido gradualmente analisados por pesquisadores brasileiros. Estudos recentes investigaram o papel das características socioeconômicas dos domicílios na mortalidade jovem (PEREIRA; QUEIROZ, 2016), o diferencial de mortalidade adulta segundo a escolaridade do responsável pelo domicílio (RIBEIRO, 2016; SILVA *et al.*, 2016) e o diferencial de mortalidade por renda (MORAIS, 2019).

O presente estudo contribui de pelo menos três formas para a literatura existente. A primeira inovação é a investigação dos diferenciais educacionais de mortalidade adulta com dados do Censo 2010, a partir da escolaridade do indivíduo que morreu, em vez do uso de *proxies* domiciliares. Para isso, propõe-se a imputação da informação de escolaridade faltante usando a técnica *expectation-maximization* (algoritmo *E-M*) e o modelo de regressão multivariado normal. A amostra de São Paulo, que é a maior no país, foi selecionada para uma primeira exploração da consistência das imputações e das estimativas de mortalidade. Os mesmos métodos poderão ser reproduzidos para estudos de diferenciais educacionais de mortalidade para todo o Brasil, considerando a heterogeneidade entre os estados. Além disso, o estudo contribui para o debate sobre o verdadeiro gradiente educacional em mortalidade adulta, ao comparar a estratégia metodológica de imputação proposta com taxas estimadas a partir da escolaridade do responsável pelo domicílio, método utilizado

em pelo menos um trabalho anterior (SILVA *et al.*, 2016). A análise também é inovadora ao comparar as estimativas de mortalidade com base em dados censitários com taxas calculadas a partir dos óbitos registrados no SIM. Para isso, foram utilizadas informações de 2013, ano em que a proporção de dados de escolaridade faltantes se apresentou relativamente baixa no SIM, seguindo uma tendência de declínio desde 2010.

Os resultados indicam que educação e mortalidade adulta são inversamente relacionadas no Brasil, confirmando trabalhos para inúmeros outros países. No entanto, as estimativas variam em função da definição de escolaridade utilizada. O emprego de informações domiciliares como *proxy* para características individuais pode resultar em estimativas inconsistentes. O método de imputação alternativo aqui proposto gerou, de forma independente, padrões de mortalidade consistentes com as estimativas do SIM, reforçando sua relevância para futuras análises.

Escolaridade e mortalidade adulta

Abordagens teóricas principais

Segundo Crimmins (1993), os estudos demográficos carecem de teorias unificadas sobre os diferenciais socioeconômicos em mortalidade. As abordagens explicativas mais discutidas são os modelos econômicos e psicossociais (PRESTON; TAUBMAN, 1994). Os modelos econômicos, considerados mais sistematizados para relacionar saúde e condições socioeconômicas (PRESTON; TAUBMAN, 1994), são inspirados no modelo de produção de saúde de Grossman (1972). Nessa abordagem, os indivíduos tomam decisões que ajudam a formar seu estoque de saúde. A morte ocorre quando o estoque atinge um valor abaixo do nível crítico. Há duas funções envolvidas nesse processo: a função orçamentária e a de produção de saúde. Essa última indica como o estoque de saúde varia de acordo com diferentes insumos, incluindo a quantidade consumida de bens de saúde, a quantidade de conhecimento ou tecnologia médica disponível para a população, a disponibilidade de serviços médicos, a dotação genética e as características do ambiente em que os indivíduos se encontram inseridos. O consumo de bens de saúde é determinado por escolhas individuais, enquanto os demais insumos são definidos exogenamente (ROSENZWEIG; SCHULTZ, 1983; PRESTON; TAUBMAN, 1994). Como a educação afeta o nível de capital humano, indivíduos com escolarização mais elevada têm maior renda, o que lhes assegura maior dotação de recursos para o consumo de bens e serviços relacionados à produção de saúde: alimentos saudáveis; maiores espaços para viver; áreas menos poluídas para residirem; serviços de saúde de melhor qualidade; e outros. Além disso, a educação tem um efeito direto sobre os níveis de saúde ao longo do ciclo de vida, pois melhora a capacidade de discernimento dos indivíduos e a forma como processam informações que afetam suas condições de saúde e seus riscos de morte (PRESTON; TAUBMAN, 1994; SICKLES; TAUBMAN, 1997; GROSSMAN, 2008).

Os preços dos produtos e serviços de saúde também são centrais nos modelos econômicos, pois influenciam os diferenciais de mortalidade por meio de pelo menos dois mecanismos. De um lado, ainda que se assuma que os preços de mercado são iguais para todos, o acesso a bens e serviços de saúde por meio de programas públicos ou seguros de saúde pode minimizar as diferenças no consumo entre indivíduos de diferentes níveis de escolaridade. O segundo mecanismo é o custo de oportunidade para se obter cuidado médico, que depende tanto da renda dos indivíduos quanto do tempo de espera pelo cuidado. Mesmo no caso de indivíduos de baixa renda, o custo de oportunidade pode se elevar se o tempo de espera for grande. Algumas políticas públicas de acesso aos serviços de saúde podem mudar os preços relativos (PRESTON; TAUBMAN, 1994), reduzindo os diferenciais de mortalidade.

Uma das principais limitações dos modelos econômicos é não incorporar os fatores ambientais (tais como incidência de doenças transmissíveis), que podem alterar o estoque de saúde da população, e o papel do contexto social na formação das preferências e hábitos individuais (PRESTON; TAUBMAN, 1994; WILLIAMS, 1990).

Em contraposição, as abordagens psicossociais privilegiam os fatores negligenciados pelos modelos econômicos. Embora a literatura que investiga o papel das relações sociais na saúde date do século XIX, o tema foi revitalizado na década de 1970 com o surgimento da teoria do suporte social. O assunto foi impulsionado pelos resultados do relatório de um comitê consultivo do Serviço de Saúde Pública (*Public Health Services*) dos Estados Unidos (US PUBLIC HEALTH SERVICES, 1964), que relacionou morbidade e mortalidade por um conjunto de causas ao consumo de cigarro (HOUSE *et al.*, 1988), sendo o comportamento, o estilo de vida e as atitudes individuais considerados independentes da estrutura social. As teorias psicossociais, ao contrário, partem da premissa de que as estruturas sociais moldam os valores e influenciam o comportamento dos indivíduos com relação às restrições que lhes são impostas (WILLIAMS, 1990).

Mas foi a partir do início dos anos 1990 que as características socioeconômicas foram incorporadas à abordagem psicossocial (PRESTON; TAUBMAN, 1994; HOUSE *et al.*, 1990; WILLIAMS, 1990). Os fatores psicossociais e os serviços de saúde funcionam como mediadores da relação entre fatores socioeconômicos e saúde. A morbidade e a mortalidade estão associadas a fatores psicossociais e comportamentais, como estresse, prática religiosa, estado civil, hábitos de alimentação e exercício físico, tabagismo e suporte social. Um exemplo seria a evidência de que pessoas casadas, sobretudo os homens, exibem menor mortalidade do que as não casadas (ROGER, 1995; LILLARD; WAITE, 1995). Pessoas envolvidas intensamente em preceitos religiosos apresentam menor mortalidade, se comparadas às menos envolvidas, o que é visto como um resultado do papel dos laços e suporte sociais na saúde dos indivíduos (ROGERS, 1996). O estresse, medido por variáveis como desemprego, divórcio, morte de entes queridos, dentre outras, mostra uma relação positiva com a presença de uma série de doenças. Os laços sociais e o controle pessoal também

são vistos como fatores que agem nos processos biológicos e afetam a vulnerabilidade a doenças e mudam as chances de morte (WILLIAMS, 1990).

Os fatores psicossociais, por sua vez, variam em função de características socioeconômicas dos indivíduos, tais como renda, ocupação e escolaridade. A obesidade e a quantidade de cigarros fumados, por exemplo, variam negativamente com a escolaridade. Entre 1974 e 1985, a prevalência de fumantes nos Estados Unidos declinou cinco vezes mais rápido entre pessoas com curso universitário quando comparadas àquelas menos escolarizadas (NCHS, 1981). O tabagismo é apontado como uma das explicações para a expansão dos diferenciais educacionais da mortalidade entre as décadas de 1980 e 1990. Estima-se que cerca de um quarto do aumento dos diferenciais entre as mulheres seja explicado por diferenciais educacionais na prática do tabagismo (MEARA *et al.*, 2008). Pampel *et al.* (2010) realçam a influência dos fatores socioeconômicos, principalmente da escolaridade, no estilo de vida dos indivíduos, com destaque para uso de tabaco, prática de exercícios físicos e alimentação. No caso das bebidas alcoólicas, embora haja uma relação positiva entre seu consumo e escolaridade, o alcoolismo está associado negativamente a esta dimensão (NCHS, 1981).

A educação provê aos indivíduos um estoque de conhecimento que influencia na forma como eles interagem com outros, na intervenção em situações de crise e emergências, na busca por bens e serviços médicos, na prevenção de doenças e na promoção da saúde. Pessoas mais escolarizadas são mais propensas a adotar comportamento saudável, procurar cuidado médico, perceber problemas de saúde e ter uma melhor comunicação com médicos e maior entendimento para seguir suas recomendações. Em contrapartida, indivíduos com menores níveis de educação são mais suscetíveis a características psicossociais negativas, do ponto de vista da saúde, podendo adotar comportamento de risco, apresentar estresse, isolamento social e falta de controle social, que são importantes determinantes da morbidade e mortalidade (ROGERS *et al.*, 1999).

A principal crítica com relação à abordagem psicossocial é a ausência de uma teoria unificada que explique as interações entre os fatores socioeconômicos, psicossociais e os determinantes próximos da mortalidade. Em geral, são apontados padrões e tendências, mas com dados insuficientes para ir além da especulação sobre as relações causais por traz destes achados.

Diferenciais de mortalidade adulta por escolaridade

O crescimento do interesse pelos diferenciais socioeconômicos da mortalidade é datado do início do século XIX. Até meados do século XX houve uma predominância do uso de classes ocupacionais na investigação destes diferenciais (ANTONOVSKY, 1967; ELO, 2009). A partir dos anos 1960, a escolaridade passou a ser uma dimensão socioeconômica mais frequentemente utilizada, por se tratar de uma variável com menor grau de causalção reversa com a mortalidade, ser a menos afetada pela proximidade com a morte, se manter

relativamente estável ao longo da vida adulta e ter papel expressivo na determinação da ocupação, da renda e da riqueza (KITAGAWA; HOUSE, 1963; PRESTON; TAUBMAN, 1994; ELO; PRESTON, 1996; MONTEZ *et al.*, 2012). Quanto à qualidade, trata-se de informação reportada com maior precisão em pesquisas de inquérito (LIBERATOS *et al.*, 1988; MONTEZ *et al.*, 2012), embora não esteja imune a erros de declaração (NEPOMUCENO; TURRA, 2020). Em países desenvolvidos, sobretudo Estados Unidos e Reino Unido, o monitoramento das desigualdades socioeconômicas da mortalidade tem sido sistemático e a educação tem sido um indicador socioeconômico central (CUTLER; LLERAS-MUNEY, 2007).

O estudo pioneiro de Kitagawa e Houser (1973) para os Estados Unidos é um marco na agenda de pesquisa do tema. Usando dados do registro de mortes de 1960, os autores encontraram importantes diferenciais de mortalidade por escolaridade para a população adulta, sobretudo para aquela em idade ativa. Nas décadas seguintes, outros estudos também apontaram tanto um gradiente educacional da mortalidade (ELO; PRESTON, 1996; ROGERS *et al.*, 1999; ROSS *et al.*, 2012) quanto um aumento dos diferenciais ao longo do tempo (PAPPAS *et al.*, 1993; ELO; PRESTON, 1996).

No século XXI, a discussão continua presente e relevante na agenda de pesquisa dos EUA. Até final dos anos 2000, o crescimento do diferencial de mortalidade era atribuído a uma queda mais acelerada da mortalidade entre os mais escolarizados em relação aos menos escolarizados (PAPPAS *et al.*, 1993; PRESTON; ELO, 1995). Mas estudos recentes mostram, também, um crescimento da mortalidade para grupos populacionais menos escolarizados (MONTEZ *et al.*, 2011; HAYWARD *et al.*, 2015). Sasson (2016) indica que, entre 1990 e 2010, a esperança de vida aos 25 anos para as mulheres menos escolarizadas diminuiu cerca de 3 anos, em grande parte devido a um aumento da mortalidade entre as idades de 45 e 64 anos. Em contrapartida, houve melhorias consideráveis na esperança de vida dos mais escolarizados. No mesmo sentido, Case e Deaton (2015) apontam um crescimento da mortalidade entre adultos brancos não hispânicos de 45 a 54 anos, no período de 1999 a 2013, explicado pelo aumento da mortalidade por causas externas, sobretudo intoxicação por drogas, álcool e suicídio. Essas mortes foram concentradas entre indivíduos com educação secundária ou inferior. Olshansky e colegas (2005) chamam a atenção para uma possível queda da esperança de vida nos EUA, em grande parte em função do efeito negativo do crescimento da obesidade.

Em países desenvolvidos da Europa, como Dinamarca, Noruega, Suécia, Finlândia, Inglaterra e Gales, os diferenciais de mortalidade por escolaridade são muito semelhantes, embora o nível da mortalidade varie entre os países (VALKONEN, 1989). Aqueles de economia em transição apresentam os maiores diferenciais de mortalidade por educação (MACKENBACH *et al.*, 2008; SHKOLNIKOV *et al.*, 1998; DENISOVA, 2009; LEINSALU, 2003) e experimentaram um aumento da desigualdade a partir da década de 1990, com o fim do regime socialista (SHKOLNIKOV *et al.*, 1998; LEINSALU, 2003). Também já há inúmeras evidências de diferenciais de mortalidade por escolaridade em outras regiões do mundo, incluindo Ásia (LIANG *et al.*, 2000; SUBRAMANINAN *et al.*, 2006; HURT *et al.*, 2004), África

(BERHANE *et al.*, 2002; WALQUE; FILMER, 2013) e América Latina e Caribe (KOCH *et al.*, 2007; MANZELLI, 2014; SANDOVAL; TURRA, 2015).

Ao longo do tempo, os estudos internacionais se beneficiaram do aprimoramento das bases de dados e metodologias disponíveis. Quatro grandes grupos de abordagens metodológicas têm sido empregados na pesquisa sobre mortalidade por escolaridade. Quando há ausência de informações individuais, é comum o uso de *proxies* no nível domiciliar ou da comunidade, muitas vezes combinadas com técnicas de estimação indireta. Trata-se do conjunto mais frágil de estratégias, como salientaram Kitagawa e Hauser (1973), uma vez que a agregação da informação introduz efeitos de confundimento nas estimativas. Outra possibilidade é o cálculo de medidas demográficas clássicas, como as taxas de mortalidade, quando a informação de escolaridade é coletada nos registros vitais e censos demográficos. Nesse caso, as estimativas baseiam-se em dados individuais, mas dependem da qualidade das informações de escolaridade registradas/reportadas e do grau de consistência entre elas. Uma terceira possibilidade, muito comum nos estudos internacionais mais recentes, é o uso de dados de painel, construídos a partir do pareamento de registros de óbitos com pesquisas domiciliares transversais ou longitudinais. Essa talvez seja a melhor alternativa metodológica, especialmente nos casos em que o pareamento das bases de dados é possível de ser feito a partir de registros de identificação únicos (como o *social security number* nos EUA), propiciando maior precisão. A maior vantagem dessa estratégia é garantir que a informação sobre escolaridade seja a mesma tanto para ocorrências quanto para exposição. Por fim, há estudos que se baseiam em fontes alternativas de dados, incluindo registros administrativos e dados de redes sociais, e estudos longitudinais cujas amostras são limitadas a populações locais ou grupos etários específicos. Embora elucidativos, esses estudos tendem a ter menor validade externa por se referirem a subgrupos populacionais.

A escassez de dados de qualidade no Brasil tem sido um empecilho para o cálculo dos diferenciais de escolaridade para a população adulta. A despeito da limitação de informação, há esforços recentes por meio da combinação de técnicas indiretas, métodos de imputação e informações sobre educação reportadas no nível individual (RENTERIA; TURRA, 2008; RENTERIA, 2010; GUEDES *et al.*, 2011; TURRA; RENTERIA; GUIMARÃES, 2016; NEPOMUCENO; TURRA, 2020; TURRA *et al.*, 2018). Renteria (2010), por exemplo, estimou a mortalidade para a população feminina de 20 a 69 anos no Brasil, a partir de informações de morte e escolaridade das mães dos entrevistados, reportadas na Pesquisa Nacional por Amostra de Domicílios (PNAD) e na Pesquisa sobre Padrões de Vida (PPV), ambas do IBGE. Com base em uma variante do método de orfandade, a autora encontrou taxas decrescentes de mortalidade em relação à escolaridade. Os diferenciais estimados foram expressivos: as mulheres sem escolaridade apresentaram mortalidade até 70% maior do que aquelas com escolaridade entre 1 e 8 anos.

Em outra frente, Silva *et al.* (2016) investigaram os diferenciais de mortalidade por educação para pessoas de 15 a 59 anos e de 60 a 80 anos, usando dados do Censo Demográfico 2010. Os autores pressupuseram que a escolaridade dos mortos era igual

àquela do responsável pelo domicílio, na tentativa de resolver o problema da falta de informações socioeconômicas para os óbitos. Os resultados indicaram probabilidades de morte decrescentes com a escolaridade para o grupo etário mais jovem. Além disso, em domicílios em que os responsáveis tinham ensino superior, a mortalidade estimada foi menor comparativamente aos sem instrução ou com ensino fundamental incompleto. No entanto, para os grupos de idade maiores de 70 anos, o comportamento encontrado foi o contrário do esperado: uma associação positiva entre mortalidade e escolaridade, exceto para a região Sudeste. Esses resultados sugerem que a escolha metodológica dos autores pode ter afetado a qualidade das estimativas, já que ela depende da ausência de mobilidade intergeracional de educação, algo que não ocorreu no Brasil nas últimas décadas.

Apesar dos esforços realizados por pesquisadores brasileiros, é preciso investir na coleta e no relacionamento de bases de dados, na melhoria da qualidade da informação existente e na estimação e comparação de novas medidas de mortalidade. Só assim teremos estimativas mais precisas dos níveis e padrões dos diferenciais de mortalidade por escolaridade. O objetivo deste artigo é exatamente contribuir para este avanço.

Metodologia

Neste estudo, são estimadas as taxas de mortalidade adulta por nível de escolaridade no município de São Paulo, a partir de modelos Poisson, segundo três diferentes conjuntos de informações: dados de óbitos do Censo Demográfico 2010, assumindo que a escolaridade dos mortos é igual àquela declarada para o responsável pelo domicílio; dados de óbitos do Censo Demográfico 2010, com escolaridade dos mortos imputada por métodos estatísticos; e dados de óbitos por escolaridade extraídos do Sistema de Informação de Mortalidade (SIM)/Datasus.

A análise considera a população entre 25 e 59 anos residente no município de São Paulo, por volta do ano de 2010. A adoção de um limite de idade inferior é necessária para garantir que a maior parte dos indivíduos já tenha concluído sua vida escolar e, assim, evitar os efeitos das transições entre níveis e anos de ensino sobre as estimativas de mortalidade. A exclusão da população idosa, acima de 60 anos de idade, tem como objetivo mitigar os problemas na qualidade dos dados do Censo 2010. Nas seções a seguir, detalhamos as fontes de dados e os métodos utilizados em cada um dos conjuntos de informações.

Óbitos no Censo de 2010

Para os dois primeiros conjuntos de estimativas, foram utilizados dados do Censo Demográfico de 2010.¹ Trata-se de uma fonte de dados pouco usual para estudos de mortalidade no Brasil, já que apenas em 2010 o censo incluiu, em seu questionário básico, a ocorrência de morte de residentes em domicílios particulares. Uma das vantagens dos

¹ A extração foi realizada usando o programa *Stata* 14.0, por meio da plataforma *Datazoom*, desenvolvida pelo Departamento de Economia da PUC-Rio.

dados censitários é poder trabalhar com uma mesma fonte de dados para óbitos e população. Outra vantagem é que o censo brasileiro provê uma gama ampla de informações socioeconômicas e geográficas no nível do domicílio e da comunidade. No entanto, nada se sabe sobre as características socioeconômicas individuais dos mortos, apenas seu sexo e idade. Além disso, dentre as desvantagens, destaca-se a possibilidade de sub e/ou superenumeração de óbitos. Como a pergunta é realizada no nível do domicílio, os óbitos ocorridos em domicílios dissolvidos no período de referência do censo ou de domicílios unipessoais não são captados. Há ainda a possibilidade de os óbitos serem reportados mais de uma vez, caso a pessoa falecida tenha transitado em mais de um domicílio no período de referência. O erro de período de referência por parte do respondente também pode gerar subestimação de óbitos. A fim de minimizar parte dos problemas de enumeração, optou-se pela exclusão da população idosa na presente análise, conforme já mencionado, uma vez que, no Brasil, há uma maior concentração de domicílios unipessoais entre os mais velhos (WAJNMAN, 2012).

Imputação da escolaridade nos dados censitários

Em geral, em uma matriz de dados em que as linhas representam unidades de análise e as colunas correspondem a variáveis, as entradas para as quais não se têm dados são consideradas dados faltantes (LITTLE; RUBIN, 2002). No entanto, muitos problemas não envolvem necessariamente perda de dados. Esse é o caso do Censo 2010, visto que a falta da informação da escolaridade para a pessoa que sofreu óbito é fruto do desenho da coleta. Para estimar taxas de mortalidade com base nos dados de óbitos do Censo 2010, é preciso imputar a informação de escolaridade para a totalidade dos mortos. Mas, mesmo nesses casos, as informações podem ser reformuladas e tratadas com métodos originalmente desenvolvidos para lidar com dados faltantes (HEITJAN; RUBIN, 1990; HEITJAN; RUBIN, 1991; SCHAFFER, 1997; ENDERS, 2010; MISTLER; ENDERS, 2012; LITTLE; RHEMUTULLA, 2013).

Uma estratégia mais simples para lidar com essa limitação do Censo 2010 seria simplesmente adotar a escolaridade de algum outro membro do domicílio como sendo a escolaridade do morto. Utilizamos essa estratégia no primeiro conjunto de estimativas, uma vez que ela foi empregada anteriormente na literatura e precisava ser analisada à luz de novos resultados. Como forma de empregar uma categorização de escolaridade que fosse compatível tanto no Censo 2010 quanto no SIM, agrupamos os anos de estudo do responsável pelo domicílio em três categorias: baixa escolaridade (fundamental I completo ou menos); média (fundamental II completo ou ensino médio completo); alta (pelo menos um ano de ensino superior, incluindo superior completo). Essa categorização segue a forma de disponibilização dos dados do SIM, conforme é discutido mais à frente.

A segunda estratégia foi imputar estatisticamente os valores, com base em informações domiciliares e comunitárias. Para a escolha do método de tratamento, os mecanismos de ligação entre os dados faltantes e os seus determinantes são importantes ou haverá viés

expressivo nas estimativas pretendidas. Rubin (1976)² criou um sistema de classificação com base nas hipóteses sobre fatores associados aos dados faltantes. São três: dados faltantes completamente aleatórios (*missing completely at random* – MCAR); dados faltantes aleatórios, também denominado ignorável (*missing at random* – MAR); e dados faltantes não aleatórios (*missing not at random* – NMAR) (LITTLE; RUBIN, 2002; ENDERS, 2010). Pode-se considerar o *missing* de escolaridade no Censo 2010 como MAR/ignorável. As razões da ausência de informações de Y para certas unidades de análises estão relacionadas a outras variáveis com respostas completas, mas não ao próprio Y .

As imputações da escolaridade (Y) foram realizadas usando o algoritmo *expectation-maximization* (algoritmo E-M), que é um método iterativo para estimação de dados incompletos, formalizado e discutido por Dempster *et al.* (1977) e descrito e detalhado por Schafer (1997), Little e Rubin (2002) e Enders (2010). Ele é apontado entre os mais eficazes para tratamento de dados faltantes, mesmo sob hipótese de NMAR (RUBIN, 1996; SCHAFFER; GRAHAM, 2002; ALISSON, 2009). Outros métodos de imputação foram testados no trabalho de Ribeiro (2016) e o E-M se mostrou mais adequado.

Cada interação procedida envolve uma etapa de estimação da expectativa (passo E) e outra de maximização (passo M). Tomando a função log-verossimilhança $l(\theta|Y)$ como o logaritmo natural da função de verossimilhança $L(\theta|Y)$, os passos de E-M consistem em: substituir os valores faltantes usando algum método mais simples (imputação pela média ou assumir uma distribuição uniforme, por exemplo); estimar parâmetros desejados por uma função de verossimilhança $l(\theta^{(0)}|Y_{\text{obs}})$; reestimar os valores faltantes usando os parâmetros estimados no item anterior; reestimar os parâmetros $l(\theta^{(1)}|Y_{\text{obs}})$ e assim por diante, $l(\theta^{(2)}|Y_{\text{obs}})$... $l(\theta^{(t)}|Y_{\text{obs}})$, até que o valor de θ alcance a convergência na t -ésima interação.³

O modelo para aplicação do método E-M foi estimado por regressão multivariada normal (MVN) considerando as variáveis explicativas do Quadro 1. A seleção das variáveis não segue critério de relação de causa e efeito com a variável imputada. Para fins de imputação, o importante é incorporar fatores que estejam altamente correlacionados com os dados faltantes. Como a amostra do Censo 2010 é suficientemente grande para modelos mais inclusivos, a estratégia foi incluir o máximo de variáveis com relação significativa com a escolaridade. Essa seleção é explicada com mais detalhes em Ribeiro (2016). O método MVN utiliza *data augmentation* (DA) com iterações de Monte Carlo via cadeias de Markov (MCMC). A estimação ocorre por máxima verossimilhança, porém, acrescenta-se um componente aleatório para criar uma amostra *bootstrap*. Primeiro foram estimados os valores esperados para os parâmetros que descrevem a relação entre as variáveis independentes e a escolaridade, partindo de uma distribuição uniforme, por meio do algoritmo E-M. A distribuição uniforme assume que todos os valores dos coeficientes são igualmente prováveis. Em seguida, partindo de amostras *bootstrap*, foram obtidos os valores esperados para os

² Rubin (1976) é responsável por desenvolver uma teoria sobre dados faltantes que se tornou referência na literatura sobre o tema.

³ Para mais detalhes sobre o procedimento de imputação, sugerimos consultar Ribeiro (2016).

dados faltantes. O método é iterativo e as imputações são realizadas por método de Monte Carlo via cadeias de Markov (MCMC). Quando os valores dos parâmetros convergem para uma distribuição estacionária, os dados faltantes são preenchidos. Não há uma regra de parada que garanta a distribuição estacionária e há possibilidade de dependência serial. Mas como a taxa de convergência depende da fração de dados faltantes, que neste caso é baixa, isso não foi um problema. No modelo de imputação por MNV, o ideal é trabalhar com anos de escolaridade, dada a suposição de normalidade da variável dependente. Após as imputações, a escolaridade foi agrupada nas três categorias já mencionadas: baixa, média e alta escolaridade.

QUADRO 1
Variáveis selecionadas para as regressões

Nível/classificação	Escolaridade/descrição	Medida/agrupamento
Variável dependente	Escolaridade	Anos de estudo
Indivíduo com óbito	Idade agrupada	25-29 anos 30-34 anos 35-39 anos 40-44 anos 45-49 anos 50-54 anos 55-59 anos
	Sexo	1 se homem, 0 se mulher
Características do responsável pelo domicílio	Sexo do responsável pelo domicílio	1 se homem, 0 se mulher
	Escolaridade do responsável pelo domicílio	anos de estudo
	Idade do responsável pelo domicílio	<30 anos 30-39 anos 40-49 anos 50-59 anos >=60
	Estado civil do responsável pelo domicílio	1 se tem cônjuge, 0 se não
	Responsável pelo domicílio tem pelo menos uma deficiência total ou grande?	1 se sim, 0 se não
	Raça/cor do responsável pelo domicílio	Branca Preta Amarela Parda Indígena
	Diferença entre grupo etário do responsável do domicílio e do indivíduo (adulto entre 25 e 59 anos)	Diferença > - 10 & < 10 anos Responsável mais jovem entre 10 e 19 anos Responsável 20 anos ou mais jovem que o indivíduo Responsável mais velho entre 10 e 19 anos Responsável 20 anos ou mais velho que o indivíduo

(continua)

(continuação)

Nível/classificação	Escolaridade/descrição	Medida/agrupamento
Domicílio	Indicador de condição socioeconômica do domicílio composto pelas variáveis: - existência de TV no domicílio - existência de carro no domicílio - existência de moto no domicílio - existência de máquina de lavar no domicílio - existência de geladeira no domicílio - existência de água canalizada em pelo menos um cômodo - existência de computador com internet no domicílio - existência de celular no domicílio - material permanente predominante (alvenaria) - se a moradia é adequada (ter abastecimento de água por rede geral, esgotamento sanitário por rede geral ou fossa séptica, coleta de lixo direta ou indireta e, no máximo, dois moradores por dormitório) - presença de pelo menos um morador do domicílio beneficiário de programa de transferência de renda (Bolsa Família, Peti ou outro)	4 categorias – quartis
	Razão dependência (≤ 14 ou ≥ 60)/(>15 e ≤ 59)	1 se >1 , 0 se ≤ 1
	Ocorrência de óbito infantil no domicílio	1 se sim, 0 se não
	Domicílios compostos exclusivamente por estrangeiros ou naturalizados ou tem brasileiro nato?	1 se tem algum brasileiro nato; 0 se só tem estrangeiros ou naturalizados
Comunidade	Proporção de pessoas vivendo em áreas subnormais	Nenhuma <50 ≥ 50 e < 90 ≥ 90
	Região metropolitana	1 se sim, 0 se não
	Situação do domicílio	1 se urbana, 0 se rural
	Escolaridade (educ) de indivíduo selecionado aleatoriamente por idade e sexo	

Óbitos do Sistema de Informações sobre Mortalidade (SIM)

Os dados do Sistema de Informações sobre Mortalidade (SIM) foram empregados no terceiro conjunto de estimativas, a fim de permitir a análise das consistências dos resultados estimados a partir dos dados censitários. Utilizaram-se os óbitos ocorridos entre 1 de agosto de 2012 e 31 de julho de 2013.

Os registros do SIM para São Paulo possuem alta confiabilidade, com cobertura de praticamente 100% dos óbitos (QUEIROZ; SAWYER, 2012). No entanto, as informações de escolaridade no SIM têm sido consideradas de baixa qualidade em todo o país, em razão, principalmente, do excesso de dados faltantes. Em 2009, houve uma mudança no quesito de escolaridade, que passou a conter ciclos de ensino (sem escolaridade, ensino

fundamental I, ensino fundamental II, nível médio e superior), seguidos da série concluída. A implementação da nova metodologia nas declarações de óbito (DO) começou a ser adotada a partir de 2011 (BRASIL, 2009, 2011). Como ainda circulam os dois modelos de declaração, os dados continuam sendo disponibilizados no formato antigo: 0, 1 a 3, 4 a 7, 8 a 11 e 12 ou mais anos de escolaridade. Assim, categorizamos a variável de escolaridade em três grupos, conforme mencionado na seção anterior: baixa escolaridade (0 a 3 anos de estudo); média (4 a 11 anos de estudo); e alta (12 anos ou mais de estudo). O percentual de perda ainda é elevado nas unidades da federação, mas houve redução na proporção de informação faltante para algumas capitais. No município de São Paulo, a escolaridade foi reportada para 92,3% dos óbitos de adultos de 25 a 59 anos, entre agosto de 2012 e julho de 2013. A baixa proporção de dados faltantes torna a base atrativa para investigação dos diferenciais de mortalidade por escolaridade e avaliação da consistência das estimativas censitárias. Para tratar os 7,7% de informações faltantes no SIM, realizou-se uma imputação simples por *hotdeck*, supondo que eles sejam MCAR. O procedimento de imputação levou em conta a idade e o sexo, o que indica que os óbitos foram redistribuídos aleatoriamente por estes atributos. O método do *hotdeck* e a explicação para aplicação nos dados do SIM estão em Ribeiro (2016). Após feitas as correções na base, agrupou-se a escolaridade em baixa, média e alta.

O modelo de Poisson para a estimação de taxas de mortalidade

O modelo de regressão de Poisson é um dos mais utilizados para o cálculo de taxas de mortalidade (COLEMAN, 1964) e permite estimar o número de mortes em função do risco de morte e do tempo de exposição. Assume-se que existe uma taxa de incidência na qual o evento ocorre, que pode ser multiplicada por um tempo de exposição para obter o número de eventos ocorridos, sendo as exposições não sobrepostas independentes entre si. O modelo é dado por:

$$\log Y = \beta X_i + \log[\text{exposição}] \quad (1)$$

Onde: a variável dependente é o logaritmo da taxa de mortalidade (taxa de incidência); β é um vetor de coeficientes estimados por máxima verossimilhança; e X é uma matriz de variáveis independentes. No presente trabalho, X é representado pela idade em grupos quinquenais, entre 25 e 59 anos, e pela escolaridade medida pelos grupos já mencionados (baixa, média e alta).

Não foi possível incluir variável de interação entre idade e escolaridade, pois o número de óbitos para São Paulo na amostra do censo é pequeno para o grupo de alta escolaridade e idades mais jovens, para o qual as taxas de mortalidade são relativamente baixas e mais sujeitas a erros (ver Tabela 2). Portanto, optou-se por modelos com efeitos puros de idade e escolaridade por sexo, o que permite avaliar os diferenciais de nível de mortalidade, mas restringe a investigação dos padrões etários de mortalidade por escolaridade.

Foram estimadas taxas de mortalidade a partir de três modelos Poisson, separadamente para homens e mulheres. Conforme indicado no início dessa seção, o primeiro modelo utilizou dados de óbitos do Censo 2010 assumindo que a escolaridade dos mortos é igual àquela declarada para o responsável pelo domicílio. O segundo modelo também fez uso de dados de óbitos do Censo 2010, mas com a escolaridade dos mortos imputada pelos métodos estatísticos descritos anteriormente. Finalmente, o terceiro modelo utilizou dados de óbitos por escolaridade extraídos do SIM/Datasus.

Quanto ao tempo de exposição ao evento, adotou-se a população do meio do período em todas as estimativas, assumindo que os óbitos estavam uniformemente distribuídos ao longo do ano-calendário. Embora fosse possível calcular o tempo de exposição aproximado de cada óbito a partir do seu mês de ocorrência, os resultados de nossos modelos seriam muito similares qualquer que fosse a estratégia adotada. Optamos pela forma mais parcimoniosa de cálculo e que também compensasse parte dos efeitos da migração ocorrida no município de São Paulo durante o período de análise.

Para as estimativas com SIM, trabalhamos com dados agregados para a população exposta, já que essa fonte fornece apenas microdados dos óbitos. Seguindo a mesma lógica utilizada para o Censo, foi considerada a população de 1º de julho de 2012, estimada pelo IBGE, e projetada para o meio do período da análise. Aplicou-se a mesma estrutura por sexo, idade e escolaridade observada para o Censo 2010, assumindo que três anos é um tempo pequeno o suficiente para que não tenha havido mudanças expressivas nessas distribuições.

Além disso, ao final, corrigiu-se o nível das taxas de mortalidade estimadas com os dados do Censo 2010, a partir do nível da taxa bruta de mortalidade por sexo estimada com base no SIM. O objetivo dessa correção foi reduzir eventuais problemas de enumeração dos óbitos do Censo 2010 (cobertura e erros de declaração), alterando seu nível, mas mantendo os padrões das taxas de mortalidade por idade e escolaridade da forma como foram captados originalmente pela base censitária.

Resultados

Na Tabela 1, apresenta-se uma análise inicial dos dados utilizados, sem considerar a escolaridade, por meio de taxas brutas de mortalidade (TBM) segundo sexo, calculadas a partir de óbitos do Censo 2010 e do SIM 2013. O nível geral de mortalidade é maior no SIM 2013 do que no Censo 2010, respectivamente 14% e 18% para homens e mulheres. Os intervalos de confiança estimados para os dados do Censo 2010 sugerem que as diferenças entre as estimativas das duas bases de dados são estatisticamente diferentes, reforçando a necessidade de correção do nível de mortalidade medido pelo censo.

TABELA 1
Taxas brutas de mortalidade (TBM) da população de 25 a 59 anos, segundo sexo
Município de São Paulo – 2010-2013

Sexo	Censo – amostra		Censo – amostra expandida	SIM	Razão SIM/ Censo
	TBM (1)	IC (95%)	TBM (1)	TBM (1)	
Mulheres	1,69	[1,50 ; 1,92]	1,70	2,01	1,18
Homens	3,79	[3,47 ; 4,14]	3,82	4,35	1,14
Total	2,68	[2,50 ; 2,88]	2,70	3,11	1,15

Fonte: IBGE. Censo Demográfico 2010; Sistema de Informações sobre Mortalidade (SIM) 2013.

(1) Por mil habitantes de 25 a 59 anos.

A Tabela 2 apresenta a distribuição dos óbitos segundo diferentes fontes e métodos utilizados. A distribuição dos óbitos por escolaridade imputada estatisticamente no Censo 2010 mostrou ser semelhante àquela observada no SIM. Quanto maior a idade, maiores são as amostras para óbitos no Censo 2010 e mais as duas distribuições se aproximam, o que está relacionado à variação da consistência das imputações por idade. Por outro lado, no geral, os óbitos do Censo 2010 segundo a escolaridade do responsável pelo domicílio estão mais concentrados no grupo de média escolaridade, diferindo das distribuições imputadas no Censo 2010 e registradas no SIM. No caso dos óbitos de indivíduos mais jovens, a distribuição de escolaridade medida pela escolaridade do responsável pelo domicílio favorece o nível baixo, comparativamente às outras duas metodologias. Nas idades mais velhas, o padrão se inverte, dando mais peso relativo ao nível médio. Esses padrões são explicados por mudanças tanto na composição domiciliar ao longo do ciclo de vida, quanto na distribuição de escolaridade por gerações no Brasil, conforme será abordado na seção Discussão, mais adiante.

TABELA 2
Distribuição dos óbitos, por sexo e categoria de escolaridade, segundo métodos de estimação e grupos de idade
Município de São Paulo – 2010-2013

Métodos de estimação e grupos de idade	Homens						Mulheres						
	Categoria de escolaridade (%)				N	Amostra expandida	Categoria de escolaridade (%)				N	Amostra expandida	
	Baixa	Média	Alta	Total			Baixa	Média	Alta	Total			
Escolaridade do responsável pelo domicílio – Censo													
25-34 anos	57,1	33,3	9,5	100,0	42	897	41,2	52,9	5,9	100,0	17	335	
30-34 anos	42,1	55,3	2,6	100,0	38	754	33,3	38,9	27,8	100,0	18	358	
35-39 anos	55,1	32,7	12,2	100,0	49	1.030	35,7	64,3	0,0	100,0	14	293	
40-44 anos	54,3	38,6	7,1	100,0	70	1.466	31,3	50,0	18,8	100,0	32	642	
45-49 anos	52,1	38,0	9,9	100,0	71	1.514	45,7	41,3	13,0	100,0	46	932	
50-54 anos	50,5	40,2	9,3	100,0	107	2.205	43,3	45,0	11,7	100,0	60	1.269	
55-59 anos	43,3	39,2	17,5	100,0	120	2.536	41,3	38,1	20,6	100,0	63	1.286	
Total	49,9	39,2	10,9	100,0	497	10.403	40,4	44,4	15,2	100,0	250	5.115	

(continua)

(continuação)

Métodos de estimação e grupos de idade	Homens						Mulheres					
	Categoria de escolaridade (%)				N	Amostra expandida	Categoria de escolaridade (%)				N	Amostra expandida
	Baixa	Média	Alta	Total			Baixa	Média	Alta	Total		
Escolaridade imputada – Censo												
25-34 anos	33,3	33,3	33,3	100,0	42	897	17,6	41,2	41,2	100,0	17	335
30-34 anos	44,7	34,2	21,1	100,0	38	754	27,8	44,4	27,8	100,0	18	358
35-39 anos	51,0	28,6	20,4	100,0	49	1.030	64,3	7,1	28,6	100,0	14	293
40-44 anos	58,6	30,0	11,4	100,0	70	1.466	37,5	31,3	31,3	100,0	32	642
45-49 anos	56,3	35,2	8,5	100,0	71	1.514	45,7	39,1	15,2	100,0	46	932
50-54 anos	63,6	22,4	14,0	100,0	107	2.205	56,7	31,7	11,7	100,0	60	1.269
55-59 anos	72,5	21,7	5,8	100,0	120	2.536	69,8	20,6	9,5	100,0	63	1.286
Total	58,8	27,6	13,7	100,0	497	10.403	51,2	30,4	18,4	100,0	250	5.115
Escolaridade e óbitos do SIM												
25-34 anos	43,7	45,8	10,5	100,0	911	-	32,8	48,1	19,1	100,0	293	-
30-34 anos	49,4	38,5	12,1	100,0	1.014	-	36,6	42,2	21,2	100,0	415	-
35-39 anos	61,2	28,6	10,2	100,0	1.127	-	43,0	38,1	18,9	100,0	535	-
40-44 anos	63,5	28,8	7,7	100,0	1.490	-	54,1	33,1	12,8	100,0	712	-
45-49 anos	61,3	28,2	10,5	100,0	1.892	-	53,9	30,6	15,5	100,0	1.066	-
50-54 anos	61,7	26,9	11,4	100,0	2.576	-	58,0	27,7	14,3	100,0	1.486	-
55-59 anos	61,8	27,1	11,1	100,0	3.341	-	63,9	22,8	13,3	100,0	1.930	-
Total	59,5	29,9	10,6	100,0	12.351	-	54,9	30,0	15,1	100,0	6.437	-

Fonte: IBGE. Censo Demográfico 2010; Sistema de Informações sobre Mortalidade (SIM) 2013.

Os resultados dos modelos Poisson, que controlam, simultaneamente, por grupos de idade e escolaridade, encontram-se na Tabela 3. Como já destacado, são apresentadas estimativas para homens e mulheres, separadamente, além de modelos individuais para cada conjunto de dados.

Nos três modelos, os níveis de mortalidade crescem com a idade, o que está dentro do esperado. No entanto, as razões entre as taxas sugerem níveis de mortalidade relativamente mais altos no grupo etário 55 a 59 anos no SIM 2013 do que no Censo 2010, tanto para homens quanto para mulheres. Embora as estimativas do SIM 2013 estejam dentro dos intervalos de confiança calculados com os dados censitários, deve-se destacar que os estimadores das IRRs por idade têm precisão menor (erro padrão e intervalos de confiança relativamente grandes) em razão do baixo número de observações, sobretudo, para as mulheres (Tabela 2). Isso é uma limitação para a estimação dos níveis reais de mortalidade no Censo 2010, mas não para identificação dos diferenciais entre os grupos, supondo que os estimadores não sejam enviesados. O diferencial observado entre SIM 2013 e Censo 2010 pode ser uma consequência de problemas de cobertura ou subnumeração de óbitos do Censo 2010. Conforme destacado anteriormente, a subnumeração tenderia a ser mais frequente entre os indivíduos que residem sozinhos, mais prevalentes em idades mais velhas, mas pode ter ocorrido também em outras idades, se tiver havido omissão de domicílios durante a coleta.

TABELA 3
Razão entre as taxas de mortalidade (IRR – Incidence Rate Ratio), por fontes ou métodos, segundo sexo, grupos de idade e categoria de escolaridade
Município de São Paulo – 2010-2013

Variáveis	Modelo 1 – Óbitos do Censo 2010, escolaridade do responsável pelo domicílio				Modelo 2 – Óbitos do Censo 2010, escolaridade imputada			Modelo 3 – Óbitos do SIM 2013	
	IRR	Erro padrão	IC-95%		IRR	Erro-padrão	IC-95%		
Homens									
<i>Idade</i>									
25-29	-	-	-	-	-	-	-	-	
30-34	1,09	0,26	0,68	1,74	1,06*	0,25	0,66	1,69	1,12*+
35-39	1,41	0,32	0,91	2,19	1,35*	0,30	0,87	2,09	1,31*+
40-44	2,25	0,47	1,50	3,38	2,12*	0,43	1,41	3,16	1,82*+
45-49	2,46	0,51	1,64	3,69	2,29*	0,47	1,53	3,41	2,62*+
50-54	4,60	0,92	3,11	6,82	4,22*	0,85	2,85	6,26	3,92*+
55-59	5,93	1,16	4,04	8,70	5,34*	1,02	3,68	7,75	6,12*+
<i>Escolaridade</i>									
Baixa	-	-	-	-	-	-	-	-	-
Média	0,73	0,08	0,59	0,90	0,43	0,05	0,34	0,55	0,45+
Alta	0,31	0,05	0,23	0,42	0,32*	0,05	0,25	0,43	0,23
<i>Constante</i>	0,0000	0,00	0,0000	0,0000	0,0000	0,00	0,0000	0,0000	0,0018+
Mulheres									
<i>Idade</i>									
25-29	-	-	-	-	-	-	-	-	-
30-34	0,99	0,36	0,49	2,01	0,96*	0,35	0,47	1,95	1,40*+
35-39	0,78	0,30	0,36	1,68	0,74*	0,29	0,35	1,59	1,93*
40-44	2,25	0,75	1,17	4,33	2,10*	0,69	1,10	4,01	2,67*+
45-49	3,08	0,97	1,67	5,70	2,84*	0,88	1,54	5,22	4,19*+
50-54	5,49	1,98	2,70	11,14	4,98*	1,79	2,46	10,06	6,19*+
55-59	5,81	1,86	3,10	10,88	5,17*	1,61	2,81	9,50	9,21*+
<i>Escolaridade</i>									
Baixa	-	-	-	-	-	-	-	-	-
Média	0,83	0,17	0,88	1,52	0,46	0,09	0,32	0,67	0,50+
Alta	0,40	0,10	0,39	0,82	0,43*	0,09	0,28	0,65	0,37+
<i>Constante</i>	0,0000	0,00	0,0000	0,0000	0,0000	0,00	0,0000	0,0000	0,0005*+

Fonte: IBGE. Censo Demográfico 2010; Sistema de Informações sobre Mortalidade (SIM) 2013.

* Indica que as IRRs do modelo estão dentro do IC do modelo 1.

+ Indica que as IRRs do modelo estão dentro do IC do modelo 2.

Nota: IRR em negrito: p -valor < 0,05. Óbitos sem correção nos três modelos.

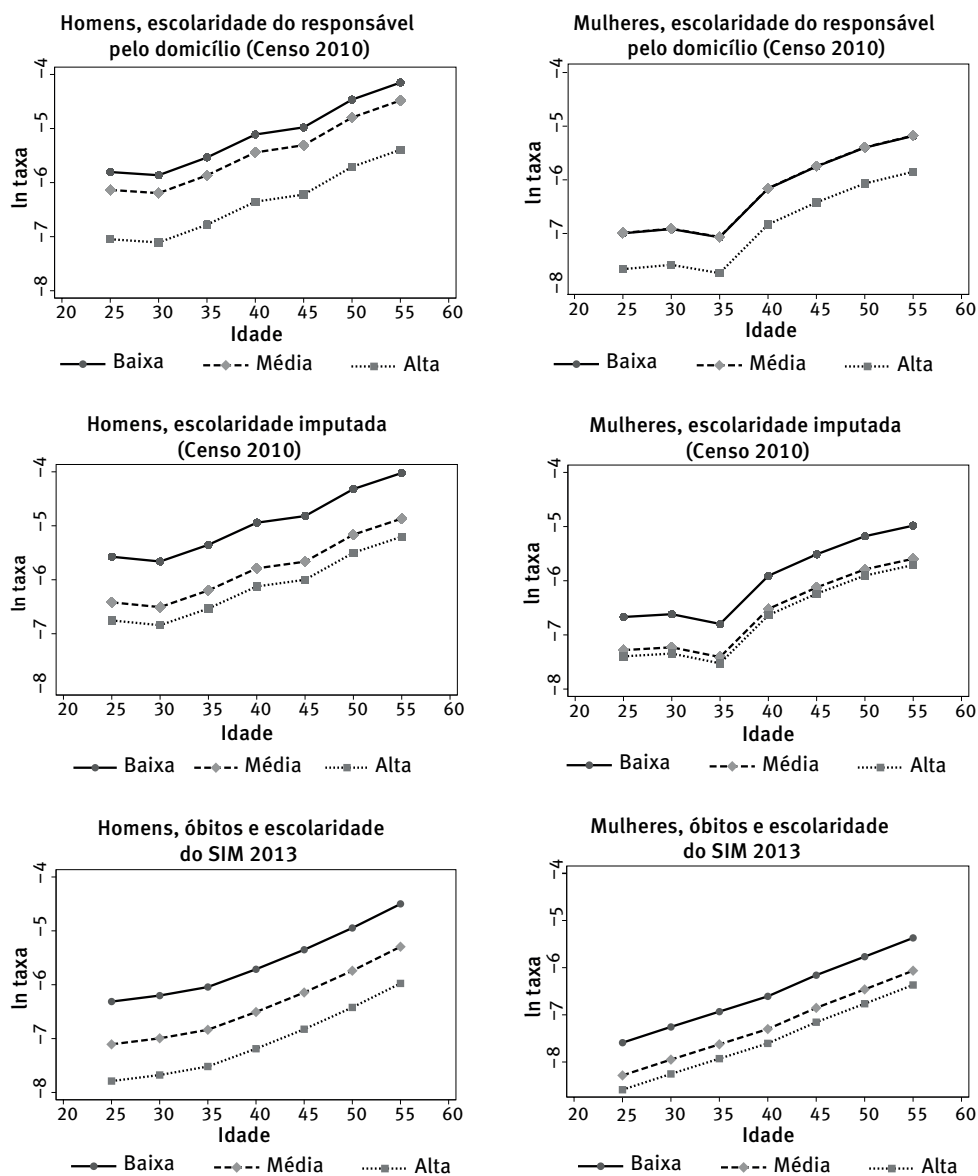
As diferenças de mortalidade por categorias de educação são evidentes nos três modelos, indicando taxas de mortalidade mais baixas para os grupos mais escolarizados. No entanto, as razões entre as taxas, em cada fonte de dados, indicam padrões de mortalidade por escolaridade muito distintos. No modelo 1, que considera óbitos do Censo 2010 e a escolaridade do chefe do domicílio para os mortos, o ganho está concentrado na alta escolaridade, com razões de mortalidade em relação ao grupo de baixa escolaridade de, aproximadamente, 0,31 e 0,40, para homens e mulheres, respectivamente.

Entre os grupos de baixa e média escolaridade, o diferencial é relativamente pequeno para os homens e estatisticamente não significativo para as mulheres.

No modelo 2, baseado também em dados do Censo 2010, mas com escolaridade imputada por métodos estatísticos, os maiores ganhos relativos se concentram no grupo de média escolaridade, sendo as razões de mortalidade em relação ao grupo de baixa escolaridade iguais a cerca de 0,43 e 0,46, para homens e mulheres, respectivamente. O gradiente de mortalidade por escolaridade observado no modelo 3, estimado com base em dados do SIM 2013, é semelhante ao do modelo 2. No entanto, o efeito marginal da alta escolaridade é ligeiramente mais pronunciado no modelo 3 do que no modelo 2, particularmente no caso das mulheres.

A partir dos coeficientes dos modelos Poisson, foram estimadas as taxas de mortalidade para homens e mulheres, por idade, grupos de escolaridade e metodologia adotada. O nível das taxas estimadas com base nos modelos 1 e 2 (dados do Censo 2010) foi corrigido utilizando-se a razão entre as TBM por sexo, do SIM 2013 e do Censo 2010, apresentadas na Tabela 1. O Gráfico 1 evidencia quatro resultados principais. Além do esperado aumento da taxa de mortalidade com a idade e dos maiores níveis de mortalidade para os homens do que para as mulheres, a utilização da escolaridade dos responsáveis pelos domicílios como *proxy* para a escolaridade dos mortos (modelo 1) resulta em um gradiente atípico, que não foi observado nas metodologias alternativas. Embora os gradientes calculados a partir dos modelos 2 e 3 não sejam idênticos, ambos sugerem um papel relativamente mais importante do ensino médio, como fator de proteção em relação à morte. Finalmente, vale notar que a escolaridade desempenha um papel tão significativo para sobrevivência de adultos no Brasil, que mulheres de baixa escolaridade apresentam taxas de mortalidade similares às dos homens de educação média e alta, nas mesmas idades. Esse resultado ocorre a despeito da conhecida vantagem feminina na mortalidade adulta no Brasil, sugerindo que a baixa escolaridade elimina eventuais proteções biológicas e comportamentais que estão mais presentes entre as mulheres.

GRÁFICO 1
Taxas de mortalidade, por sexo e fontes de dados dos óbitos, segundo idade e categorias de escolaridade
Município de São Paulo – 2010-2013



Fonte: IBGE. Censo Demográfico 2010; Sistema de Informações sobre Mortalidade (SIM) 2013. Estimativas a partir dos coeficiente dos modelos 1, 2 e 3 apresentados na Tabela 1 com correção do nível das taxas pelo SIM.

Discussão

Este trabalho apresentou novos resultados para os diferenciais educacionais na mortalidade adulta no país, contribuindo para uma área da demografia brasileira que ainda necessita avançar. Em virtude da inexistência de uma base de dados com qualidade reconhecidamente superior, foram comparados três conjuntos distintos de informações. O objetivo dessa comparação foi oferecer uma análise crítica sobre o verdadeiro gradiente de mortalidade adulta por educação no Brasil. De maneira geral, os resultados indicam que a aquisição de escolaridade pode reduzir em até 77% as taxas de mortalidade entre 25 e 59 anos de idade. Além disso, em um país em que a população tem baixa escolaridade, obter ensino médio ainda representa um ganho significativo do ponto de vista da sobrevivência. Ressalta-se, entretanto, que a forma de coleta de dados e a necessidade de imputação de informações faltantes afetam a distribuição de ganhos por níveis de escolaridade, o que faz com que as questões metodológicas permaneçam centrais na análise de diferenciais de mortalidade no Brasil, gerando incertezas sobre os verdadeiros padrões e exigindo atenção dos pesquisadores.

A despeito de todas as limitações existentes nos dados censitários, o trabalho também demonstrou que as informações do Censo 2010 podem ser úteis para os estudos de diferenciais de mortalidade, desde que os devidos cuidados sejam tomados com relação às formas de imputação das variáveis faltantes. Em trabalho anterior, Silva *et al.* (2016) mediram as probabilidades de morte e expectativa de vida para a população entre 15 e 60 anos, segundo a escolaridade do responsável pelo domicílio. A análise aqui realizada sugere que a escolaridade do responsável pelo domicílio não é uma boa *proxy* para a escolaridade dos mortos, resultando em um diferencial de mortalidade por educação distinto do encontrado com dados do SIM e com dados censitários imputados por um modelo estatístico mais robusto, como o proposto aqui.

Uma hipótese para as diferenças encontradas com base na escolaridade do responsável do domicílio é a transição de educação que vem acontecendo no Brasil. Houve um aumento do acesso à educação desde a década de 1990, com a instituição da cobertura universal do ensino primário. Entre 1986 e 2008, a escolaridade média da população de 7 a 25 anos de idade passou de 4 para 6 anos de estudo (RIOS-NETO; GUIMARÃES, 2010). Portanto, embora a escolaridade dos indivíduos mantenha uma relação direta com a escolaridade dos responsáveis pelos domicílios em que residem, variações geracionais no nível de escolaridade são inevitáveis no Brasil. Além disso, para cada fase do ciclo de vida, medida pela idade dos residentes mortos, há uma composição domiciliar típica. Esse padrão por idades, combinado com as variações na escolaridade entre gerações, provoca as distorções demonstradas neste trabalho. Por exemplo, constatamos (em resultados não apresentados) que os responsáveis são, em média, mais velhos do que os mortos, nos domicílios em que ocorreram mortes de indivíduos jovens. Portanto, nesses casos, os responsáveis tendem a ter escolaridade mais baixa do que a provável escolaridade dos mortos. O padrão de

idade e escolaridade se inverte nos domicílios em que ocorreram óbitos em idades mais velhas, favorecendo uma superestimação do nível médio de escolaridade. Como resultado desses efeitos combinados para todas as idades, constatamos que uma parte dos óbitos que deveria ser classificada como pertencente a indivíduos de baixa escolaridade foi erroneamente identificada como sendo de média escolaridade. Consequentemente, as taxas de mortalidade dos dois níveis de escolaridade tornaram-se mais parecidas, suavizando o diferencial de mortalidade. Silva *et al.* (2016) reconhecem que a existência de algum grau de mobilidade intergeracional de educação poderia subestimar os diferenciais de mortalidade por nível educacional.

Já a escolaridade imputada por métodos estatísticos oferece uma estimativa que tem como base a escolaridade do indivíduo, o que torna os diferenciais de mortalidade medidos por meio dessa estratégia mais parecidos com os obtidos com dados do SIM. Associamos o êxito da imputação estatística proposta por nós ao tamanho da amostra e à disponibilidade de um número considerável de características individuais (sobreviventes) e domiciliares no Censo 2010. Ainda assim, para que se possa aplicar os métodos desenvolvidos nesse trabalho para todo o Brasil, é necessário testá-los para as diferentes regiões e analisar como se comportam diante de diferentes populações. Além disso, há que se considerar a existência de problemas na qualidade da informação de escolaridade dos entrevistados nos censos, padrão que foi recentemente analisado para o Brasil (NEPOMUCENO; TURRA, 2020). De todo modo, a similaridade dos padrões de mortalidade por escolaridade estimados com dados do Censo 2010 (imputados) e SIM 2013 sugere que a solução metodológica apresentada nesse trabalho pode ser bastante útil para futuras análises a partir de dados de óbitos dos censos brasileiros.

Apesar do êxito da imputação estatística, para estudos de mortalidade, os dados do SIM ainda são preferíveis aos do Censo 2010. Mas é necessário lembrar que as taxas de mortalidade obtidas com óbitos do SIM também não estão livres de vies. Ainda não sabemos, efetivamente, qual é a qualidade da informação de escolaridade nas declarações de óbito que pode ser reportada por alguém que não conhece bem o indivíduo falecido. Além disso, como persistem dois modelos de declaração de escolaridade no SIM, o processo de compatibilização para sistematização dos dados pode ser fonte de erro. Há ainda problemas de consistência entre o numerador e o denominador das taxas de mortalidade, já que para a utilização dos dados de óbitos do SIM, segundo a metodologia convencional, é necessário utilizar outra fonte de dados para o denominador (população exposta). Não é certo que a escolaridade esteja declarada de forma consistente nos óbitos do SIM (numerador) e na população censitária (denominador).

A despeito de eventuais erros nas fontes de dados, os diferenciais de mortalidade estimados neste trabalho, quando calculados segundo a escolaridade dos próprios indivíduos, são comparáveis aos apresentados em outros estudos. Sandoval e Turra (2015) estimaram a mortalidade por escolaridade no Chile para o período 2001-2003, utilizando os mesmos intervalos de idade adotados por nós. Embora os grupos de escolaridade

empregados pelos autores sejam ligeiramente distintos dos nossos (≥ 8 , ≥ 9 e ≤ 12 , ≥ 13 anos de escolaridade), os diferenciais estimados para um país vizinho podem ajudar a detectar padrões irrealistas no Brasil. Como exemplo, os chilenos de 55 a 59 anos apresentam taxas de 13,1 (baixa escolaridade), 8,5 (média) e 3,7 por mil (alta escolaridade), indicando diferenciais semelhantes aos encontrados no nosso trabalho para homens residentes em São Paulo, com base em dados do SIM 2013. Para as mulheres do mesmo grupo de idade, os diferenciais entre os níveis de escolaridade médio e baixo são similares, mas os ganhos de sobrevivência com o ensino superior são superiores no Chile comparativamente a São Paulo. É esperado que os efeitos marginais da escolaridade sobre a mortalidade variem em cada país, em função do seu contexto econômico, social e cultural. Confirmando nossos padrões por sexo, Gomes *et al.* (2013) também encontram uma associação negativa entre nível de escolaridade e mortalidade maior para homens do que para mulheres, utilizando dados de painel da Sabe para idosos residentes em São Paulo.

Em função das limitações dos dados de óbitos do Censo 2010, não estimamos os diferenciais de mortalidade por educação para pessoas com mais de 60 anos. No entanto, há registros na literatura de que os diferenciais diminuem com o avançar da idade em outros países (Chile e EUA, por exemplo). Algumas explicações para esse padrão incluem os efeitos de seletividade da mortalidade por escolaridade, variações na relação entre escolaridade e mortalidade por coorte de nascimento e existência de programas públicos de saúde específicos para idosos em alguns países (LAUDERDALE, 2001; PRESTON; ELO, 1995). No caso do Brasil, o aumento do acesso à escolaridade para as coortes mais jovens (RIOS-NETO; GUIMARÃES, 2010) pode ter implicações importantes para os padrões etários dos diferenciais de mortalidade por educação e, portanto, merecem ser analisados em pesquisa futura. Além disso, com a expansão do sistema de ensino no país, as relações da escolaridade com o nível de renda e o *status* ocupacional podem mudar entre as gerações, com repercussões para a saúde e mortalidade em cada período. Há que se considerar, também, que no Brasil o SUS pode desempenhar um papel relevante na intermediação entre saúde e escolaridade. A presença de um sistema público universal e integral pode diminuir os diferenciais de mortalidade nas idades em que a prevalência de morbidades é maior e a saída dos indivíduos do mercado de trabalho diminui seu acesso à saúde suplementar. Nesse sentido, a estrutura de oferta de serviços de saúde em cada localidade pode afetar os padrões regionais dos diferenciais de mortalidade por escolaridade, outro aspecto não examinado no presente trabalho.

Finalmente, uma agenda futura de pesquisa que pretenda ajudar a elucidar o comportamento futuro dos ganhos de longevidade no Brasil deve ir além da mensuração de diferenciais educacionais na mortalidade total. É preciso identificar as causas de morte que são responsáveis pelos gradientes observados e considerar outras medidas socioeconômicas além da escolaridade, como ocupação, renda e condições socioeconômicas anteriores à fase adulta, experimentadas na infância.

Referências

- ALLISON, P. D. Missing data. *In*: MILLSAP, R. E.; MAYDEU-OLIVARES, A. **The SAGE handbook of quantitative methods in psychology**. SAGE Publications, 2009. p 72-89.
- ANTONOVSKY, A. Social class life expectancy and overall mortality. **The Milbank Quarterly**, v. 45, n. 2, p. 31-73, 1967.
- BERHANE, Y.; HOGBERG, U.; BYASS, P.; WALL, S. Gender, literacy, and survival among ethiopian adults, 1987-96. **Bulletin of the World Health Organization**, New York, v. 80, n. 9, p. 714-720, Sept. 2002.
- BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. **Portaria nº 116, de 11 de fevereiro de 2009**. Regulamenta a coleta de dados, fluxo e periodicidade de envio das informações sobre óbitos e nascidos vivos para os Sistemas de Informações em Saúde sob gestão da Secretaria de Vigilância em Saúde. Brasília, 2009.
- BRASIL. Ministério da Saúde. Coordenação Geral de Informações e Análise Epidemiológica – CGIAE. **Sistema de Informações sobre Mortalidade – SIM**: consolidação da base de dados de 2011. Disponível em: http://tabnet.datasus.gov.br/cgi/sim/Consolida_Sim_2011.pdf.
- CASE, A.; DEATON, A. Rising morbidity and mortality in midlife among white non-Hispanic Americans in the 21st century. **Proceedings of the National Academy of Sciences**, v. 112, n. 49, p. 15078-15083, Dec. 2015.
- CRIMMINS, E. Demography: the 30 past years, the present, and the future. **Demography**, v. 30, n. 4, p. 579-591, 1993.
- CUTLER, D. M.; LLERAS-MUNEY, A. Education and health: evaluating theories and evidence. *In*: HOUSE, J. S.; SCHOENI, R. F.; KAPLAN, G. A.; POLLACK, H. (ed.). **The health effects of social and economic policy**. New York: Russel Sage Foundation, 2007.
- DEMPSTER, A. P.; LAIRD, N. M.; RUBIN, D. B. Maximum likelihood from incomplete data via the EM algorithm. **Journal of the Royal Statistical Society. Series B (Methodological)**, v. 39, n. 1, p. 1-38, 1977.
- DENISOVA, I. **Mortality in Russia**: microanalysis. Center for Economic and Financial Research and New Economic School, March, 2009. (Working Paper, 128).
- ELO, I. T.; PRESTON, S. H. Educational differentials in mortality: United States, 1979-1985. **Social Science and Medicine**. v. 42, p. 47-57, 1996.
- ELO, I. T. Social class differentials in health and mortality: patterns and explanations in comparative perspective. **Annual Reviews Of Sociology**, v. 35, p. 553-572, 2009.
- ENDERS, C. K. **Applied missing data analysis**. New York: The Guilford Press, 2010.
- GOMES, M. M. F.; TURRA, C.; FÍGOLI, M. G. B.; DUARTE, Y. A. O.; LEBRÃO, M. L. Associação entre mortalidade e estado marital: uma análise para idosos residentes no Município de São Paulo, Brasil, Estudo SABE, 2000 e 2006. **Cadernos de Saúde Pública**, Rio de Janeiro, v. 29, n. 3, p. 566-578, mar. 2013.
- GOMES, M. M. F.; TURRA, C.; FÍGOLI, M. G. B.; DUARTE, Y. A. O.; LEBRÃO, M. L. Passado e presente: condições de vida na infância e mortalidade de idosos. **Revista de Saúde Pública**, v. 49, 2015.
- GUEDES, G. R.; SIVIERO, P. C. L.; QUEIROZ, B. L.; MACHADO, C. J. Approximating the educational differences in mortality: demographic indirect techniques. **Cadernos de Saúde Coletiva**, Rio de Janeiro, v. 19, n. 2, p. 240-3, 2011.

GROSSMAN, M. On the concept of health capital and the demand for health. **The Journal of Political Economy**, Chicago, v. 80, n. 2, p. 223-55, 1972.

GROSSMAN, M. The relationship between health and schooling. **Eastern Economic Journal**, v. 34, n. 3, p. 281-292, 2008.

HAYWARD, M. D.; HUMMER, R. A.; SASSON, I. Trends and group differences in the association between educational attainment and U.S. Adult mortality: implications for understanding education's causal influence. **Social Science & Medicine**, v. 127, p. 8-18, 2015.

HEITJAN, D. F.; RUBIN, D. B. Inference from coarse data via multiple imputation with application to age heaping. **Journal of the American Statistical Association**, v. 85, n. 410, p. 304-314, 1990.

HEITJAN, D. F.; RUBIN, D. B. Ignorability and coarse data. **Annals of Statistics**, v. 19, n. 4, p. 2244-2253, 1991.

HOUSE, J. S.; LANDIS, K. R.; UMBERSON, D. Social relationship and health. **Science**, v. 241, p. 540-545, 1988.

HOUSE, J. S.; KESSLER, R. C.; HERZOG, R. Age, socioeconomic status and health. **The Milbank Quarterly**, v. 68, n. 3, p. 383-411, 1990.

HURT, L. S.; RONSMANS, C.; SAHA, S. Effects of education and other socioeconomic factors on middle age mortality in rural Bangladesh. **Journal of Epidemiology & Community Health**, v. 58, n. 4, p. 315-20, 2004.

KITAGAWA, E. M.; HAUSER, P. M. **Differential mortality in the United States: a study in socioeconomic epidemiology**. Cambridge, MA: Harvard University Press, 1973.

KOCH, E. *et al.* Desigualdad educacional y socioeconómica como determinante de mortalidad en Chile: análisis de sobrevivencia en la cohorte del proyecto San Francisco. **Revista Médica de Chile**, v. 135, n. 11, p. 1370-1379, 2007.

LAUDERDALE, D. S. Education and survival: birth cohort, period, and age effects. **Demography**, v. 38, n. 4, p. 551-561, 2001.

LEINSALU, M.; VAGERO, D.; KUNST, A. E. Estonia 1989-2000: enormous increase in mortality differences by education. **International Journal of Epidemiology**, v. 32, n. 6, p. 1081-87, 2003.

LIANG, J. *et al.* Socioeconomic gradient in old age mortality in Wuhan, China. **Journal of Gerontology: Social Sciences**, v. 55B, n. 4, p. S222-S233, 2000.

LIBERATOS P.; LINK, B. G.; KELSEY, J. L. The measurement of social class in epidemiology. **Epidemiologic Review**, v. 10, n. 1, p. 87-121, 1988.

LILLARD, L. A.; WAITE, L. J. 'Til death do us part': marital disruption and mortality. **American Journal of Sociology**, v. 100, p. 1131-56, 1995.

LITTLE, R. J. A.; RUBIN, D. B. **Statistical analysis with missing data**. 2. ed. Hoboken, NJ: Wiley, 2002.

LITTLE, T. D.; RHEMTULLA, M. Planned missing data designs for developmental researchers. **Child Development Perspectives**, v. 7, n. 4, p. 199-204, 2013.

MACKENBACH, J. P.; STIRBU, I.; ROSKAM, A.J. R.; SCHAAP, M. M.; MENVIELLE, G.; LEINSALU, M.; KUNST, A. E. Socioeconomic inequalities in health in 22 European countries. **New England Journal of Medicine**, v. 358, p. 2468-81, 2008.

MANZELLI, H. Educational attainment and adult mortality differentials in Argentina. **Revista Latinoamericana de Población**, v. 8, n. 14, p. 129-163, 2014.

MEARA, E.; RICHARDS, S.; CUTLER, D. The gap gets bigger: changes in mortality and life expectancy by education, 1981-2000. **Health Aff (Millwood)**, v. 27, n. 2, p. 350-360, 2008.

MISTLER, S. A.; ENDERS, C. K. Planned missing data designs for developmental research. *In*: LAURSEN, B.; LITTLE, T. D.; CARD, N. A. (ed.). **Handbook of developmental research methods**. New York, NY: Guilford Press, 2012. p. 742-754.

MONTEZ, J. K.; HUMMER, R. A.; HAYWARD, M. D.; WOO, H.; ROGERS, R. G. Trends in the educational gradient of U.S. adult mortality from 1986 through 2006 by race, gender, and age group. **Research on Aging**, v. 33, n. 2, p. 145-71, 2011.

MONTEZ, J. K.; HUMMER, R. A.; HAYWARD, M. D. Educational attainment and adult mortality in United States: a systematic analysis of functional form. **Demography**, v. 49, p. 315-336, 2012.

MORAIS, C. C. B. **Diferenciais de mortalidade adulta por nível de renda no estado de São Paulo: variações por sexo e fase do ciclo de vida, a partir de três estratégias de imputação**. Dissertação (Mestrado em Demografia) – Centro de Desenvolvimento e Planejamento Regional, Universidade Federal de Minas Gerais, Belo Horizonte, 2019.

NCHS – National Center for Health Statistics. **Highlights from Wave 1 of the National Survey of Personal Health Practices and Consequences: United States, 1979**. Vital and Health Statistics, Series 15(1), 1981.

NEPOMUCENO, M. R.; TURRA C. M. Assessing the quality of education reporting in Brazilian censuses. **Demographic Research**, v. 42, p. 441-460, 2020.

OLSHANSKY, S. J. *et al.* A potential decline in life expectancy in the United States in the 21st century. **New England Journal of Medicine**, v. 352, n. 11, p. 1138-1145, 2005.

PAMPEL, F. C.; KRUEGER, P. M.; DENNEY, J. T. Socioeconomic disparities in health behaviors. **Annual Review of Sociology**, v. 36, p. 349-70, 2010.

PAPPAS, G.; QUEEN, S.; HADDEN, W.; FISHER, G. The increasing disparity on mortality between socioeconomic groups in the United States, 1960 and 1986. **New England Journal of Medicine**, v. 329, n. 2, p. 103-109, 1993.

PEREIRA, F. N. A.; QUEIROZ, B. L. Diferenciais de mortalidade jovem no Brasil: a importância dos fatores socioeconômicos dos domicílios e das condições de vida nas localidades. **Cadernos de Saúde Pública**, v. 32, n. 9, 2016

PEREZ, E. R.; TURRA, C. M. Desigualdade social na mortalidade no Brasil: diferenciais por escolaridade entre mulheres adultas. *In*: XVI ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS. **Anais [...]**. Caxambu: Abep, 2008.

PEREZ, E. R. **Estimativas de mortalidade adulta feminina por nível de escolaridade no Brasil**. 2010, 146f. Tese (Doutorado em Demografia) – Centro de Desenvolvimento e Planejamento Regional, Universidade Federal de Minas Gerais, Belo Horizonte, 2010.

PRESTON, S. H.; TAUBMAN, P. Socioeconomic differences in adult mortality and health status. *In*: MARTIN, L.; PRESTON, S. H. **The demography of aging**. Washington, DC: National Academy Press, 1994. p. 279-318.

PRESTON, S. H.; EIO, I. T. Are educational differentials in adult mortality increasing in the United States? **Journal of Aging Health**, v. 7, p. 476-496, 1995.

QUEIROZ, B. L.; SAWYER, D. O. T. O que os dados de mortalidade do Censo de 2010 podem nos dizer? **Revista Brasileira de Estudos de População**, v. 29, n. 2, p. 225-238, 2012.

RIBEIRO, M. M. **Mortalidade adulta por níveis de escolaridade no estado e no município de São Paulo: uma proposta de estimação a partir do Censo Demográfico de 2010**. Tese (Doutorado em Demografia) – Centro de Desenvolvimento e Planejamento Regional, Universidade Federal de Minas Gerais, Belo Horizonte, 2016.

RIOS-NETO, E.; GUIMARÃES, R. R. M.; PIMENTA, P. S.; MORAES, T. A. **Análise da evolução de indicadores educacionais no Brasil: 1981 a 2008**. Belo Horizonte: Cedeplar, Universidade Federal de Minas Gerais, 2010. (Working Paper, n. 386).

RIOS-NETO, E.; GUIMARÃES, R. R. M. The demography of education in Brazil: inequality of educational opportunities based on Grade Progression Probability (1986-2008). **Vienna Yearbook of Population Research**, v. 8, n. 1, p. 283-312, 2010.

ROGER, R. G. Marriage, sex and mortality. **Journal of Marriage and the family**, v. 57, n. 2, p. 515-526, 1995.

ROGER, R. G. The effects of family composition, health, and social support linkages on mortality. **Journal of Health and Social Behavior**, v. 37, n. 4, p. 326-338, 1996.

ROGERS, R. G.; HUMMER, R. A.; NAM, C. B. **Living and dying in the USA: behavioral, health, and social differences of adult mortality**. Academic Press, 1999. Cap 7. The effects of basic socioeconomic factors on mortality, p. 115-139.

ROSENZWEIG, M. R.; SCHULTZ, T. P. Estimating a household production function: heterogeneity, the demand for health inputs, and their effects on birthweight. **Journal of Political Economy**, v. 91, n. 5, p. 723-46, 1983.

ROSS C. E.; MASTERS, R. K.; HUMMER, R. R. A. Educational and the gender gaps in health and mortality. **Demography**, v. 49, p. 1157-1183, 2012.

RUBIN, D. B. Inference and missing data. **Biometrika**, v. 63, n. 3, p. 581-592, 1976.

RUBIN, D. B. Multiple imputation after 18+ years. **Journal of the American Statistical Association**, v. 91, n. 434, p. 473-489, 1996.

SANDOVAL, M. H.; TURRA, C. M. El gradiente educativo en la mortalidad adulta en Chile. **Revista Latinoamericana de Población**, v. 17, n. 9, p. 7-35, 2015.

SASSON, I. Trends in life expectancy and lifespan variation by educational attainment: United States, 1990-2010. **Demography**, v. 53, p. 269-293, 2016.

SCHAFFER, J. L. **Analysis of incomplete multivariate data**. London: Chapman & Hall, 1997.

SCHAFFER, J. L.; GRAHAM, J. W. Missing data: our view of the state of the art. **Psychological Methods**, v. 7, p. 147-177, 2002.

SHKOLNIKOV, V. M. *et al.* Educational level and adult mortality in Russia: an analysis of routine data 1979 to 1994. **Social Science & Medicine**, v. 47, n. 3, p. 357-369, 1998.

SICKLES, R. C.; TAUBMAN, P. Mortality and morbidity among adults and the elderly. *In*: ROSENZWEIG, M.; STARK, O. **Handbook of population and family economics**. Elsevier, 1997.

SILVA, L. E.; FREIRE, F. H. M. A.; PEREIRA, R. H. M. Diferenciais de mortalidade por escolaridade da população adulta brasileira. **Cadernos de Saúde Pública**, v. 32, n. 4, p. 1-12, 2016.

SUBRAMANIAN, S. V. *et al.* The mortality divide in India: the differential contributions of gender, caste, and standard of living across the life course. **American Journal of Public Health**, v. 96, n. 5, p. 818-825, 2006.

TURRA, C. M.; PEREZ, E. R.; GUIMARÃES, R. The effect of changes in educational composition on adult female mortality in Brazil. **Research on Aging**, v. 38, n. 3, p. 283-298, 2016.

US PUBLIC HEALTH SERVICES. **US Surgeon General's Advisory Committee on Smoking and Health, Smoking and Health**. Washington DC, 1964.

VALKONEN, T. Adult mortality and level of education: a comparison of six countries. *In*: FOX, J. **Health inequalities in European countries**. Aldershot: Gower, 1989. p. 142-162.

WAJNMAN, S. **Demografia da família e dos domicílios brasileiros**. Tese (Professor titular) – Cedeplar, Universidade Federal de Minas Gerais, Belo Horizonte, 2012.

WALQUE, D.; FILMER, D. Trends and socioeconomic gradients in adult mortality around the developing world. **Population and Development Review**, v. 39, n. 1, p. 1-29, 2013.

WILLIAMS, D. F. Socioeconomic differentials in health: a review and redirection. **Social Psychology Quarterly**, v. 53, n. 2, p. 81-99, 1990.

Sobre os autores

Mirian Martins Ribeiro é doutora e mestre em Demografia pela Universidade Federal de Minas Gerais (UFMG) e graduada em Ciências Econômicas pela UFMG. Professora adjunta da Universidade Federal de Ouro Preto, Instituto de Ciências Sociais Aplicadas, Departamento de Ciências Econômicas.

Cassio Maldonado Turra é doutor em Demografia pela University of Pennsylvania, mestre em Demografia pelo Centro de Desenvolvimento e Planejamento Regional, da Universidade Federal de Minas Gerais (Cedeplar/UFMG) e graduado em Economia pela UFMG. Professor associado do Departamento de Demografia da UFMG. Bolsista de produtividade do CNPq.

Cristine Campos de Xavier Pinto é doutora em Economia pela University of California, Berkeley, mestre em economia pela PUC-Rio, mestre em Estatística pela University of California e graduada em Ciências Econômicas pela Universidade Federal de Minas Gerais (UFMG). Professora assistente da Escola de Economia de São Paulo, da Fundação Getúlio Vargas (FGV).

Endereço para correspondência

Mirian Martins Ribeiro

Universidade Federal de Ouro Preto, Instituto de Ciências Sociais Aplicadas
Rua do Catete, 166, Centro
35420-000 – Mariana-MG, Brasil

Cassio Maldonado Turra

Universidade Federal de Minas Gerais, Centro de Desenvolvimento e Planejamento Regional
Av. Antônio Carlos, 6626, sala 3006, Pampulha
31270-901 – Belo Horizonte-MG, Brasil

Cristine Campos de Xavier Pinto

Fundação Getúlio Vargas
Rua Itapeva, 474, 12º andar, Bela Vista
01332-000 – São Paulo-SP, Brasil

Abstract

Adult mortality by education level in São Paulo: estimates from different methodological strategies

In this paper, we estimate educational differentials for adult mortality in São Paulo. We offer a comparative analysis of estimates based on data from the 2010 Census and Sistema de Informações sobre Mortalidade (SIM) – Datasus, and three different ways of measuring schooling:

registered in the SIM, reported in the Census for the head of household, and statistically imputed in the Census of deceased individuals. For imputations of schooling, we used the Dempester method (1977), which proposes the use of the expectation-maximization algorithm (E-M algorithm) to deal with missing data. We considered three levels of education (low, medium, and high) and we estimated mortality rates based on Poisson models. Results indicate that schooling can reduce mortality rates between 25 and 59 years of age by up to 77%. In addition, in a country where the population has a low educational level, secondary education represents a significant gain from the point of view of adult survival (about 50%). We found similar patterns of mortality by education for estimates with data recorded in the SIM and data imputed in the 2010 Census. In addition, our analysis suggests that estimates assuming the educational level of the head of the household result in atypical mortality gradient by education, probably distorted by the educational transition in Brazil. We hope that the imputation model proposed here can be used in future analyses of mortality from the 2010 Census.

Keywords: Adult mortality. Missing data. Imputation. Socioeconomic differentials in mortality. Brazil. Education.

Resumen

Mortalidad adulta por nivel de escolaridad en San Pablo: análisis comparativo a partir de diferentes estrategias metodológicas

En este artículo estimamos los diferenciales educativos de la mortalidad de adultos en San Pablo. Ofrecemos un análisis comparativo de estimaciones con base en datos del censo de 2010 y el Sistema de Información de Mortalidad (SIM) – Datasus, y tres formas diferentes de medir la escolaridad: registrada en el SIM, declarada en el censo por el jefe de hogar e imputado estadísticamente en el censo para las personas fallecidas. Para las imputaciones de escolaridad se utilizó el método de Dempester (1977), que propone el uso del algoritmo de maximización de esperanza (algoritmo E-M) para tratar los datos faltantes. Consideramos tres niveles de educación (bajo, medio y alto) y estimamos las tasas de mortalidad con base en los modelos de Poisson. Los resultados indican que la escolarización puede reducir las tasas de mortalidad entre los 25 y 59 años hasta en un 77 %. Además, en un país donde la población tiene bajo nivel de educación, completar la educación secundaria representa una ganancia significativa desde el punto de vista de la supervivencia de los adultos (alrededor del 50%). Encontramos patrones similares de mortalidad por educación para las estimaciones obtenidas con datos registrados en el SIM y datos imputados en el Censo de 2010. Además, nuestro análisis sugiere que las estimaciones asumiendo la educación del jefe de hogar dan como resultado diferenciales de mortalidad atípicos, probablemente distorsionados por la transición de educación en Brasil. Esperamos que el modelo de imputación propuesto aquí se pueda utilizar en futuros análisis de mortalidad del Censo de 2010.

Palabras clave: Mortalidad de adultos. Datos faltantes. Imputación. Diferenciales socioeconómicos en la mortalidad. Brasil. Educación.

Recebido para publicação em 09/08/2020

Aceito para publicação em 29/12/2020