

Qualidade da informação da idade nos registros de óbito no Brasil, 1996-2015

Quality of the information about age on the death registers in Brazil, 1996-2015

Calidad de la información de la edad en los registros de óbito en Brasil, 1996-2015

Fatima Valéria Lima Jacques^{1,a}

fatimajacques@yahoo.com.br | <https://orcid.org/0000-0002-4568-2485>

José Henrique Costa Monteiro-da-Silva^{2,b}

jose.henrique@nepo.unicamp.br | <https://orcid.org/0000-0001-9674-3041>

Raphael Mendonça Guimarães^{3,c}

raphael24601@gmail.com | <https://orcid.org/0000-0003-1225-6719>

¹ Universidade Federal do Rio de Janeiro, Instituto de Pediatria e Puericultura Margatão Gesteira, Serviço de Emergência Pediátrica. Rio de Janeiro, RJ. Brasil.

² Universidade Estadual de Campinas, Instituto de Filosofia e Ciências Humanas, Núcleo de Estudos de População. Campinas, SP. Brasil.

³ Fundação Oswaldo Cruz, Escola Politécnica de Saúde Joaquim Venâncio, Laboratório de Educação Profissional em Informação e Registro em Saúde. Rio de Janeiro, RJ. Brasil.

^a Mestrado em Saúde Pública pela Universidade Federal do Rio de Janeiro.

^b Mestrado em Demografia pela Universidade Estadual de Campinas.

^c Doutorado em Saúde Pública pela Universidade Federal do Rio de Janeiro.

Resumo

Este artigo apresenta os resultados de uma pesquisa que teve como objetivo analisar a qualidade da declaração da idade nos registros de óbito no Brasil, de 1996 a 2015. Foi realizada uma análise por 'idade simples' dos microdados de óbitos no Brasil no período mencionado. A preferência por dígitos terminais 0 e 5 foi avaliada usando o índice de Whipple (IW). Já a preferência pelos dígitos terminais de 0 a 9 foi expressa usando o método de Myers (IM). A qualidade dos dados de idade foi alta no período [IW_{tot} = 0,55 – 0,83 (masculino) e 0,71 – 0,93 (feminino); IM = 0,388 – 1,004 (masculino) e 0,430 – 1,589 (feminino)]. A qualidade da informação foi mais satisfatória entre homens e não houve tendência significativa a uma melhora, sugerindo sua estabilidade durante os 20 anos analisados. Foi encontrada preferência pelo dígito terminal 0 (zero) principalmente entre mulheres. Concluiu-se que os dados de óbito no Brasil, com relação à idade, são satisfatórios, podendo ser utilizados em análises demográficas e epidemiológicas.

Palavras-chave: Qualidade de dados; Índice de Whipple; Índice de Myers; Mortalidade; Sistemas de informação.

Abstract

This article presents the results of a research aimed at analyzing the quality of the information about the age on the death registers in Brazil, from 1996 to 2015. An analysis was performed by simple age of the deaths microdata in Brazil for that period. The preference for the last digits 0 and 5 was evaluated using the Whipple index (IW), while the preference for the last digits from 0 to 9 was expressed using the Myers (IM) method. The quality of the age data was high in the period [IWTtot = 0.55 - 0.83 (male) and 0.71 - 0.93 (female); IM = 0.388 - 1.004 (male) and 0.430 - 1.589 (female)]. The quality of the information was more satisfactory among men, and there was not a significant trend in the improvement suggesting stable quality during the 20 years analyzed. The preference was given to the last digit 0, mainly among women. It was concluded that data from death registers in Brazil regarding the age are satisfactory and can be used in demographic and epidemiological analyses.

Keywords: Data quality; Whipple index; Myers index; Mortality; Information systems.

Resumen

Este artículo presenta los resultados de una investigación que tuvo como objetivo analizar la calidad de la declaración de la edad en los registros de óbito en Brasil, desde 1996 hasta 2015. Se realizó el análisis por edad simple de los microdatos de óbitos en Brasil en el periodo mencionado. La preferencia por dígitos finales 0 y 5 fue evaluada usando el índice de Whipple (IW). La preferencia por los dígitos finales desde 0 hasta 9 fue expresada usando el método de Myers (IM). La calidad de los datos de edad fue alta en el período [(IWTot = 0,55 - 0,83 (masculino) y 0,71 - 0,93 (femenino), IM = 0,388 - 1,004 (masculino) y 0,430 - 1,589 (femenino)]. La calidad de la información fue más satisfactoria entre hombres y no hubo tendencia significativa en la mejora, sugiriendo estabilidad en la calidad en los 20 años analizados. Se encontró una preferencia por el dígito terminal 0 principalmente entre mujeres. Se concluye que los datos de óbito en Brasil con relación a la edad son satisfactorios y pueden ser utilizados en análisis demográficos y epidemiológicos.

Palabras clave: Calidad de datos; Índice de Whipple; Índice de Myers; Mortalidad; Sistemas de información.

INFORMAÇÕES DO ARTIGO

Contribuição dos autores:

Concepção e desenho do estudo: José Henrique Costa Monteiro da Silva e Raphael Mendonça Guimarães.

Aquisição, análise ou interpretação dos dados: Fátima Valéria Lima Jacques e Raphael Mendonça Guimarães.

Redação do manuscrito: Fátima Valéria Lima Jacques, José Henrique Costa Monteiro da Silva e Raphael Mendonça Guimarães.

Revisão crítica do conteúdo intelectual: Fátima Valéria Lima Jacques e Raphael Mendonça Guimarães.

Declaração de conflito de interesses: este trabalho não apresenta conflito de interesses.

Fontes de financiamento: nenhuma.

Considerações éticas: o presente estudo utilizou dados secundários provenientes de bases públicas, sem qualquer tipo de identificação individual. Desta forma, em conformidade com a resolução 466/2012, e com a resolução 510/2016, o projeto ficou isento de submissão ao Comitê de Ética em Pesquisa.

Agradecimento/Contribuições adicionais: não há.

Histórico do artigo: submetido: 06 set. 2018 | aceito: 27 dez. 2018 | publicado: 29 mar. 2019.

Apresentação anterior: não houve.

Licença CC BY-NC atribuição não comercial. Com essa licença é permitido acessar, baixar (download), copiar, imprimir, compartilhar, reutilizar e distribuir os artigos, desde que para uso não comercial e com a citação da fonte, conferindo os devidos créditos de autoria e menção à Reciis. Nesses casos, nenhuma permissão é necessária por parte dos autores ou dos editores.

Introdução

Na maioria das pesquisas, os dados quantitativos são coletados por meio de questionários e entrevistas. Quando esses instrumentos são autoaplicados, frequentemente, ocorre subestimação ou superestimação sistemática de certas informações e parâmetros clínicos, o que pode levar a erros de interpretação¹.

A idade é uma variável demográfica amplamente utilizada para análises descritivas e estatísticas da estrutura populacional e para previsão do crescimento da população, já que muitos dados demográficos e socioeconômicos são atribuídos à idade e ao sexo². Além disso, a relevância da informação da idade na epidemiologia e saúde pública é evidente, uma vez que a idade do indivíduo é considerada em diversas situações, desde o desenvolvimento do perfil de risco até o diagnóstico, manejo e prognóstico de uma doença³.

O registro incorreto de informações é um fenômeno comum nos países em desenvolvimento e constitui um dos maiores desafios à demografia⁴. As irregularidades mais comuns relacionadas ao registro de idade são as preferências por certos dígitos e o aumento da idade em torno de algumas idades^{5,6}. Esses dados frequentemente exibem alta ocorrência em idades redondas ou atraentes, como números pares e múltiplos de cinco⁵. Desta forma, a avaliação da informação sobre a idade é considerada uma medida de qualidade e consistência dos dados⁶.

Quando medida adequadamente, a probabilidade de frequência de cada dígito terminal (0 a 9) é de cerca de 10%. Há, contudo, uma maior frequência observada para os dígitos finais zero ou cinco^{5,7}. Essa preferência por arredondar números pode mascarar ou exagerar as diferenças reais entre as populações e também explicar por que as diferenças entre as estimativas medidas e relatadas pelo próprio indivíduo variam entre as culturas. Para contabilizar esse viés e explorar todo o potencial de detecção da fabricação de dados, as preferências por declaração de dígitos finais devem ser avaliadas em diferentes fontes de informação⁸.

A precisão dos dados relativos ao registro de idade varia de um país para outro e depende de inúmeros fatores, associados principalmente a problemas na coleta e tabulação, bem como à falta de um fluxo de informações eficiente nos registros administrativos e pesquisas populacionais⁹⁻¹¹. É importante destacar que a boa acurácia da informação sobre a idade é vital para a formulação adequada de políticas públicas. No que diz respeito aos censos demográficos e às pesquisas sobre outras informações populacionais, como a Pesquisa Nacional por Amostras de Domicílio (Pnad), há uma rotina de avaliação da qualidade das declarações, assim como da magnitude e do dimensionamento de eventuais erros existentes¹. Esta rotina inclui estratégias que vão desde o desenho da amostra mestra, através do planejamento do marco amostral, passando pela criação de perguntas, e o uso de ferramentas e sistemas no processamento de dados. Os registros administrativos, no entanto, devido à sua natureza, possuem uma lógica distinta de avaliação da qualidade dos dados, e isto representa muitas vezes um desafio para a obtenção de estatísticas confiáveis de mortalidade, que muitas vezes consistem nos únicos indicadores disponíveis para a realização de análises da situação de saúde da população¹². Neste sentido, o objetivo deste estudo foi analisar a qualidade da declaração da idade nos registros de óbito no Brasil, segundo o sexo de 1996 a 2015.

Métodos

Desenho de estudo

Trata-se de estudo ecológico para a avaliação da qualidade da informação.

Fonte de dados

Foram utilizados os microdados referentes aos óbitos no Brasil, por unidade da federação (UF), posteriormente compilados em um banco de dados nacional, no período de 1996 a 2015. Os dados foram extraídos do Sistema de Informações sobre Mortalidade (SIM), disponibilizados pelo Ministério da Saúde.

Análise de dados

Com o objetivo de assegurar maior precisão na análise, considerou-se o diferencial por sexo. Primeiramente, foi realizada uma inspeção visual da pirâmide etária por idade simples, verificando distorções em torno de dígitos específicos (como o 0 e 5, por exemplo), para idades como 40, 45, 50, 60 anos. Nestas idades, quando há problemas na qualidade dos dados, é possível observar pontos de inflexão na distribuição etária para ambos os sexos.

Para avaliar a qualidade das declarações das idades nos registros de óbito, no período analisado, foram utilizados dois métodos que indicam o nível de qualidade desses dados. Estes indicadores são extensamente descritos no Manual II das Nações Unidas¹³. A preferência por dígitos na declaração da variável ‘idade’, segundo o sexo, foi determinada usando os métodos de Whipple e Myers, e avaliada estatisticamente utilizando-se teste de aderência baseado no teste de qui-quadrado de Pearson, assumindo que a frequência esperada de cada dígito seja a mesma (ou seja, 10%). Estes índices são aplicáveis quando a idade é referida em ‘anos simples’.

Índice de Whipple (IW)

O método de Whipple é aplicado para detectar preferências por dígitos terminais 0 e 5 no grupo populacional de 23 a 62 anos de idade. A opção por este intervalo de idade se dá pelo fato de se assumir que as idades mais jovens e mais idosas apresentam maior imprecisão na informação¹³.

O cálculo do Índice de Whipple para um dígito específico segue a seguinte lógica: o numerador representa a soma das populações com idades terminadas no dígito pelo qual elas têm preferência e que se deseja avaliar, dentro do intervalo de 23 a 62 anos; o denominador representa a soma de todas as populações deste intervalo etário dividida por 10, por se tratar de uma avaliação de um dígito entre 10 possíveis. A preferência pelo dígito 5, índice IW_5 , por exemplo, é ilustrada na equação (1). As variáveis P_x representam as populações com declaração de idade x .

$$IW_5 = \frac{P_{25} + P_{35} + \dots + P_{55}}{\frac{1}{10} * (P_{23} + P_{24} + \dots + P_{62})} \times 100 \quad (1)$$

Considera-se, para a classificação da qualidade, a partir deste índice, o disposto na Tabela 1.

Tabela 1 – Classificação de atração geral pelo Índice de Whipple

Nível de Atração	Índice de Whipple
Preciso	99,0 – 104,9
Pouco preciso	105,0 – 109,9
Aproximado	110,0 – 124,9
Grosseiro	125,0 – 174,9
Muito grosseiro	175,0 ou mais

Fonte: United Nations (1955)¹³.

O cálculo para o conjunto de idades que terminam em 0 ou 5 foi feito de maneira análoga ao descrito anteriormente e é apresentado na equação (2).

$$IW_{0,5} = \frac{P_{25} + P_{30} + \dots + P_{60}}{\frac{1}{5} * (P_{23} + P_{24} + \dots + P_{62})} \times 100 \quad (2)$$

Adicionalmente, foi utilizado o Índice de Whipple modificado, que permite a análise de atração global por todos os dígitos¹⁴. A modificação pôs no denominador a soma da população por grupos quinquenais nos quais a idade com o dígito a ser analisado é o ponto médio do grupo, calculado da forma demonstrada pela equação (3):

$$IW_{m_1} = \frac{5 * (P_{31} + P_{41} + \dots + P_{61})}{({}_5P_{29} + {}_5P_{39} + \dots + {}_5P_{59})} \quad (3)$$

Finalmente, o índice total de Whipple, modificado, apresenta-se na equação (4). Conforme a equação ilustra, os resultados deste índice são discutidos a partir do desvio do índice modificado apresentado na equação (3) com relação à unidade.

$$W_{tot} = \sum_{i=0}^9 |IW_{m_i} - 1| \quad (4)$$

Neste caso, considera-se que, quanto mais próximo de zero, melhor é a qualidade da informação.

Índice de Myers (IM)

O método de Myers, por sua vez, permite determinar a preferência para cada dígito terminal (0 a 9) nas idades que variam de 10 a 99 anos. Este método assume que a distribuição por dígito terminal é uniforme (ou seja, 10% de frequência para cada dígito). Através deste índice, é possível verificar a atração por dígitos, o que implica na qualidade da informação. Para a implementação deste método, a Organização das Nações Unidas propôs, em 1955, a soma da população que possui o mesmo dígito final na idade declarada (P_i) para o grupo de 10 a 89 anos (G1) e para o grupo de 20 a 99 anos (G2). Neste método, a parcela da população com idade acima de 100 anos não é contabilizada, pois supõe-se que ela não afete significativamente os resultados¹⁴. Como a população tende a ser menor com o avanço dos dígitos finais (ou seja, para cada dígito sucessivo, a população é mais velha e menor que a população anterior), essas populações são multiplicadas por coeficientes (x) no intervalo de 1 a 10 para o grupo G1, e pelos complementos (10 – x) para o grupo G2, conforme as equações (5) e (6):

$$G_1(i) = (i + 1) * \sum_{\alpha=10}^{97} P_i, \text{ para todo } i \in \{0, 1, \dots, 9\} \quad (5)$$

$$G_2(i) = (9 - i) * \sum_{\alpha=20}^{99} P_i, \text{ para todo } i \in \{0, 1, \dots, 9\} \quad (6)$$

Considera-se que a frequência (f_i) que cada dígito i tem no total é dada pela soma de G_1 e G_2 . Supõe-se, ainda, que a proporção esperada de cada dígito equivale a 10% do total (suposição da distribuição uniforme)¹³. O índice de Myers para cada dígito final da idade declarada é calculado como o desvio de f_i em relação aos 10% esperados, enquanto o índice geral é dado pelo somatório dos índices de cada dígito, conforme as fórmulas (7) e (8):

$$IM_i = |100 \times f_i - 10| \quad (7)$$

$$IM = \sum_{i=0}^9 IM_i, \text{ para } 0 \leq IM_i \leq 180 \quad (8)$$

Considera-se, para efeito de classificação, os patamares definidos na Tabela 2.

Tabela 2 – Classificação de atração geral pelo Índice de Myers

Nível de Atração	Índice de Myers
Baixo	até 4,9
Mediano	de 5,0 a 14,9
Alto	de 15,0 a 29,9
Muito alto	de 30,0 a 180,0

Fonte: United Nations (1955)¹³.

Com a obtenção dos índices de Whipple e Myers por ano, no período de 1996 a 2015, foi realizada uma análise de séries temporais, para avaliar a evolução da qualidade da informação da idade nos registros de óbito. Inicialmente, através do teste de Wald-Wolowitz, verificou-se que a série temporal ficou estacionada e, através do teste de Cox-Stuart, foi averiguado o efeito da tendência. Em seguida, após estas verificações, a tendência de mortalidade foi analisada utilizando-se a técnica de regressão polinomial, com a qual foi avaliada a tendência dos índices no período. A variável considerada independente, a fim de evitar colinearidade, foi o ano-centralizado pelo ponto médio do período ($x-2006$). Foram testados os modelos de regressão linear simples de segundo e terceiro grau. Foi escolhido como modelo mais adequado aquele que apresentou o melhor ajuste do coeficiente de determinação (R^2), significância estatística (considerando o nível de 5%) e análise dos resíduos. Nos casos em que os modelos apresentaram similaridade, foram escolhidos os que apresentaram o modelo mais simples, pelo princípio da parcimônia¹⁵. O software utilizado para o estudo foi o R, versão 3.5.1.

Resultados

Em um período de 20 anos, o Brasil apresentou pouco mais de 21 milhões de óbitos, distribuídos entre as diversas faixas etárias. O Gráfico 1 apresenta as mudanças observadas por ‘idade simples’, estratificada por sexo. Os gráficos permitem dizer que há uma distinção importante entre os sexos. O sexo masculino possui um padrão de mortalidade com destaque nas faixas etárias jovens, que reduz a partir dos 30 anos, voltando a crescer nas faixas mais avançadas, entre idosos. Já o sexo feminino possui mortalidade de padrão

crescente com o avançar da idade, padrão que não se verifica nas mulheres jovens. Especula-se que esta diferença de acordo com o sexo se dê graças aos óbitos por causas externas, que possuem maior expressividade no sexo masculino. Ainda é importante destacar a mudança entre 1996 e 2015, corroborando a análise anterior, de que há, indiscutivelmente, um processo de compressão da mortalidade em curso.

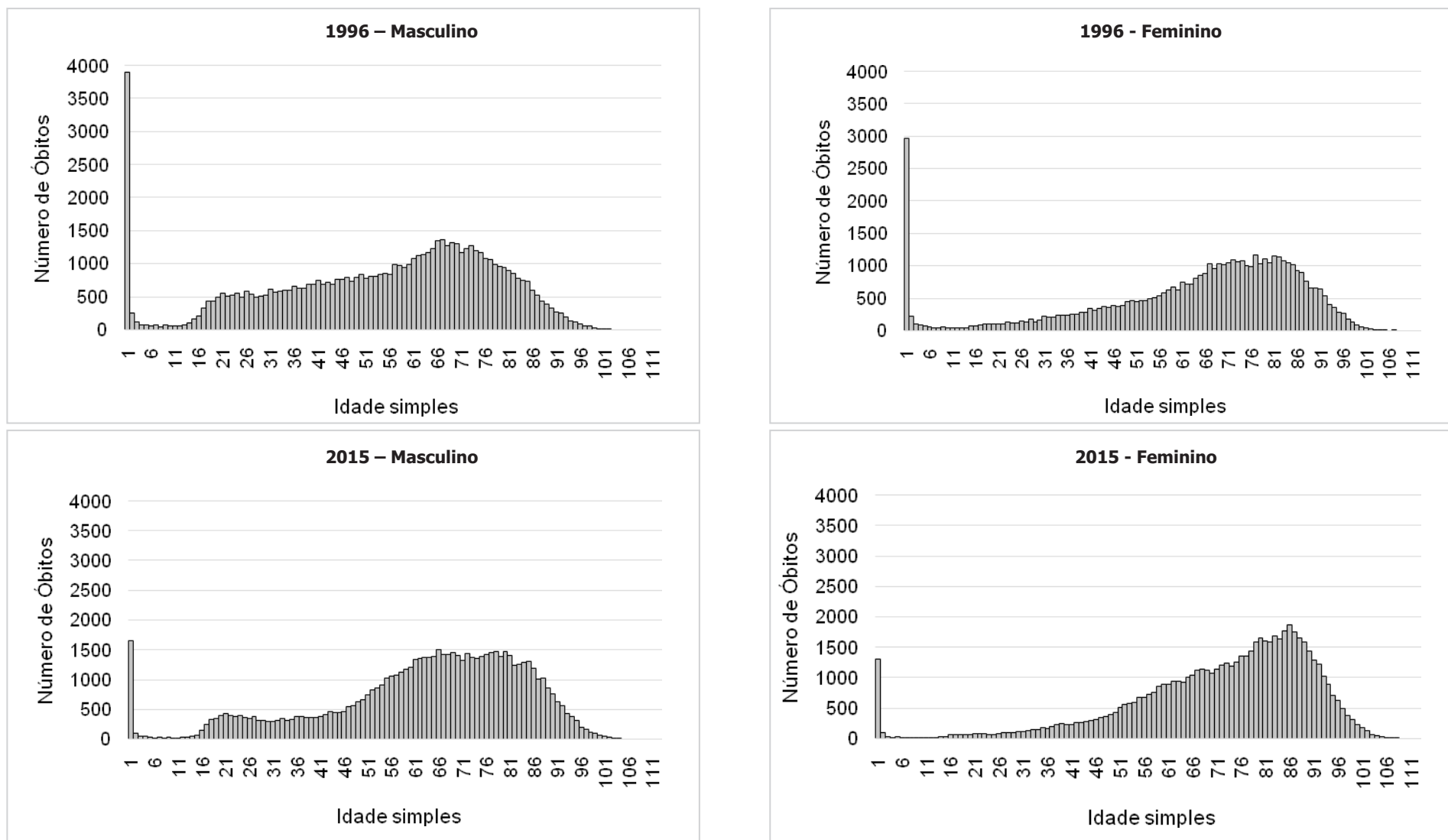


Gráfico 1 – Estrutura etária dos óbitos no Brasil, segundo sexo, 1996 e 2015
 Fonte: DATASUS (2018).

É possível dizer que há uma certa estabilidade na qualidade da informação sobre a variável idade. Isto se reflete nos dois índices avaliados (Whipple e Myers). Ao avaliar o desvio de dígitos, ou seja, a diferença entre a frequência observada do dígito terminal e a esperada (10%), através do índice de Myers (Gráfico 2), pode-se perceber que, desde 1996, a informação já apresentava boa qualidade, apresentando moderada variabilidade, de acordo com o dígito de preferência, ainda que, considerando a unidade de medida (%), os valores de desvio sejam pequenos. Notadamente, os valores calculados do índice no período sugerem uma baixa atração por dígito, conforme os patamares estabelecidos na Tabela 2. Isto sugere que esta diferença com relação ao patamar esperado de 10% para cada dígito seja reflexo da flutuação aleatória dos dados. Destaca-se maior variabilidade no dígito zero, e menor variabilidade no dígito 1. Observa-se ainda diferença entre sexos, com alguma vantagem na qualidade da informação do sexo masculino.

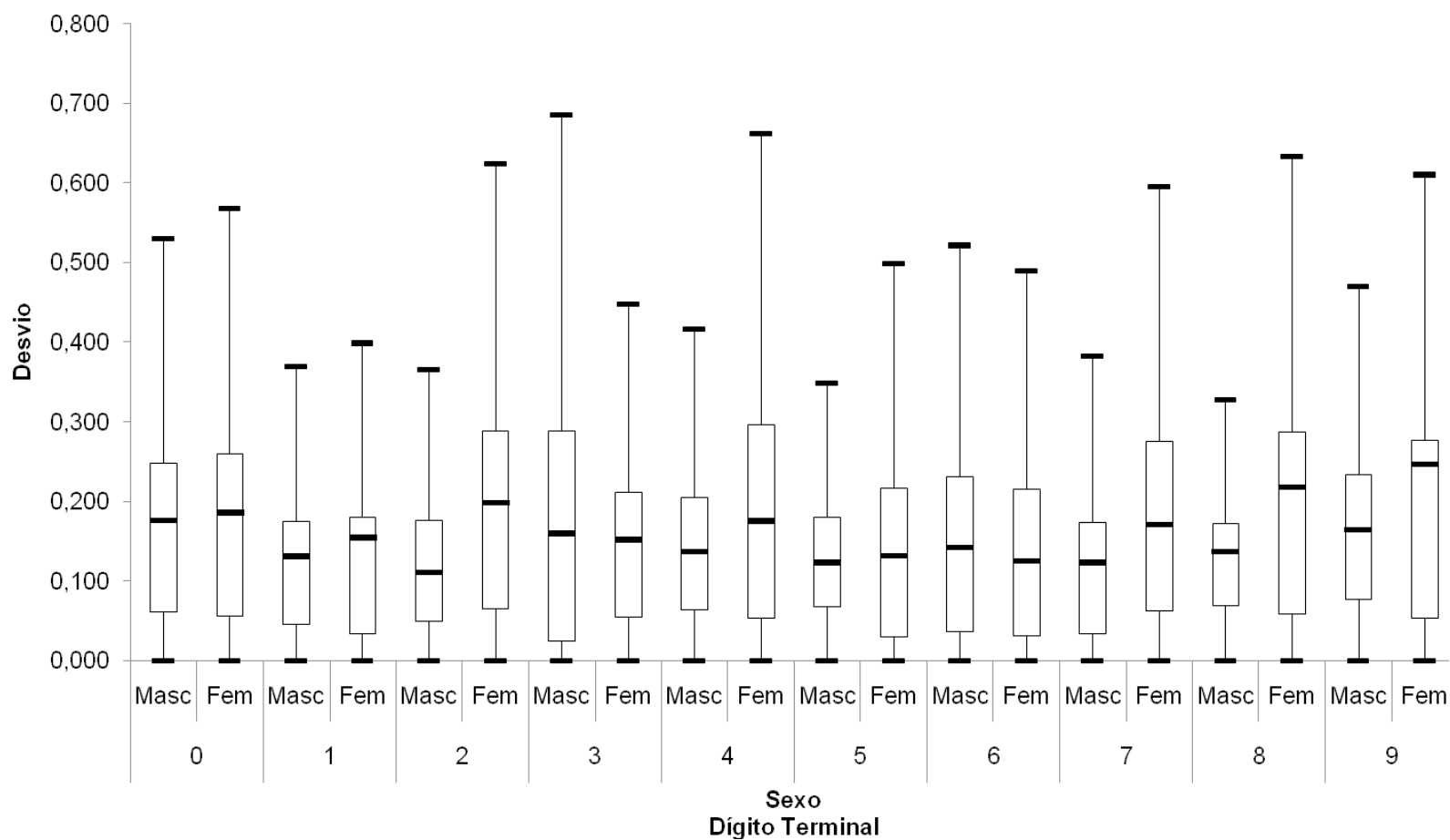
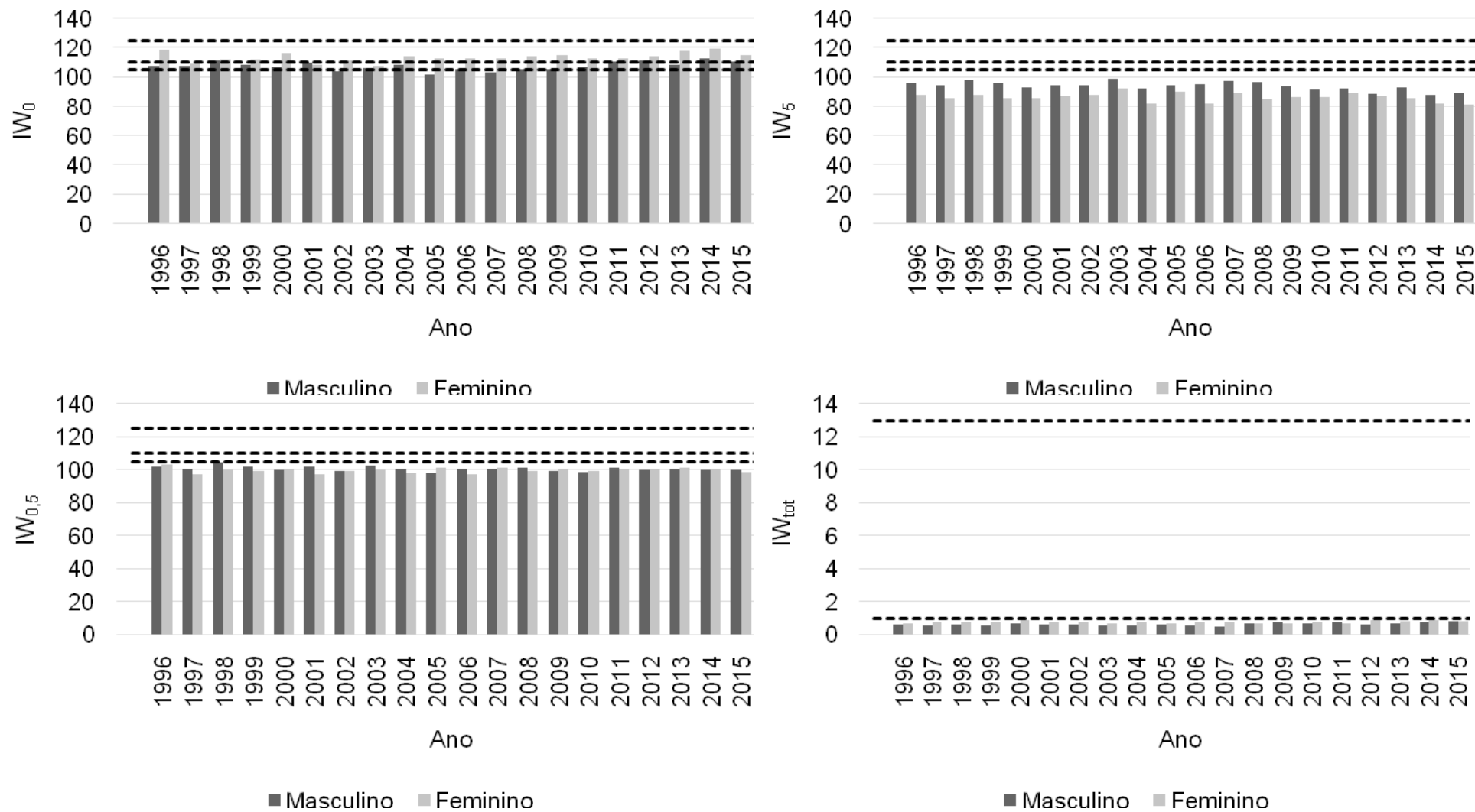


Gráfico 2 – Descrição da atração por dígito terminal pelo Índice de Myers a partir do desvio em torno do valor esperado do índice, segundo sexo e dígito terminal. Brasil, 1996 a 2015
 Fonte: DATASUS (2018).

O Gráfico 3 permite avaliar a atração por dígitos a partir do índice de Whipple e suas variações. De uma forma geral, considerando a tendência de maior empilhamento de dados entre os dígitos 0 e 5, pode-se dizer que a sua qualidade é satisfatória. Isto se reflete igualmente quando se avalia o índice total. Entretanto, ao estratificar os dígitos 0 e 5, percebe-se que a qualidade dos dados tende a reduzir quando analisado o dígito terminal zero, evidenciando alguma atração. De fato, como se observa no Gráfico 4, que fornece os índices totais modificados específicos por dígito para cada um dos sexos – nos anos de 1996 e 2015 – vê-se alguma diferença para os dígitos zero, um, dois, três, e quatro, para ambos os sexos, independentemente do ano de observação.



Limites das categorias de classificação do Índice (vide tabela 1 e equações 1 – 4)

Gráfico 3 – Indicador de atração por dígito terminal em dados de óbitos no Brasil com o Índice de Whipple em relação aos patamares de qualidade de declaração definidos na Tabela 1, 1996 a 2015
 Fonte: DATASUS (2018).

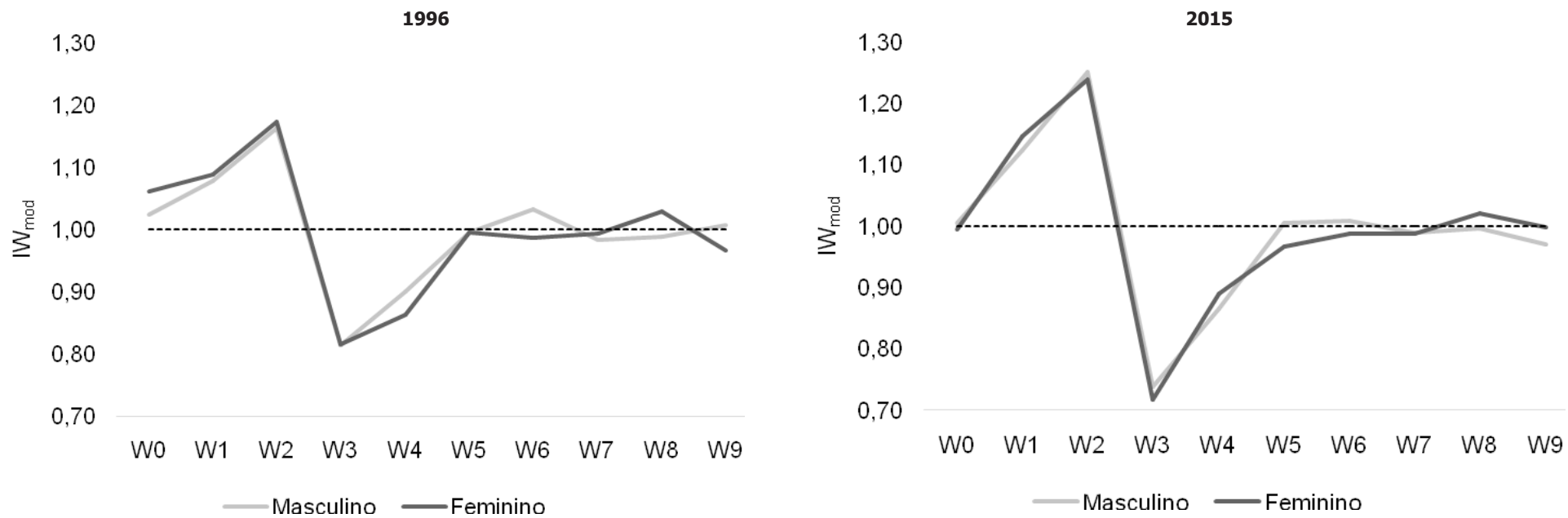


Gráfico 4 – Descrição da atração por dígito terminal pelo Índice de Whipple, segundo sexo e dígito terminal. Brasil, 1996 e 2015
Fonte: DATASUS (2018).

Finalmente, ao analisar a tendência dos dois índices (Gráfico 5), corrobora-se a hipótese da flutuação aleatória, já que a análise da série temporal mostra que não há tendência significativa no período. Cabe destacar, entretanto, que esta flutuação é mais evidente para o sexo feminino, em que a variabilidade parece ser maior.

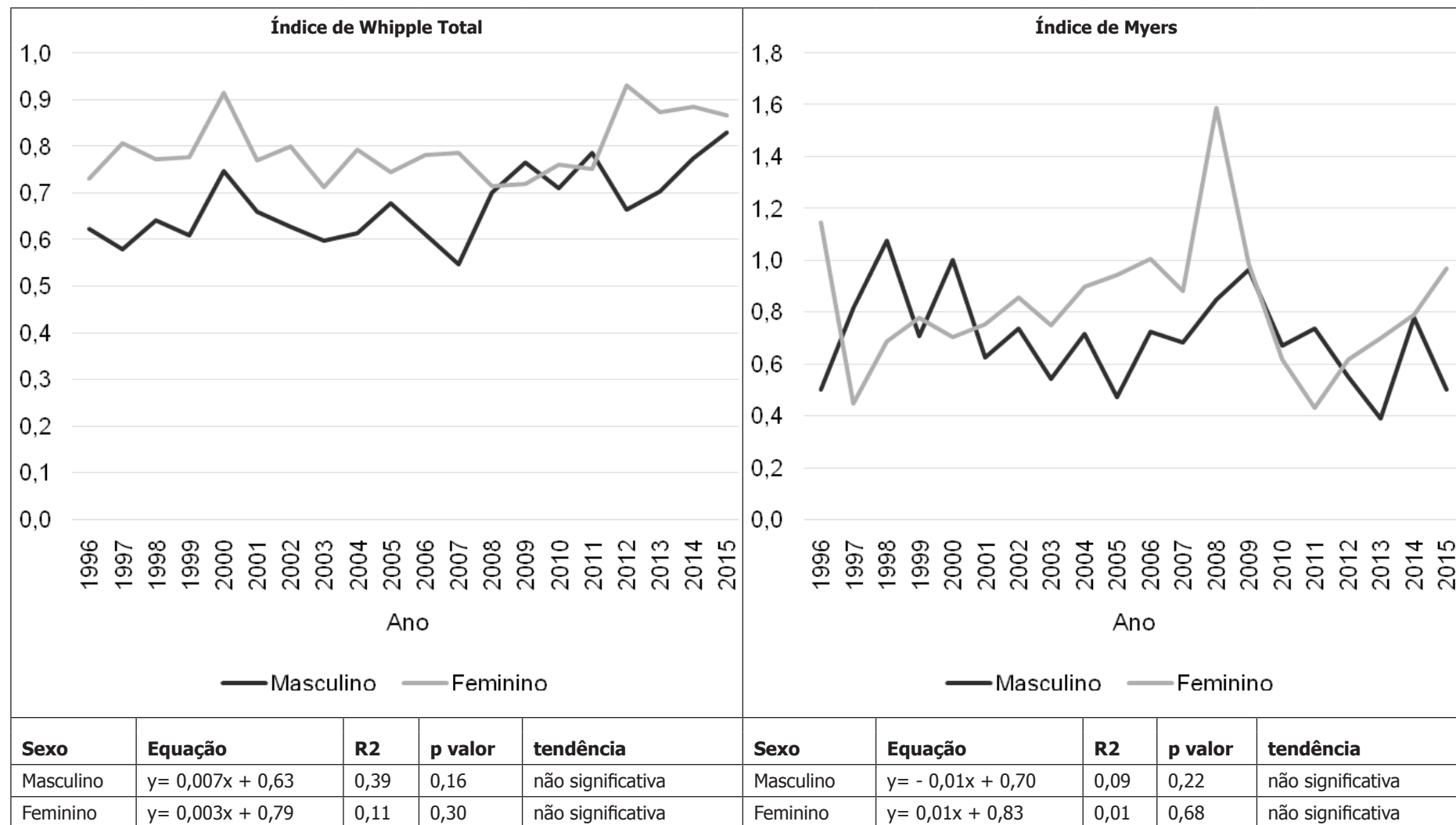


Gráfico 5 –Tendência temporal dos indicadores de atração de dígito terminal no Brasil, 1996 a 2015
 Fonte: DATASUS (2018).

Discussão

O presente estudo utilizou duas diferentes técnicas para avaliar a qualidade da variável ‘idade’ registrada nas declarações de óbito. Vale lembrar que inadequações no registro da idade são frequentes em certas fontes de dados, sendo mais comuns em países em desenvolvimento, impossibilitando o uso instantâneo dos dados sem que seja utilizada qualquer correção¹⁶.

A metodologia de coleta de dados utilizada neste estudo foi semelhante ao método utilizado em quase todos os estudos e pesquisas domiciliares. Considerando os dois índices estudados, infere-se que os dados de idade coletados nos registros de óbito podem ser considerados de boa qualidade. Houve discreta atração pelo dígito terminais 0, indicando uma preferência dos respondentes por este dígito na declaração de idade de óbito. Desta forma, a distribuição por idade sugere um nível de confiabilidade aceitável, sem a necessidade de ajustamentos ou suavizações. E, portanto, os dados de mortalidade permitem a construção de tábuas de vida, importante instrumento de avaliação da mortalidade e, para a demografia, esperança de vida. Um estudo anterior¹⁷ já sugeria este nível de qualidade para o Brasil, corroborando a presente análise longitudinal, que segue apresentando tal padrão.

Os valores obtidos para os índices IW e IM apontam uma discreta preferência por dígitos zero na base de dados do SIM. Contudo, ela não é grande o suficiente para pôr em xeque a qualidade dos registros. Uma das possíveis explicações para a boa qualidade da informação referente à idade é o fato de o preenchimento da declaração de óbito, assim como a da declaração de nascidos vivos, ser realizada mediante a apresentação de documento de identificação, ou feita diretamente a partir do prontuário do paciente^{18,19}, e não por declaração verbal dos familiares, uma prática comum em pesquisas domiciliares^{20,21}.

A qualidade da informação sobre a idade varia de acordo com certas características, como o sexo e a faixa etária. A respeito da idade, Randall e Coast²² propuseram a modificação no cálculo do Índice de Whipple (índice de Whipple totalmente modificado), para avaliar a qualidade de dados em idosos, analisando pesquisas domiciliares de países africanos. Sua conclusão é de que a qualidade da informação sobre idade é pobre para a maioria dos países da África do Sul. Os autores concluíram ainda que nos *surveys* de alguns países há omissão considerável da proporção de mulheres idosas. Para a América Latina, contudo, Romero e Freitez²³ mostraram a melhoria na qualidade da declaração de idade, ao comparar os censos da década de 1990 com os da década de 2000. Andrade, Brusse, Camargo, Pereira e De Maria²⁴, também analisando censos da América Latina, constataram que, em um contexto de estrutura etária rejuvenescida, independe a utilização do Índice de Whipple ou de Myers. Contudo, em países com estrutura etária mais envelhecida, existe maior peso da qualidade de dados dos idosos e a escolha do indicador deve ser pensada com atenção. Já com relação ao sexo, estudos prévios que utilizaram o Índice de Whipple e o de Myers demonstraram que a preferência digital ocorre de forma mais frequente quando se trata de pessoas do sexo feminino²⁵⁻²⁷. Adicionalmente, nota-se uma maior atração pelos dígitos 0 e 5²⁷⁻²⁸.

A qualidade de registro da variável ‘idade’ é importante porque a distribuição por idade não é apenas uma parte invariável de um relatório de pesquisa, mas a introdução de vieses relacionados a ela pode levar a inferências erradas. Embora haja rotina cuidadosa para obtenção dos dados, alguns problemas podem surgir; por exemplo: existem os chamados ‘erros sistemáticos’, que são aqueles cometidos por problemas na forma de coleta de dados (entre eles, a coleta de informação de pessoas por terceiros), ou na seleção inadequada de indivíduos a serem pesquisados (por exemplo, selecionar pessoas apenas de determinada faixa etária ou sexo)²⁹. Sabe-se que a aproximação da consciência etária, ou seja, a referência à idade aproximada, e não precisa, se manifesta no fenômeno da acumulação de idade, para dados de idade autorreferidos ou por procuração^{19,30}. Quando a informação é fornecida por *proxy*, este problema torna-se ainda mais evidente. Por sua vez, o impacto de tal declaração incorreta pode levar a erros de classificação e à avaliação incorreta das taxas demográficas e, por conseguinte, interferir no planejamento de intervenções efetivas²⁹. A fim de minimizar esses erros, as rotinas de coleta de informação, na circunstância do óbito, passaram a ter critérios mais rigorosos, como a obtenção

dos dados por profissional específico, a dupla checagem da declaração de óbito e a checagem do documento de identidade, ao invés de usar apenas a informação referida por parente ou responsável.

A precisão dos dados relativos a idade em registros de saúde pode ser avaliada por meio de índices demográficos de controle de qualidade. Desta forma, devem ser utilizados métodos inovadores na coleta de dados, bem como técnicas estatísticas para a minimização de erros, visando à garantia da precisão dos dados de idade. O presente trabalho tem como objetivo a apresentação de um panorama geral da qualidade de declaração de idade dos óbitos entre 1996 e 2015 para o Brasil como um todo. Em vista disso, a sua principal limitação é a não desagregação das tendências regionais do país. Ainda assim, a atual análise representa um diagnóstico inicial da avaliação etária da qualidade dos dados do SIM e abre espaço para uma análise futura acerca de possíveis desigualdades regionais que poderiam ser observadas.

Considerações finais

A avaliação da qualidade dos dados de mortalidade se mostra importante, pois para diversas situações de saúde, envolvendo uma série de agravos, os dados de óbito são as únicas informações disponíveis. Por esta razão, é importante manter uma observação sistemática dos dados a fim de assegurar a qualidade do registro. Os resultados apontam para a boa qualidade dos dados, ainda que haja uma diferença entre os sexos. Isso indica a necessidade de avaliar situações específicas para a qualidade dos dados como, por exemplo, a causa de óbitos (assumindo que há riscos concorrentes e padrões distintos de mortalidade para cada causa). Uma vez que o uso de ferramentas demográficas parece adequado para sistemas de informações em saúde, aponta-se a necessidade de sua aplicação em outras bases que não apenas a de mortalidade, mas também as de informações sobre nascimentos, procedimentos ambulatoriais e internações hospitalares.

Referências

1. Hakkert R. Fontes de dados demográficos [Internet]. Belo Horizonte: Abep; 1996 [citado em 2018 abr. 25]. Disponível em: <http://www.ernestoamaral.com/docs/IndSoc/biblio/Hakkert1996.pdf>
2. Jannuzzi PM. Indicadores sociais no Brasil: conceitos, fontes de dados e aplicações para formulação e avaliação de políticas públicas, elaboração de estudos socioeconômicos. 2 ed. Rio de Janeiro: Alínea Editora; 2014.
3. Calazans ATS. Qualidade da informação: conceitos e aplicações. Transinformação. 2008;20(1): 29-45. doi: <http://dx.doi.org/10.1590/S0103-37862008000100003>
4. Del popolo F. Los problemas en la declaración de la edad de la población adulta mayor en los censos. Santiago de Chile: Cepal; 2000. (Serie Población y Desarrollo; n.8).
5. Luy M. Estimating Mortality Differences in Developed Countries From Survey Information on Maternal and Paternal Orphanhood. Demography. 2012;49(12):607-27. doi: <https://doi.org/10.1007/s13524-012-0101-4>
6. Bello Y. Error Detection in Outpatients' Age Data Using Demographic Techniques. Int J Pure Appl Sci Technol. 2012;10(1):27-36.
7. Burch TK. Error in Demographic and Other Quantitative Data and Analyses. Document de travail. Population Change and Lifecourse Strategic Knowledge Cluster Discussion Paper Series [Internet]. 2015 [cited 2019 jan 31];3(3):1-14. Available from: <https://goo.gl/JgGAXy>
8. Lyons-Amos M, Stones T. Trends in Demographic and Health Survey data quality: an analysis of age heaping over time in 34 countries in Sub Saharan Africa between 1987 and 2015. BMC Res Notes. 2017;10:760. doi: <https://dx.doi.org/10.1186%2Fs13104-017-3091-x>
9. Agrawal G, Khanduja P. Influence of Literacy on India's Tendency for Age Misreporting: Evidence from Census 2011. J Popul Soc Stud. 2015;23(1): 47-66. doi: <https://dx.doi.org/10.14456/jpss.2015.1>
10. Singh M. Understanding digit preferences in India using modified whipple index: an analysis of 640 districts of India. Int J Curr Res. 2017;9(1):45144-52.
11. ICF International. Demographic and Health Surveys Methodology: Questionnaires: Household, Woman's, and Man's [Internet]. Calverton, Maryland: ICF; 2011 [cited 2018 may 28]. Available from: <https://goo.gl/sSqC5K>

12. Lima EEC, Queiróz BL. Evolution of the deaths registry system in Brazil: associations with changes in the mortality profile, under-registration of death counts, and ill-defined causes of death. *Cad Saude Pública*. 2014;30(8):1721-30. doi: <http://dx.doi.org/10.1590/0102-311X00131113>
13. United Nation. Manuals on methods of estimating population: manual II: methods of appraisal of quality of basic data for population estimates. New York, NY: United Nations; 1955.
14. Andrade PG, Pereira ACR, Camargo KCM, Brusse GPL, Guimarães RM. Calidad de la declaración de la edad de las personas mayores en países de América Latina y el Caribe: análisis de los censos demográficos de las décadas de 1960 a 2010. *Notas de Población*. 2017;44(105): 53-84.
15. Latorre MRDO, Cardoso MRA. Análise de séries temporais em epidemiologia: uma introdução sobre os aspectos metodológicos. *Rev Bras Epidemiol*. 2001;4(3):145-52. doi: <http://dx.doi.org/10.1590/S1415-790X2001000300002>.
16. Denic S, Khatib F, Saadi H. Quality of age data in patients from developing countries. *J Public Health*. 2004;26(2):168-71. doi: <http://dx.doi.org/10.1093/pubmed/fdh131>
17. Paes NA, Albuquerque ME. Evaluation of population data quality and coverage of registration of deaths for the Brazilian regions. *Rev Saude Pública*. 1999;33(1):33-43. doi: <http://dx.doi.org/10.1590/S0034-89101999000100006>
18. Kanso S, Romero DE, Leite IC, Moraes EN. Diferenciais geográficos, socioeconômicos e demográficos da qualidade da informação da causa básica de morte dos idosos no Brasil. *Cad Saude Pública*. 2011;27(7):1323-39. doi: <http://dx.doi.org/10.1590/S0102-311X2011000700008>
19. Romero DE, Cunha CB. Avaliação da qualidade das variáveis epidemiológicas e demográficas do Sistema de Informações sobre Nascidos Vivos, 2002. *Cad Saude Pública*. 2007;23(3):701-14. doi: <http://dx.doi.org/10.1590/S0102-311X2007000300028>
20. Borkotoky K, Unisa S. Indicators to examine quality of large scale survey data: An example through District Level Household and Facility Survey. *PLoS One*. 2014; 9(3):1-11. doi: <https://doi.org/10.1371/journal.pone.0090113>
21. Pimienta R, Bolaños M. La declaración de la edad: un análisis comparativo de su calidad en los censos generales de población y vivienda. Zinacantepec: El Colegio Mexiquense;1999. (Documentos de Investigación; n.33)
22. Randall S, Coast E. The quality of demographic data on older Africans. *Demographic Research*. 2016;34(1):143-74. doi: <https://doi.org/10.4054/DemRes.2016.34.5>
23. Romero DE, Freitez A. Problemas de calidad de la declaración de edad de la población adulta mayor en los censos de América Latina de la ronda del 2000. In: 3º Congreso de la Asociación Latinoamericana de Población. Córdoba; 2008 sep. 23-27; [local desconhecido]: [editora desconhecida]; 2008.
24. Andrade PG, Brusse GPL, Camargo KCM, Pereira ACR, De Maria PF. Evolução da qualidade da declaração da idade na América Latina e Caribe: uma proposta de escolha de métodos a partir da estrutura etária. In: 7º Congresso da Associação Latino-Americana de População; 20º Encontro Nacional de Estudos Populacionais. Foz do Iguaçu; 2016 out. 17-22; [local desconhecido]: [editora desconhecida]; 2016.
25. Bailey M, Makannah TJ. Patterns of digit preference and avoidance in the age statistics of some recent African censuses: 1970-1986. *J Off Stat*. 1993;9(3):705-15. doi: <http://dx.doi.org/10.1590/1413-812320182311.21832016>
26. Yazdanparast A, Pourhoseingholi MA, Abadi A. Digit preference in Iranian age data. *Ital J Public Health*. 2012;9(1):64-70. doi: <https://doi.org/10.1590/1413-812320182311.21832016>
27. Barua RK. Detection of Digit Preference and Age Misreporting by using Demographic Techniques. Dhaka: East West University; 2015.
28. Dahiru T, Dikko HG. Digit preference in Nigerian censuses data of 1991 and 2006. *Epidemiol Biostat Public Heal*. 2013;10(2):6-10. doi: <https://doi.org/10.2427/8843>
29. Groves RM, Fowler FJ, Couper MP, Lepkowski JM, Singer E, Tourangeau R. *Survey Methodology*. 2nd ed. New Jersey: John Wiley and Sons; 2009.
30. Brestoff JR, Van Den Broeck J. Reporting data quality. In: *Epidemiology: Principles and Practical Guidelines*. Dordrecht: Springer; 2013. p. 557-70.