

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
CENTRO DE CIÊNCIAS DA SAÚDE
FACULDADE DE MEDICINA
INSTITUTO DE ATENÇÃO À SAÚDE SÃO FRANCISCO DE ASSIS
Mestrado Profissional em Atenção Primária à Saúde

FATIMA VALÉRIA LIMA JACQUES

**QUALIDADE DA INFORMAÇÃO SOBRE IDADE NOS REGISTROS DE
ÓBITO NO BRASIL, 1996 – 2015.**

Rio de Janeiro
2018



UNIVERSIDADE FEDERAL
DO RIO DE JANEIRO

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

CENTRO DE CIÊNCIAS DA SAÚDE

FACULDADE DE MEDICINA

INSTITUTO DE ATENÇÃO À SAÚDE SÃO FRANCISCO DE ASSIS

Mestrado Profissional em Atenção Primária à Saúde

FATIMA VALÉRIA LIMA JACQUES

**QUALIDADE DA INFORMAÇÃO SOBRE IDADE NOS REGISTROS DE
ÓBITO NO BRASIL, 1996 – 2015.**

Orientador: Prof. Dr. Raphael Mendonça Guimarães

Dissertação apresentada ao Curso de
Mestrado em Atenção Primária à Saúde
da Universidade Federal do Rio de
Janeiro como requisito para obtenção do
título de Mestre em Saúde Coletiva.

Rio de Janeiro
Abril 2018

1. JACQUES, FATIMA VALÉRIA LIMA

2. QUALIDADE DA INFORMAÇÃO SOBRE A IDADE NOS REGISTROS DE ÓBITO NO BRASIL, 1996-2015. / FATIMA VALÉRIA LIMA JACQUES. - RIO DE JANEIRO: UFRJ / FACULDADE DE MEDICINA, INSTITUTO DE ATENÇÃO À SAÚDE SÃO FRANCISCO DE ASSIS - HESFA, 2018.

3.

V; 36 f.; il.: 31 cm.

Orientador: Raphael Mendonça Guimarães

Dissertação (mestrado) – UFRJ, Faculdade de Medicina, Instituto de Atenção à Saúde São Francisco de Assis – HESFA, Mestrado Profissional de Atenção Primária a Saúde, 2018.

Referências: f. 33-36.

1. Dados Demográficos. 2. Registros de Mortalidade. 3. Saúde Pública. 4. Acurácia dos Dados. 5. Sistemas de Informação em Saúde. 6. Assistência Primária à Saúde - Tese. I. Guimarães, Raphael Mendonça. II. UFRJ, FM, Instituto de Atenção à Saúde, São Francisco de Assis- HESFA. III. Título.



UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
PROGRAMA MESTRADO PROFISSIONAL
ATENÇÃO PRIMÁRIA À SAÚDE



ATA DA DEFESA DE DISSERTAÇÃO PARA A OBTENÇÃO DO GRAU DE MESTRE EM
SAÚDE COLETIVA, ÁREA DE CONCENTRAÇÃO ATENÇÃO PRIMÁRIA À SAÚDE

Às 9 horas do dia 16 de abril de 2018, teve início a Defesa de Dissertação para a obtenção do Grau de Mestre em Saúde Coletiva, área de Concentração: Atenção Primária à Saúde, na Linha de Pesquisa: Evidência Clínica, Modelos Assistenciais e qualidade na APS. A dissertação intitulada "Qualidade da Informação da Variável Idade nos Registros de Óbito no Brasil, 1996-2015". foi apresentada pela candidata Fátima Valéria Lima Jacques, regularmente matriculada no Curso de Mestrado Profissional em Atenção Primária à Saúde da Faculdade de Medicina da UFRJ/HESFA, sob registro nº 117001571 (SIGA). A banca de examinadores foi composta pelos doutores: Raphael Mendonça Guimarães (CPF:053468547-11), Tatiana Rodrigues de Araujo Lima (CPF:092744837-83), Lívia Maria Santiago (CPF:036489566-71), em sessão pública, no prédio do HESFA/UFRJ, no Auditório, sito a Av. Presidente Vargas, 2863 - Cidade Nova, Rio de Janeiro, RJ. Após a exposição em sessão pública, de trinta minutos, a candidata foi arguida. Os membros da banca deliberaram nesta ocasião, considerando a dissertação (X) APROVADA () REPROVADA () APROVADA COM RESTRIÇÃO. A banca emitiu, então, o seguinte parecer:

Após defesa da candidata, a mesma foi arguida satisfatoriamente. Resaltou-se a qualidade do trabalho, recomendando-o para submissão de artigo para publicação.

Na forma regulamentar esta ata foi lavrada e assinada pelos membros da Banca e pela aluna examinada.

Presidente/Orientador: Raphael Mendonça Guimarães

1º Examinadora: Tatiana Lima

2º Examinadora: Livia Maria Santiago

Candidata: Fátima Valéria Lima Jacques

Dedico este trabalho aos meus pais, Conceição e Milton Jacques, pelo desmedido esforço para educar suas filhas e por serem pais tão dedicados.

AGRADECIMENTOS

Agradeço primeiramente a Deus, por me permitir acordar todos os dias com a chance de recomeçar e por todas as oportunidades colocadas no meu caminho.

Ao meu orientador, professor e grande amigo Raphael Mendonça Guimarães, pelo incentivo desde sempre, por ter me oferecido a oportunidade de adquirir conhecimentos que não pensava descobrir sobre a atenção primária e sua importância para sociedade brasileira. Por sua colaboração e contribuição para meu crescimento profissional. Muito obrigada!

Aos meus pais e minhas irmãs, pessoas que fazem todo meu esforço valer à pena, toda tristeza virar alegria, todos os momentos difíceis virar ensinamento. É por eles e para elas todas as minhas vitórias.

Aos meus amigos, pela torcida e aos novos amigos feitos durante esta trajetória, em especial à Gaby e Zé, pela companhia e parceria neste caminho cansativo, trazendo um pouco de leveza em momentos tão difíceis.

As parceiras Alice Castro e Cristina Barcelos, e ao parceiro de todos os momentos Ronaldo Sampaio que me ajudaram muito durante todo o curso no desenvolvimento das minhas atividades laborativas e me deram sempre aquele incentivo e carinho importante em todas as horas. A amiga Natácia Guimarães pela força nos treinos, que faziam liberar todas as endorfinas possíveis para continuar tendo o gás necessário para um novo dia.

A amiga Fabiana Carolino, pelo grande incentivo e por não ter me deixado desistir quando tudo parecia muito difícil, suas palavras foram muito importantes.

A Eliane Mendes, pelo incentivo permanente e por sempre acreditar em mim.

As professoras Tatiana Lima e Livia Santiago, que gentilmente aceitaram o convite para não somente participar da banca, mas de todo o processo para meu aprimoramento profissional. Obrigada!

A todos os professores do Mestrado Profissional em Atenção Primária – UFRJ, que proporcionaram meu encantamento pela atenção primária em saúde.

A Telma Galvão, Chefe de Enfermagem do Instituto de Puericultura e Pediatria Martagão Gesteira, a Rejane Araújo, Coordenadora de Enfermagem do Hospital Universitário Pedro Ernesto, a Marcela Bedran, Gerente do Departamento de Auditoria da Caixa de Assistência da CEDAE, que tiveram um papel importante nesta minha formação.

A todos aqueles que participaram, direta e indiretamente, desta etapa importante da minha vida.

“Vivemos num ritmo célere. Temos uma rotina enfadonha. Não damos um passo sem que não tenhamos que olhar o relógio e avaliar o tempo disponível para realizarmos nossas tarefas diárias, tantas tão repetidas que agimos muitas vezes como robôs. Isso tem tornado o próximo cada vez mais distante de nossa percepção. Diante das máquinas nos confundimos muitas vezes com elas que nos respondem, nos alertam, nos viciam! É através delas que nos comunicamos com os amigos, agora virtuais. É através delas que desejamos bom dia, bom final de semana, feliz aniversário.... Não temos tempo de perceber se nosso colega de trabalho está mais alegre ou triste, se necessita de nosso abraço, de nosso ombro amigo. Quantos grandes amigos ainda restam ao seu redor? Quantas oportunidades aproveitaram ou foram criadas para descobrir um novo amigo? (...) A velocidade tecnológica nos envolveu no seu ritmo. Nossos olhares são velozes, nossos abraços escassos, nosso amor efêmero. Nos entregamos a essa velocidade e já não experimentamos o aconchego do abraço, o afago, a intensidade dos olhos nos olhos.... Contudo, sempre é tempo de acordar, de refazer o caminho e reencontrar o olhar do outro, do nosso sentimento reascender pelo outro e por nós. É sempre tempo de abrir o peito e deixar expandir a essência de cada um. É sempre tempo de amar!”

Antônio Maciel

RESUMO

A falta de qualidade na informação sobre a idade é um fenômeno comum nos países em desenvolvimento. As irregularidades mais comuns são as preferências de dígitos e o aumento da idade em torno de algumas idades atraentes. Ainda, há uma distorção quando a análise é realizada por sexo. Esta avaliação é relativamente comum para dados censitários e pesquisas demográficas. Contudo, é pouco frequente em dados de registro de óbitos, principalmente no Brasil, onde há um fluxo de informações com revisão de registros e, portanto, onde se espera encontrar pouca ou nenhuma inacurácia. Neste sentido, o objetivo geral desta dissertação foi analisar a qualidade da declaração da idade nos registros de óbito no Brasil, segundo sexo, de 1996 a 2015. Como objetivos específicos, tem-se estimar os indicadores de qualidade da declaração da idade nos registros de óbito no Brasil, segundo sexo; descrever a tendência temporal dos indicadores de qualidade da informação da variável idade nos registros de óbito no Brasil, segundo sexo; e, discutir a acurácia dos indicadores de qualidade da declaração da idade nos registros de óbito no Brasil, segundo sexo, de 1996 a 2015. Os dados de idade para os óbitos ocorridos no Brasil entre 1996 e 2015 foram coletados nos microdados do Sistema de Informações sobre Mortalidade, reunidos em um único banco de dados, e a idade foi recodificada, de forma a se obter a idade simples de cada óbito. Em seguida, foi realizada a avaliação dos dígitos terminais. Foram utilizados para a avaliação os índices de Whipple (para atração dos dígitos zero, cinco e a atração em geral, através do Whipple modificado) e o Índice de Myers. Em seguida, a tendência da qualidade dos dados ao longo de 20 anos de observação (1996 a 2015) foi avaliada através de modelos de séries temporais, realizados através de regressão polinomial. A qualidade dos dados de idade foi alta no Brasil 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 na informação foi mais satisfatória entre mulheres, e não houve tendência significativa na melhora, sugerindo estabilidade na avaliação de qualidade ao longo dos 20 anos. Foi encontrada preferência por idade terminando com dígitos '0' principalmente entre homens. A avaliação dos dados entre 1996 e 2015 do Sistema de Informações sobre Mortalidade, utilizando as técnicas demográficas, permite qualificá-lo como adequado. Não foi encontrada evidência significativa de erro na notificação de idade no Brasil. O Sistema de Informações sobre Mortalidade, neste quesito, apresenta, portanto, boa qualidade, com relativa estabilidade ao longo do tempo. Destaca-se a diferença entre sexos na atração por dígito terminal. Espera-se que os resultados sejam úteis para demógrafos e epidemiologistas nas análises que considerem idade como um fator de estudo.

Palavras-Chave: Qualidade de dados. Índice de Whipple. Índice de Myers. Mortalidade. Sistemas de Informação.

ABSTRACT

The lack of quality in information on age is a common phenomenon in developing countries. The most common irregularities are digit preferences and increasing age around some attractive ages. Yet, there is a distortion when the analysis is performed by sex. This evaluation is relatively common for census data and demographic research. However, it is infrequent in death registry data, mainly in Brazil, where there is a flow of information with review of records and, therefore, where it is expected to find little or no inaccuracy. In this sense, the general objective of this dissertation was to analyze the quality of the declaration of the age in the records of death in Brazil, according to sex, from 1996 to 2015. As specific objectives, it is necessary to estimate the quality indicators of the age declaration in the records of death in Brazil, according to sex; to describe the temporal tendency of the indicators of information quality of the variable age in the death registries in Brazil, according to sex; and to discuss the accuracy of the quality indicators of the age declaration in the records of death in Brazil, according to sex, from 1996 to 2015. The data of age for the deaths occurred in Brazil between 1996 and 2015 were collected in the microdata of the Information System on Mortality, gathered in a single database, and the age was recoded, in order to obtain the simple age of each death. Then, the terminal digits were evaluated. Whipple indices (for zero digit attraction, five and general attraction, using the modified Whipple) and the Myers Index were used for the evaluation. Then, the trend of data quality over 20 years of observation (1996 to 2015) was evaluated through time series models, performed through polynomial regression. The quality of the data of age was high in Brazil in the period ($IW_{tot} = 0.55 - 0.83$ (male) and $0.71 - 0.93$ (female); $IM = .388 - .004$ (male) and $.004 - .004$ (female)). The quality of the information was more satisfactory among women, and there was no significant trend in the improvement, suggesting stability in the quality evaluation throughout the 20 years. of the data between 1996 and 2015 of the Mortality Information System, using the demographic techniques, allows to qualify it as adequate, and thus, good quality, with relative stability over time, the difference between sexes in the terminal digit attraction stands out. The results are expected to be useful for demographers and epidemiologists in the analyzes that consider age as a study factor.

Keywords: Data quality. Whipple Index. Myers Index. Mortality. Information systems.

SUMÁRIO

	Página
Dedicatória	i
Agradecimentos	ii
Epígrafe	iii
Resumo	iv
Abstract	v
Sumário	vi
Capítulo 1 Introdução	1
Capítulo 2 Objetivo	4
Capítulo 3 Referencial Teórico	5
Capítulo 4 Metodologia	14
Capítulo 5 Resultados	22
Capítulo 6 Discussão	29
Capítulo 7 Conclusão	32
Referências	33

CAPÍTULO 1

INTRODUÇÃO

Os dados demográficos são representações de fenômenos da realidade. Podem ser considerados a manifestação do modo de ser do homem em coletividade, e por isto têm o mesmo grau de objetividade que a população a que vão refletir (PINTO, 1973). Neste contexto, os dados demográficos e suas fontes serão tão próximos à realidade quanto maior for a sua disponibilidade, sua regularidade (frequência), e sua qualidade (abrangência).

Por sua vez, as fontes de dados demográficos são importantes instrumentos para pesquisa, análise e promoção de políticas públicas, contribuindo para o entendimento de nossa realidade social. Deve-se destacar que o conhecimento das fontes de dados é fundamental para um diagnóstico detalhado e verdadeiro da realidade, o que levará à formulação de programas e políticas públicas específicas e de melhor qualidade. Cabe salientar que estas informações, mais que um conjunto de dados demográficos, taxas e índices presentes em tabelas e banco de dados, são expressões concretas da própria realidade (MAGALHÃES, 2015).

Programas computacionais simples coletam e organizam os dados de forma sistemática, facilitando o acesso, melhorando a captação da informação e sua análise (RISI JÚNIOR, 2006), em concomitância com o desenvolvimento e aprimoramento das tecnologias da informação.

São comuns os estudos que abordam a mortalidade brasileira por meio de estimativas indiretas, apoiadas em severas restrições quanto ao comportamento demográfico das populações e quanto ao uso de pressupostos¹. A Fundação IBGE, por exemplo, descarta as estatísticas vitais e recorre a procedimentos indiretos no cálculo de indicadores de mortalidade) (IBGE, 2010).

A Declaração de Óbito é um instrumento padronizado pelo Ministério da Saúde para o registro das informações de óbito, possibilitando sua utilização como fonte de dados para estudos sobre mortalidade. A qualidade dos registros de óbitos deve ser

¹ Pressupostos são ideias não expressas de maneira explícita num discurso, mas que podem ser percebidas a partir de certas palavras ou expressões que foram utilizadas. Quanto à utilização de pressupostos, eles devem ser sempre verdadeiros ou aceitos como verdadeiros, pois são os responsáveis por construir as informações consideradas explícitas.

constantemente incentivada, com preenchimento completo das declarações e registros de todas as informações solicitadas. Nesse contexto, é importante sensibilizar os profissionais envolvidos no preenchimento ou processamento dos dados sobre o óbito, enfatizando a importância da captação da informação na construção de bancos de dados confiáveis sobre a mortalidade, que servirão de suporte para criação, avaliação e análise de políticas públicas (LIMA E QUEIRÓZ, 2014). Entende-se que informação de qualidade é parte essencial do processo decisório, pois dados fidedignos são muito importantes para a consolidação das informações que servirão como ferramenta para um correto direcionamento das ações dos gestores, e ainda, dados de má qualidade apresentam sérias limitações para gestão dos serviços, uma vez que as bases epidemiológicas têm entre seus objetivos o conhecimento da situação de saúde da população (EWBANK, 1981).

A informação sobre a idade de uma população, sobretudo num contexto de transição demográfica diferenciada, como é o caso do Brasil, está sujeita a erros que, em certa medida, podem comprometer a validade dos indicadores estimados por essas informações, como a razão de dependência² e a taxa de mortalidade³ (PAES, 1999). A qualidade da declaração de idade, de uma forma geral, pode sofrer influências de vários aspectos, muitos intrínsecos ao processo de operacionalização e coleta da informação, e outros vinculados ao respondente (EWBANK, 1981). Estes erros são, muitas vezes, de difícil mensuração e frequentemente resultam no arredondamento da idade declarada, em geral optando-se por idades terminadas em “0” (zero) e “5” (cinco) ou, no caso específico de idosos, na superestimação da idade (COALE & KISKER, 1986; COALE & SHAONIN, 1991; PRESTON e ELO, 1997; YI et al. 1999; ROMERO & FREITEZ, 2008).

Considera-se, portanto, que a qualidade da informação da variável idade é importante para os estudos populacionais e de saúde, pois esta consiste numa variável que está presente na formulação de todos os indicadores necessários para a fundamentação das análises de políticas sociais. Em geral, a confiança na estrutura etária da população é pressuposto teórico de métodos demográficos indiretos (MOULTRIE et al., 2013). Desta forma, não considerar os possíveis erros de informação da variável

² Razão entre o segmento etário da população definido como economicamente dependente (os menores de 15 anos de idade e os de 60 e mais anos de idade) e o segmento etário potencialmente produtivo (entre 15 e 59 anos de idade) na população residente em um determinado espaço geográfico, no ano considerado (JANUZZI, 2014).

³ Índice demográfico que reflete o número de mortes registradas, em média por mil habitantes, em uma determinada região em um período de tempo. A taxa é expressa comumente em unidades de morte por 1.000 pessoas ao ano (JANUZZI, 2014).

idade podem dificultar a aplicação das técnicas demográficas e de análise de dados em saúde e, na pior das hipóteses, tornar inválidos seus resultados.

O interesse em desenvolver o estudo nessa área vem do entendimento da da potencialidade que a informação tem, enquanto ferramenta, oferecendo subsídios para contribuir com os gestores na tomada de decisão e para elaboração de políticas públicas que possam promover mudanças positivas nas condições de vida e saúde da população.

CAPÍTULO 2

OBJETIVOS

Geral

Analisar a qualidade da informação da variável idade nos registros de óbito no Brasil, segundo sexo, de 1996 a 2015.

Específicos

1. Estimar os indicadores de qualidade da informação da variável idade nos registros de óbito no Brasil, segundo sexo, de 1996 a 2015;
2. Descrever a tendência temporal dos indicadores de qualidade da informação da variável idade nos registros de óbito no Brasil, segundo sexo, de 1996 a 2015.
3. Discutir a acurácia dos indicadores de qualidade da informação da variável idade nos registros de óbito no Brasil, segundo unidade da federação, de 1996 a 2015;

CAPÍTULO 3

REFERENCIAL TEÓRICO

1. Demografia, Saúde Pública e Informação e Saúde

A demografia é uma ciência importante para a saúde pública, dentre outras razões por fornecer conceitos e medidas fundamentais sobre a saúde e sua dimensão populacional. Alguns indicadores demográficos são analisados para realização de uma avaliação direta do perfil das condições de saúde, como por exemplo a mortalidade em geral, a mortalidade infantil, etc. Outros são tidos como vinculados a essas condições, como por exemplo a fecundidade (BRITO, 2007).

A estreita relação entre a demografia e a saúde pública surgiu possivelmente em 1662 com o autor John Graunt, que contribuiu imensamente para o conhecimento humano, com estudos sobre a razão de sexos, tábuas de mortalidade e classificação de doenças (NADALIN, 2004). Vale ressaltar que um dos destaques da demografia para a saúde pública, desde os primórdios até os dias atuais, é a sua capacidade de projetar tendências futuras de forma mais correta. Através da demografia, portanto, é possível medir com maior precisão os fenômenos sociais e quantificar o que estes fenômenos representam, estabelecendo prioridades e formulação de melhores políticas para a população.

Os processos demográficos são marcados, neste aspecto, por uma inércia que lhes dá uma estabilidade, uma previsibilidade maior que as outras tendências políticas, sociais ou econômicas. Por isso, os instrumentos demográficos são a base empírica de qualquer planejamento estratégico para médio e/ou longo prazo (EWBAKN, 1981; JANUZZI, 2014).

Dentre as fontes de dados utilizadas pela demografia e pela saúde pública, o Registro Civil é a principal fonte de dados dinâmicos, isto é, de dados provenientes de eventos que alteram a composição e tamanho da população. Tal registro; fornece as chamadas “estatísticas civis”, que possuem algumas características fundamentais: são originadas de informações comunicadas e registradas em cartórios de registro civil, são levantadas e sistematizadas ao longo de todo o ano e, ainda, são divulgadas segundo o local de ocorrência e o local do registro do evento. Possui ainda exigências institucionais,

cujas principais são: registrar os eventos vitais e publicar os informes e levantamentos estatísticos do registro desses dados (HAKKERT,1996).

No Brasil, o Registro Civil é fonte de dados para dois sistemas de informação muito importantes para o estudo da população e a formulação de políticas públicas: o Sistema de Informação sobre Nascidos Vivos (SINASC), cujos dados são originados da Declaração de Nascidos Vivos e o Sistema de Informações sobre Mortalidade (SIM), cujos dados são originados da Declaração de Óbito (ABOUZHR & BOERMA, 2005).

A história sobre o conhecimento de eventos vitais, de maneira sistematizada, no âmbito da saúde, é relativamente recente no Brasil. Em 1975, o Ministério da Saúde (MS) desenvolveu e implantou o Sistema de Informação da Mortalidade (SIM). O SIM processa dados para o volume de óbitos ocorridos para as Unidades da Federação (UF), agregados por tipos de causa morte, cuja série estatística teve início em 1979. Estudos apontam que, mesmo depois de três décadas após sua implantação, ainda persistem problemas com a qualidade dos registros dos óbitos, estes influenciados por fatores que têm relação com aspectos sociais, econômicos e com as condições de vida das populações (PAES, 2005). Contudo, nos últimos vinte anos, os sistemas de informação em saúde no Brasil vêm sendo aprimorados.

2. Sistema de Informações da Mortalidade (SIM)

Um sistema de informação em saúde (SIS) pode ser definido como um conjunto de partes que se integram e articulam para uma finalidade comum (BITTAR et al, 2018). Já a informação é o meio imprescindível para a tomada de decisões, sendo a informação em saúde o apoio para a orientação, implantação, monitoramento e avaliação dos modelos de atenção e da conjuntura de saúde de determinado local. Para se construir uma informação é preciso analisar o dado, ou seja, considerar um número bruto sem tratamento estatístico e transformá-lo em matéria-prima da informação (FNS, 2001).

Conceitualmente, o sistema de informação em saúde pode ser entendido como um instrumento para adquirir, organizar e analisar dados necessários à definição de problemas e riscos para a saúde, avaliar a eficácia, eficiência e influência que os serviços prestados possam ter no estado de saúde da população, além de contribuir para produção de conhecimentos acerca da saúde e dos assuntos a ela

ligados. Portanto, os SIS constituem-se de peças articuladas que têm por objetivos selecionar dados e produzir informação, auxiliando na formulação e avaliação de políticas e no processo de tomada de decisão, visando a melhoria de saúde da população (CAVALCANTI e PINHEIRO, 2011).

A finalidade da informação em saúde consiste em identificar problemas individuais e coletivos do quadro sanitário de uma população, propiciando elementos para análise da situação encontrada e subsidiando a busca de possíveis alternativas de encaminhamento (BRANCO, 1996). As informações em saúde devem, portanto, abranger as relativas ao processo saúde-doença e as de caráter administrativo, todas essenciais ao processo de tomada de decisão dos gestores. Neste sentido, muitas áreas da ciência utilizam informação como elemento necessário para a tomada de decisão. Por conta disso, há um interesse crescente em avaliar a qualidade dessa informação (LIMA et.al., 2000). Alguns atributos de qualidade podem ser mensurados e quantificados. A completude, por exemplo, é um desses atributos e resulta da inclusão de todos os dados necessários para responder a uma questão de determinado problema (CALAZANS, 2008).

Nas últimas décadas, foram implementados no Brasil vários Sistemas de Informação em Saúde, ampliando o uso da informação para a gestão do setor saúde e consolidando uma rede de informações composta por sistemas de racionalidade epidemiológica, de assistência à saúde (produção de serviços), monitoramento de programas de saúde, gerenciamento de serviços, entre outros (COELI et. al., 2009). No limite, os dados contidos neste sistema têm como finalidade primordial a produção de indicadores de saúde que retratem as condições de saúde da população no que diz respeito ao processo saúde-doença e aos aspectos administrativos dos serviços de saúde (CORREA, PADILHA e VASCONCELOS, 2014).

Em 1975, o Ministério da Saúde (MS) desenvolveu e implantou o Sistema de Informação da Mortalidade (SIM), que além de processar dados para o volume de óbitos ocorridos para as UFs, agrega o tipo de causa de morte, cuja série estatística foi iniciada em 1979 (PAES, 2005). Graças a este sistema de informação, o MS divulga em seu sítio na internet estimativas de cobertura de óbitos por unidades da federação, cujos cálculos se baseiam na razão simples entre óbitos observados e estimados. As bases para estas estimativas são muito genéricas e pouco confiáveis.

Elas foram fundamentadas nas projeções de óbitos usadas pelo IBGE nas projeções populacionais, no que tange ao comportamento da mortalidade em cada unidade federativa, apoiadas em pressupostos raramente verificados em populações reais; apesar de tais restrições, essas estimativas servem como referenciais (PAES, 2005). Ressalta-se a importância dos dados com qualidade, que expressem com fidelidade a realidade dos óbitos ocorridos, para formulação de políticas sociais e de saúde que verdadeiramente possam atender às necessidades da população e ainda, que possibilitem a apropriação dessas informações pelos gestores e pela sociedade organizada, para subsidiar ações para benefício dos mesmos.

O amplo e contínuo emprego de um sistema de informação em saúde propicia, no limite, a quebra do círculo vicioso em que usuários consideram o dado inadequado e por isso não o utilizam, o que, por sua vez, contribui para a manutenção desta situação. O SIM apresenta ganhos cada vez mais destacados no que se refere à cobertura e à velocidade de divulgação das informações. Os dados provenientes desse sistema têm subsidiado diversas atividades e publicações relevantes para a saúde pública.

A qualidade das informações disponibilizadas pelo SIM depende de fatores diretamente ligados ao sistema de informação, como a qualidade do programa de computador que lhe dá sustentação, a capacitação dos profissionais de saúde envolvidos, mas também depende de fatores externos ao sistema. Por exemplo, o acesso da população aos serviços de saúde - quando a assistência à saúde é inadequada ou inexistente, dificilmente são coletadas informações fidedignas sobre a causa da morte.

3. Qualidade de Dados

3.1 Qualidade dos Dados de Mortalidade

No Brasil, ainda são poucos os estudos que avaliaram a qualidade dos dados dos SIS, sendo observadas, entretanto, algumas alternativas pontuais. Estudos realizados para avaliação de qualidade de dados dos SIS priorizam as análises sobre confiabilidade, validade, cobertura e completitude, e se referem com maior frequência aos dados de saúde da Região Sudeste (FRIAS et al, 2013), ou seja, são pontuais e concentram-se em determinadas localidades.

A garantia de uma informação de qualidade é condição essencial para a análise objetiva da situação sanitária, sendo de fundamental importância para a tomada de decisões baseadas em evidências e para a programação de ações de saúde. Na área da saúde, existe interesse de muitos países em investir recursos consideráveis em atividades que visem garantir a qualidade dos dados, incluindo capacitação periódica dos profissionais envolvidos com a produção e análise dos dados, além de um monitoramento regular dos dados disponibilizados pelos sistemas.

As estatísticas atuais disponíveis no Brasil oferecem um campo de exploração, cujos dados carregam um poder explicativo suficiente para apontar tendências e padrões de comportamento. Além disso, fornecem informações úteis para se entender como se processa a dinâmica da mortalidade brasileira para vários Estados (PAES, 2005). Neste contexto, a análise e o gerenciamento desses dados foram estratégicos para o fortalecimento dos objetivos e ações dos gestores e o alcance da eficácia das ações de controle social.

O estudo do nível e padrão da mortalidade, assim como a obtenção de estimativas confiáveis são de grande importância na compreensão da dinâmica demográfica e no planejamento de políticas sociais (AGOSTINHO e QUEIRÓZ, 2008). Neste contexto, cabe salientar que o SIM é uma das mais importantes fontes de informação para o planejamento e avaliação das ações de saúde e estes somente podem ser realizados de maneira adequada à saúde pública a partir de informações de boa qualidade sobre qualquer evento relacionado à saúde.

Em vários países do mundo, estimar a mortalidade é um desafio uma vez que a qualidade da informação geralmente não é satisfatória e limitações nos dados de mortalidade e população têm persistido ao longo do tempo (LUY, 2012). Contudo, tem-se investido em esforços para preparo de pessoal para preenchimento dos dados de registro e estes vem trazendo uma melhoria das informações de saúde e de mortalidade nas últimas décadas. No Brasil, por exemplo, o grau de cobertura dos registros de óbitos de homens adultos passou de 83,2% no período de 1980-1991 para 89,7% no período 2000-2010 (LIMA e QUEIRÓZ, 2014). Neste contexto, a garantia de uma informação de qualidade é condição fundamental para análise da situação de saúde, fundamentando a tomada de decisões baseada em dados confiáveis e oferecendo subsídios para o planejamento adequado da programação em saúde.

Apesar do avanço na qualidade de informação, existem ainda desafios a serem perseguidos, como o avanço na avaliação da qualidade dos dados e a correção da subenumeração de óbitos, para que as estimativas de mortalidade sejam mais confiáveis. Neste contexto, é importante salientar que estimativas inadequadas de mortalidade afetam o modo como a evolução e a tendência da morbimortalidade acontecem.

3.2 Qualidade da informação da variável idade

Em 1955, a ONU lançou um manual (*Methods of appraisal of quality of basic data dor population estimates*), com a finalidade de propôr métodos para avaliação da qualidade dos dados vinculados a estimativas populacionais. Foi a primeira publicação com visibilidade internacional a organizar estes métodos, sendo importante até os dias atuais (ONU, 1955).

No Brasil, porém, a evolução dos registros vitais de maneira sistematizada é relativamente recente. Em 1974, o IBGE, iniciou uma série de publicações anuais das estatísticas vitais para todas as UFs.

A confiança na estrutura etária da população em estudo é um pressuposto de diversos métodos e técnicas, diretas e indiretas, relacionadas a componentes da dinâmica demográfica, tais como: fecundidade, mortalidade e migração (MOULTRIE et al,2013). Porém, adverte-se que nem sempre a distribuição por idade simples possui boa qualidade (ONU, 1995). Por isso, a qualidade da declaração da idade é uma preocupação constante dos demógrafos.

Em geral, os erros na declaração da idade ocorrem por dois motivos: 1) a preferência por dígitos; 2) o erro/viés de memória do respondente. Além destes, em pesquisas domiciliares, é comum um indivíduo, além de prestar informações sobre si próprio, fazer o mesmo a respeito de outros indivíduos, não presentes no domicílio no momento da pesquisa (sendo a este dado o nome de informante *proxy*). Neste caso, o erro de memória pode estar presente quando o respondente, além de informações de si mesmo, ainda responde por outras pessoas do domicílio (ALVES et al., 2016).

Vários autores recomendam a avaliação de possíveis erros nos dados antes da geração de indicadores demográficos e socioeconômicos (PAES, 2005). A ONU sugere

algumas técnicas para avaliar a qualidade da declaração da idade. Neste contexto, o método mais simples e eficaz é a realização de uma inspeção visual da pirâmide etária por idade simples, verificando as distorções em torno de dígitos específicos (como o “0” e o “5”). Neste cenário, destaca-se que existem ainda técnicas mais sofisticadas, que ajudam a identificar e comparar erros, como a atração por dígitos específicos. Dentre elas, destaco: o Índice de Myers e o Índice de Whipple, que serão abordados neste estudo.

Os índices para análise de erros na declaração da idade foram construídos para fornecer uma medida mais objetiva, ajudando a complementar a inspeção visual realizada a partir dos gráficos, favorecendo a identificação dos dígitos pelos quais ocorre a atração, caso ela exista. De forma geral, todos os índices a serem apresentados alcançam este objetivo, variando somente em capacidade diagnóstica, uma vez que a variação considera os grupos etários utilizados no cálculo e os dígitos finais analisados (ALVES et al., 2016).

No quadro a seguir, estão colocados os indicadores para avaliação de declaração da idade e suas características importantes (ALVES et al., 2016).

Quadro 1: Indicadores de avaliação da qualidade de dados do quesito idade.

INDICADOR	INFORMAÇÃO UTILIZADA	MEMÓRIA DE CÁLCULO	INTERPRETAÇÃO
Índice de Whipple	Necessário ter a população por idade simples. A partir desta, obtém-se a soma da população com idades múltiplas de 5 e da população entre 23 e 62 anos.	$IW_5 = \frac{P_{25} + P_{35} + \dots + P_{55}}{\frac{1}{10} * (P_{23} + P_{24} + \dots + P_{62})} \times 100$	Quanto maior o resultado, maior a violação do pressuposto de distribuição uniforme dos dígitos finais
Índice Whipple-type index	Necessário ter a população por idade simples. A partir desta, obtém-se a soma da população com idades múltiplas de 5 e da população entre 23 e 62 anos.	$IW_5 = \frac{P_{25} + P_{35} + \dots + P_{55}}{\frac{1}{10} * (P_{23} + P_{24} + \dots + P_{62})} \times 100$ $IW_{m1} = \frac{5 * (P_{31} + P_{41} + \dots + P_{61})}{({}_5P_{29} + {}_5P_{39} + \dots + {}_5P_{59})} \times 100$	Quanto maior o resultado, maior a violação do pressuposto de distribuição uniforme dos dígitos finais
Índice Total Modified Whipple's Index	Necessário ter a população por idade simples. A partir desta, obtém-se a soma da população com idades múltiplas de 5 e da população entre 23 e 62 anos.	$W_{tot} = \sum_{i=0}^9 IW_{m_i} - 1 $	Quanto maior o resultado, maior a violação do pressuposto de distribuição uniforme dos dígitos finais
Índice de Myers	Propõe a soma da população que possui o mesmo dígito final na idade declarada para o grupo de 10 a 89 anos (G1) e para o grupo de 20 a 99 anos (G2). A população acima de 100 anos pode ser dada como igual a zero, pois não afetam significativamente os resultados.	$IM_i = 100 \times f_i - 10 $ $IM = \sum_{i=0}^9 IM_i, \text{ para } 0 \leq IM_i \leq 180$	Este índice varia de 0 a 180. Quanto menor for a preferência digital, mais próximo de 0, o índice foi indicativo de melhor qualidade da informação

É importante salientar, ao firmar, que PAES e ALBUQUERQUE (1999) recomendaram a avaliação de possíveis erros nos dados antes da geração de indicadores demográficos e socioeconômicos. Algumas técnicas mais sofisticadas permitem a identificação e comparação de tais erros, no caso da atração por dígitos específicos. Dentre elas, podemos destacar: o Índice de Myers e o índice de Whipple. Ambos assumem que não há atração por dígitos finais (ou seja, a distribuição seria uniforme) e, a partir disso, se propõem a identificar, por técnicas distintas, em para qual (ou quais) dígitos existe a violação deste pressuposto.

CAPÍTULO 4

METODOLOGIA

Desenho de estudo

Trata-se de estudo ecológico, que avaliou a qualidade da declaração da variável idade nos dados de mortalidade.

Fonte de dados

Foram utilizados os dados referentes aos óbitos no Brasil no período de 1996 a 2015. Os dados foram extraídos do Sistema de Informações sobre Mortalidade, disponibilizados pelo Ministério da Saúde (disponível em: <http://www2.datasus.gov.br/DATASUS/index.php?area=0901&item=1&acao=26&pad=31655>). Já os denominadores foram estimados através das projeções populacionais intercensitárias do Instituto Brasileiro de Geografia e Estatística (IBGE), calculadas através de interpolação linear. Foram obtidos os microdados por unidade da federação (UF), que foram em seguida agregados para compor o banco de dados agregado para o Brasil. A variável idade foi recodificada, de forma a se obter seu valor como variável contínua em anos.

Análise de Dados

a) Etapa 1

Primeiramente, foi realizada uma inspeção visual da pirâmide etária por idade simples, verificando distorções em torno de dígitos específicos (como “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.

b) Etapa 2

Outras técnicas mais sofisticadas ajudam a identificar e comparar erros, como atração por dígitos específicos. As principais técnicas utilizadas assumem que não existe atração por dígitos finais (ou seja, a distribuição seria uniforme) e, a partir disso, se propõem a identificar em qual ou quais dígitos existe a violação deste pressuposto. Dentre elas, podemos destacar: o Índice de Whipple e o índice de

Myers (ALVES et al, 2016).

i. *Índice de Whipple*

O índice proposto por Whipple procura por sinais de atração por dígitos na declaração de idade. Inicialmente, o índice foi proposto para análise dos dígitos 0 e 5, que usualmente são apontados como os mais comuns quando há preferência por dígitos, geralmente associada à inacurácia na declaração da idade. Para o cálculo, considera-se o intervalo entre 23 e 62 anos. Este intervalo, criado arbitrariamente, procura minimizar o erro nas idades extremas, que em geral são mais comuns. Supondo uma distribuição normal para a idade na população, acredita-se que os extremos estariam em áreas críticas da distribuição, e por isso poderiam apresentar maior variabilidade aleatória (portanto, menos associadas à erros sistemáticos, como é o caso da atração por dígitos) (ANDRADE et al, 2016).

É necessário ter, para a aplicação do método, a idade simples. Ademais, supõe-se, da mesma forma que no índice de Myers, que há uma distribuição uniforme da população por dígito terminal. Desta forma, para o índice de Whipple no dígito terminal 5, por exemplo, considera-se que a soma das populações das idades cujo dígito terminal é 5 (25, 35, ... 55) corresponde a 10% da população total. Desta forma, tem-se que:

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

Em situações extremas, em que não há uma única idade com dígito terminal 5, o índice tem valor zero. Ao contrário, quando todas as idades têm dígito terminal 5, o índice tem valor 1000. Quanto maior o valor, maior a violação do pressuposto de distribuição uniforme dos dígitos terminais. Considera-se, para classificação:

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: Formiga, Ramos e Monteiro (2000)

Usualmente, realiza-se o cálculo para o conjunto de idades que terminam em 0 ou 5, e desta forma o cálculo é:

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

Cabe ressaltar que, neste caso, assumindo dois dígitos preferenciais, o valor máximo alcançado pelo índice é de 500.

Posteriormente à criação do índice, Noubissi (1992) sugeriu um índice de Whipple modificado, que permite a análise de atração global para todos os dígitos. 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. Neste sentido, tem-se que:

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

Neste caso, quando todas as idades convergem para o mesmo dígito, tem-se que $IW_{m_1} = 5$. Ao contrário, quando nenhuma idade possui o dígito terminal analisado, o valor do índice é zero. Finalmente, o índice total, modificado, apresenta-se da seguinte forma:

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

Neste caso, considera-se que, quanto mais próximo de zero, melhor é a qualidade da informação. Este índice total serve como uma medida agregada que, de forma similar ao índice de Myers, fornece um resultado global a respeito da preferência por dígitos. Em oposição à distribuição uniforme, havendo total preferência por um determinado dígito único, o valor final do índice é 13; quando há preferência por dois dígitos terminais distintos, o valor é 16. Este seria o valor máximo do índice.

ii. Índice de Myers

O Índice de Myers é uma medida utilizada para verificar e quantificar erros na declaração da idade (SIEGEL e SWANSON, 2004). Através deste índice, é possível verificar

a atração por dígitos, o que implica na qualidade da informação. Para a implementação deste método, a ONU (1955) propôs a soma da população que possui o mesmo dígito final na idade declarada para o grupo de 10 a 89 anos (G1) e para o grupo de 20 a 99 anos (G2). Neste método, a população acima de 100 anos não é contabilizada, pois supõe-se que ela não afeta significativamente os resultados (ALVES et al, 2016). 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 abaixo:

$$G_1(i) = (i + 1) * \sum_{\alpha=10}^{89} P_i, \text{ para todo } i \in \{0, 1, \dots, 9\}$$

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

Considera-se que a frequência (f_i) que cada dígito i tem no total é dada pela soma de G1 e G2. Supõe-se, ainda, que a proporção esperada de cada dígito equivale a 10% do total (suposição da distribuição uniforme). 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 a fórmula:

$$IM_i = |100 \times f_i - 10|$$

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

O Índice de Myers varia de 0 a 180, e quanto menor for a preferência digital, mais próximo de zero foi o valor do índice e, portanto, melhor a qualidade de apresentação da variável. Considera-se, para efeito de classificação:

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: Formiga, Ramos e Monteiro (2000)

Etapa 3

Trata-se de análise de séries temporais. Por princípio, admite-se que os modelos utilizados para descrever séries temporais são processos estocásticos, isto é, processos controlados por leis probabilísticas (LATORRE e CARDOSO, 2001).

Inicialmente, as séries históricas foram escritas da seguinte forma:

$$Z_t = f(t) + a_t, t=1,2,3,\dots, N,$$

onde $f(t)$ é chamada de sinal, e a_t é chamado de ruído. Neste modelo, o sinal $f(t)$ é uma função do tempo determinada, e a_t é uma variável aleatória, independente de $f(t)$. Os pressupostos do modelo são que as variáveis aleatórias a_t não sejam correlacionadas, e tenham como parâmetros a média zero e a variância constante, ou seja,

$$E(a_t) = 0, \forall;$$

$$E(a_t^2) = \sigma_a^2, \forall t$$

$$E(a_t a_s) = 0, s \neq t$$

Supõe-se inicialmente que a componente sazonal não esteja presente, por não se tratar de evento relacionado a nenhuma característica ambiental, como é o caso de algumas doenças tropicais. Considera-se, então, o seguinte modelo

$$Z_t = T_t + a_t,$$

onde a_t é uma variável aleatória com média zero e variância a .

Deve-se verificar, primeiramente, se a série não se trata de uma série estacionária, ou seja, sem tendência. Para isso, foi aplicado o teste de Wald-Wolfowitz.

Considerou-se, para isto, as seguintes etapas:

Seja $\{Z_t, t = 1, \dots, N\}$ uma série temporal com N observações. Foi considerada M como a mediana das N observações de Z_t . Foi atribuído o símbolo “ p ” para os valores maiores ou iguais à mediana, e “ q ” menores que a mediana, para cada elemento Z . Desta forma, $N = N_a + N_b$. Ainda, ao longo da série haverá grupos de observações marcados por “ p ” ou “ q ”. Considere ainda “ T ” como o número de grupos com símbolos iguais.

Para a avaliação, foram formuladas as hipóteses:

$$H_0 = \text{"Não Existe tendência"}$$

$$H_1 = \text{"Existe tendência"}$$

A estatística de teste é mensurada a partir do número total de grupos. Assim, rejeita-se a hipótese nula (H_0) se houver um número pequeno de grupos com

símbolos iguais, ou seja, se “T” for pequeno. Para valores de N_a ou N_b superiores a 20, pode-se aproximar a distribuição de T por uma normal, ou seja,

$$T \sim N(\mu, \sigma^2), \text{ onde}$$

$$\mu = \frac{2N_a N_b}{N} + 1 \quad \text{e} \quad \sigma = \sqrt{\frac{2N_a N_b (2N_a N_b - N)}{N^2 (N - 1)}}$$

Para testar a hipótese, foi considerado o nível de significância de 95%.

Em seguida, com o objetivo de se obter uma primeira ideia da evolução dos valores dos indicadores das séries, avaliando o efeito da tendência, foi realizado o teste de Cox-Stuart. Este método, baseado na distribuição binomial, baseia-se em agrupar as observações por pares. A cada par (X_i, X_{i+c}) associa-se o sinal “+”, se $X_i < X_{i+c}$, e o sinal “-”, se $X_i > X_{i+c}$, eliminando os empates, para $c = N/2$, em que N é o número de observações da série e X_i é a observação ($i = 1, \dots, N$).

O teste é usado para verificar a seguinte hipótese:

$$H_0: P(X_i < X_{i+c}) = P(X_i > X_{i+c}) \quad \forall i$$

$$H_1: P(X_i < X_{i+c}) \neq P(X_i > X_{i+c}) \quad \forall i$$

As hipóteses testam a existência de tendência (H1) ou não (H0). Se a probabilidade de sinais “+” for igual à probabilidade de sinais “-”, não existe tendência. Para isso, considera-se que a distribuição é dada por

$$T \sim \text{Bin}(n, p), \text{ sendo } n=N \text{ e } p=0,5.$$

Finalmente, as tendências foram analisadas pelo modelo polinomial, cuja variável dependente (Y) é representada pelos indicadores de avaliação da qualidade dos dados, e a variável independente (X) é representada pelos anos-calendário. Para a análise das tendências, foram construídos diagramas de dispersão entre os indicadores e os anos de estudo, para visualizar a função que poderia expressar a relação entre eles e escolher a ordem do polinômio. Os seguintes modelos de regressão foram testados:

a) Linear (1ª ordem): $Y = \beta_0 + \beta_1 X$

b) 2ª ordem: $Y = \beta_0 + \beta_1 X + \beta_2 X^2$

c) 3ª ordem: $Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3$

A escolha do melhor modelo dependeu da análise do diagrama de dispersão, do valor do coeficiente de determinação (R^2) e da análise dos resíduos (especialmente para

ratificação da suposição de homocedasticidade). Quando dois modelos eram semelhantes, do ponto de vista estatístico, optou-se pelo modelo mais simples, ou seja, de menor ordem, ou seja, escolheu-se o modelo mais parcimonioso. E considerou-se tendência significativa aquela cujo modelo estimado apresentou valor de p inferior a 0,05. Foi realizada a análise de resíduos e, para todos os modelos, foi avaliada a suposição de homocedasticidade e aderência à distribuição normal.

A fim de se evitar a multicolinearidade entre os termos da equação de regressão, optou-se por não trabalhar com os valores de X e sim com o ponto médio da série histórica. Desta forma, para o período entre 1996 e 2015, o termo $(X - \text{ano})$ representa a variável ano-centralizada, de acordo com a seguinte formalização do modelo:

$$Y_T = \beta_0 + \beta_1 (X - 2006),$$

Onde: Y_T = Índice de Qualidade (Myers ou Whipple); X = Ano; β_0 = Coeficiente Médio do período; β_1 = Incremento Médio anual

Com a obtenção dos índices de Whipple e Myers por ano (Software utilizado para o estudo: R versão 3.5.1), no período de 1996 a 2015, foi realizada uma análise de séries temporais, para avaliar a evolução temporal da qualidade da informação da idade nos registros de óbito. Inicialmente, foram verificadas a estacionariedade da série temporal, através do teste de Wald-Wolowitz, e o efeito da tendência, através do teste de Cox-Stuart. Em seguida, após verificar estes efeitos, a tendência de mortalidade foi analisada utilizando-se a técnica de regressão polinomial (LATORRE, 2001), 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, foi escolhido o modelo mais simples, pelo princípio da parcimônia (LATORRE, 2001).

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, fica o projeto isento de submissão ao Comitê de Ética em Pesquisa.

CAPÍTULO 5

RESULTADOS

No período de 20 anos, o Brasil apresentou pouco mais de 21 milhões de óbitos, distribuídos entre as diversas faixas etárias. Destaca-se a evidente mudança na contribuição dos grupos: o grupo de 0 a 9 anos teve uma mudança na mortalidade proporcional, entre 1996 e 2015, de 10,31% para 3,67%. Já o grupo de 80 anos e mais passou de 17,88 para 28,53%. A figura 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 reduzem 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, sem este padrão em idades jovens. Especula-se que esta diferença 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.

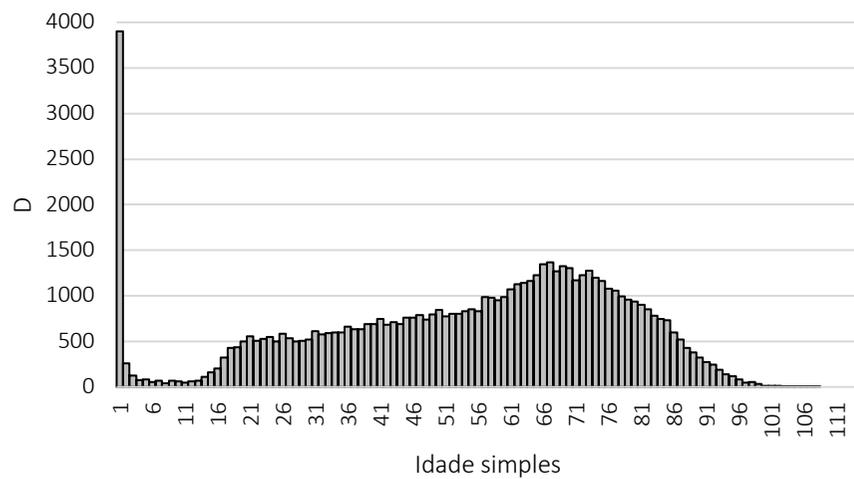
É possível dizer que há uma certa estabilidade na qualidade da informação sobre a variável idade. Isto se reflete nos dois índices avaliados (Whipple e Myers). Ao avaliar o desvio de dígitos, através do índice de Myers (figura 2), pode-se perceber que, desde 1996, a informação já apresentava boa qualidade, apresentando moderada variabilidade, de acordo com o dígito de preferência, ainda que, considerando a unidade de medida (%), os valores de desvio sejam pequenos. Isto sugere o fato de esta diferença ser devida à 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 pelo sexo masculino.

A figura 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 qualidade dos dados 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 tende a reduzir quando analisado o dígito terminal zero, evidenciando alguma atração. De fato, como se observa na figura 4, que

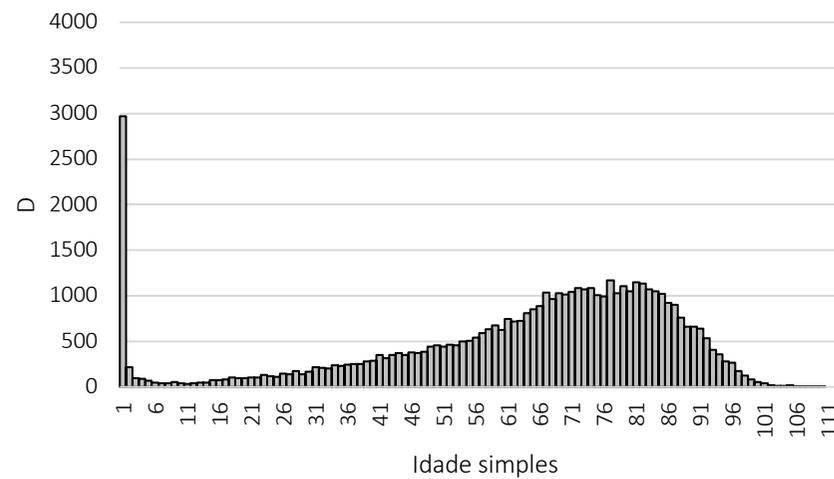
fornece os índices totais modificados específicos por dígito para cada um dos sexos, nos anos de 1996 e 2015, vê-se alguma diferença para os dígitos zero, dois e três, para ambos os sexos, independente do ano de observação.

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

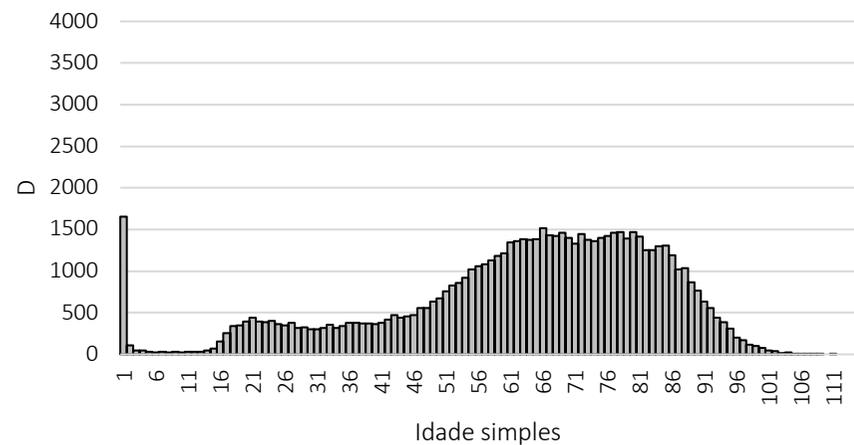
Figura 1: Estrutura etária dos óbitos no Brasil segundo sexo, 1996 e 2015.



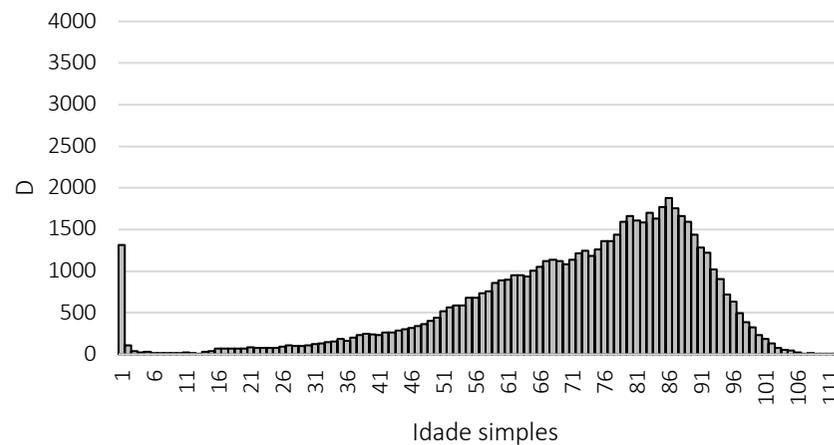
1996 - Masculino



1996 - Feminino



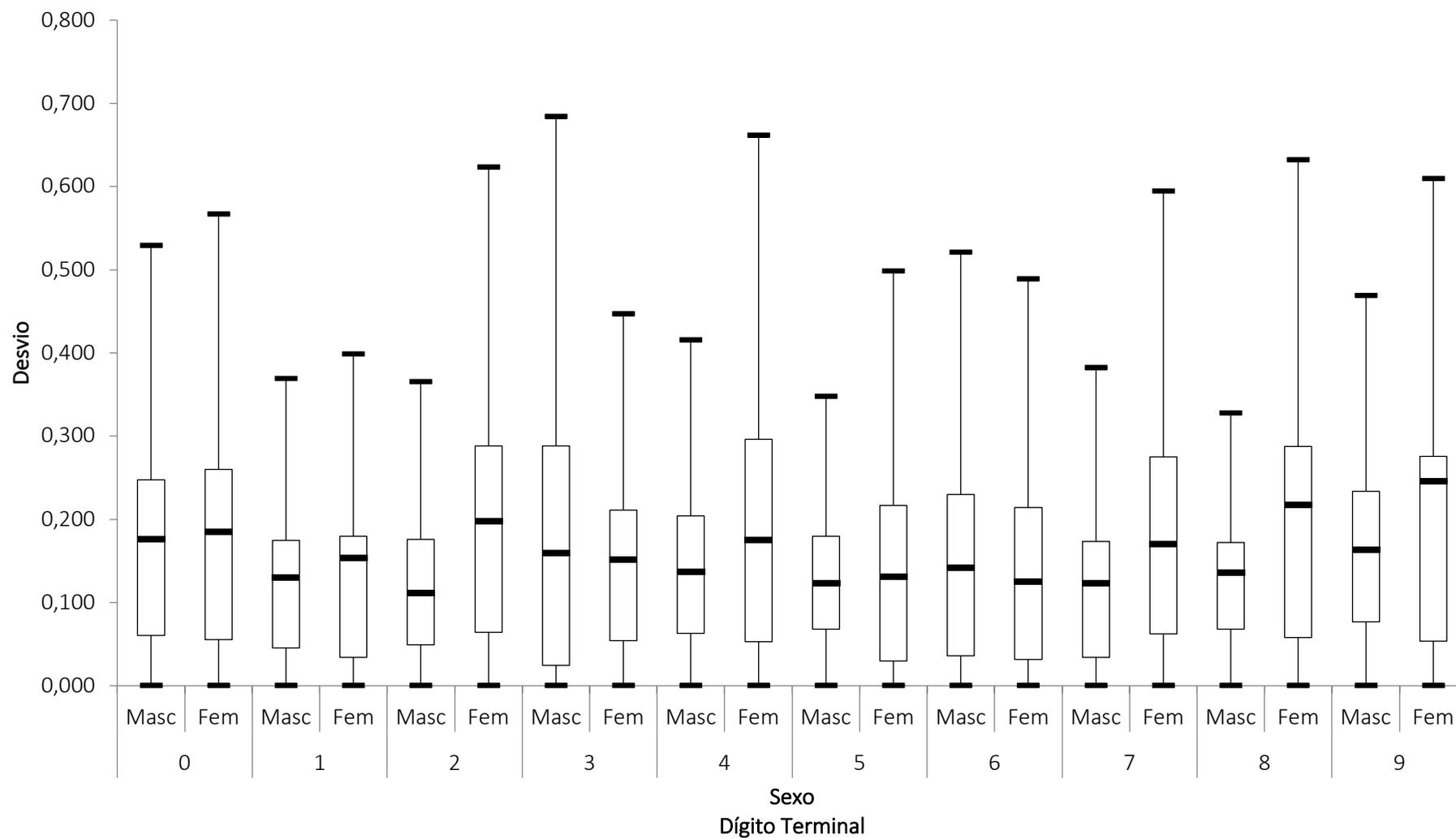
2015 - Masculino



2015 - Feminino

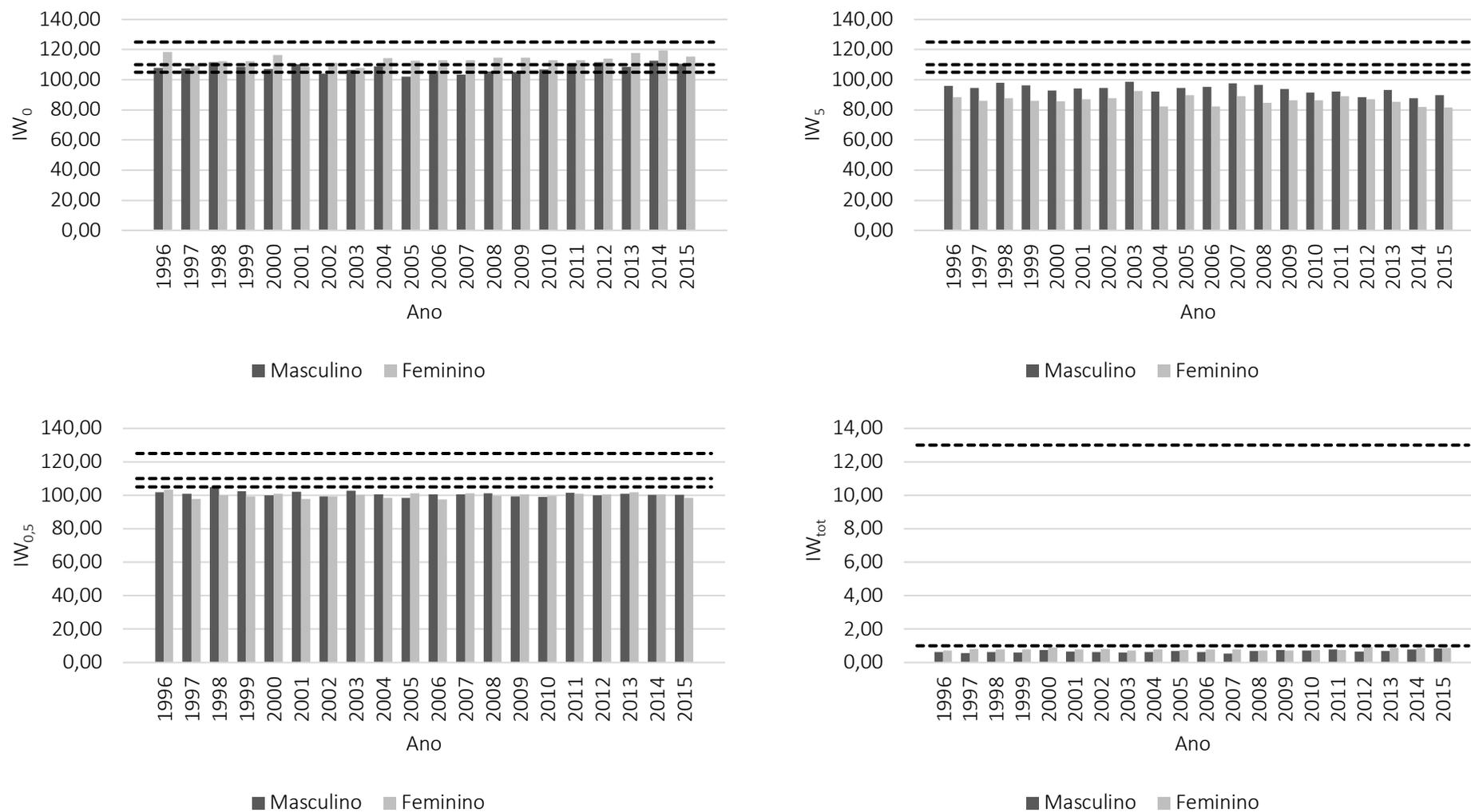
Fonte: DATASUS, 2018.

Figura 2: Descrição da atração por dígito terminal pelo Índice de Myers, segundo sexo e dígito terminal. Brasil, 1996 – 2015.



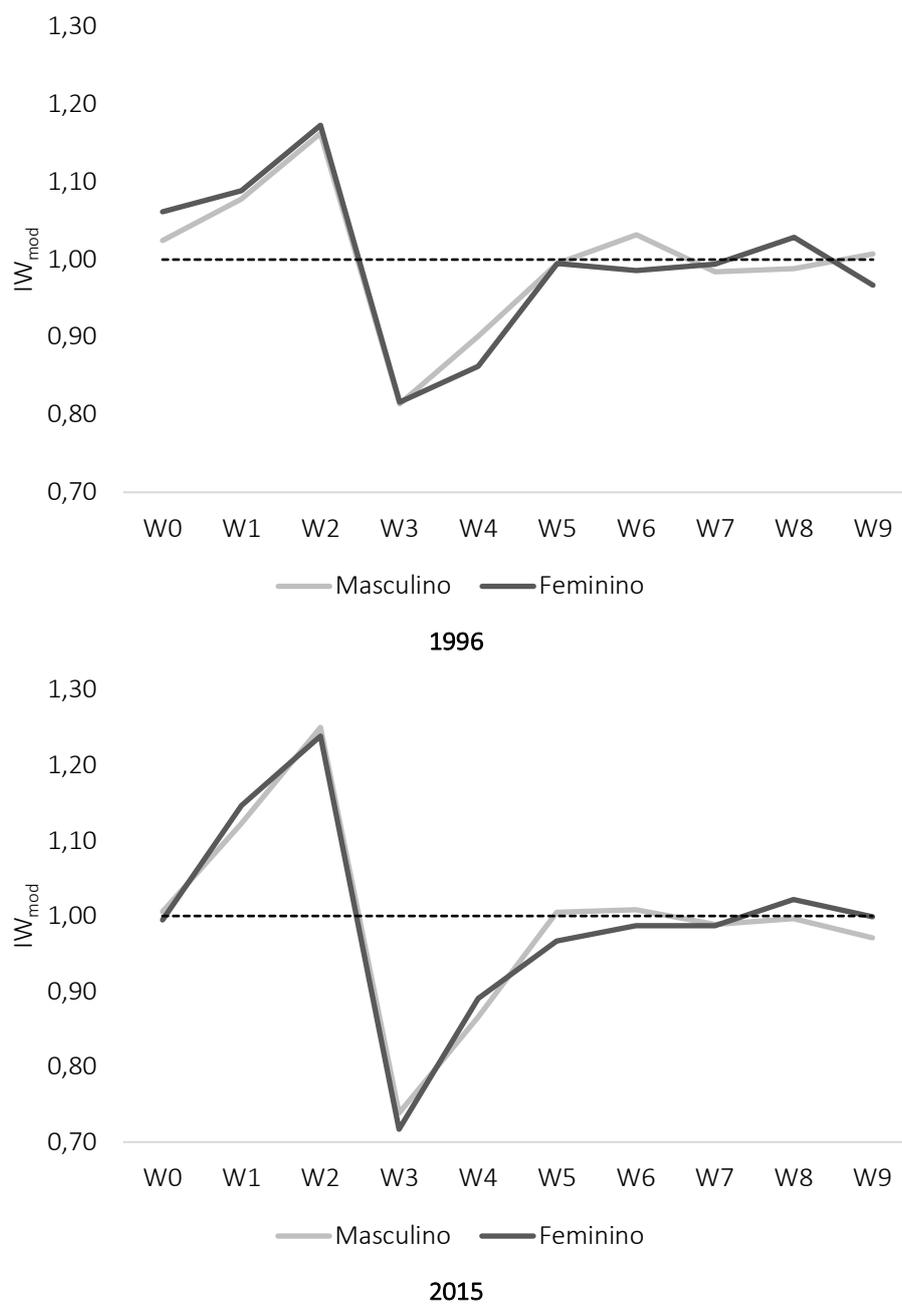
Fonte: DATASUS, 2018.

Figura 3: Indicador de atração por dígito terminal em dados de óbitos no Brasil com o Índice de Whipple, 1996 - 2015



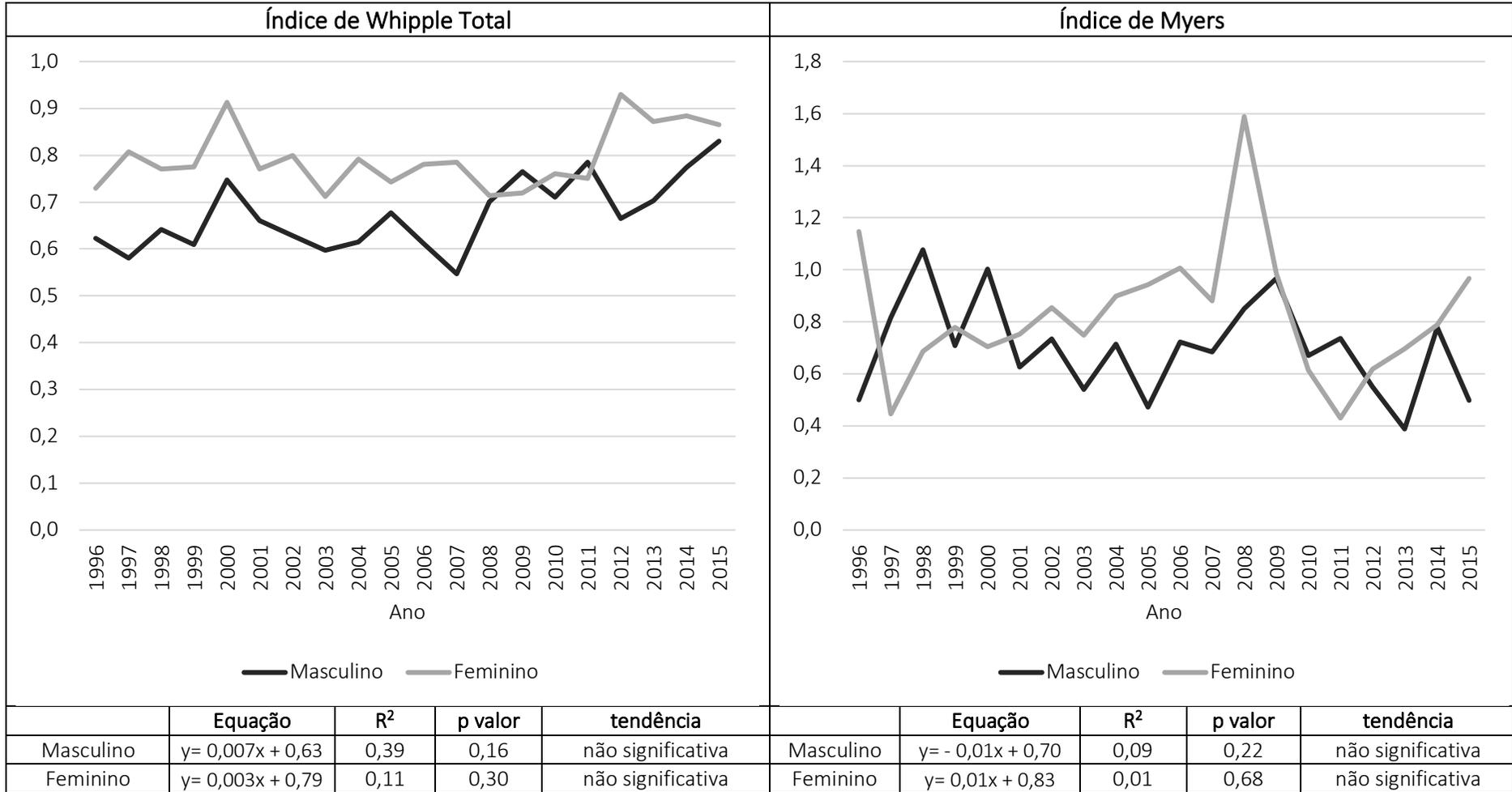
Fonte: DATASUS, 2018.

Figura 4: Descrição da atração por dígito terminal pelo Índice de Whipple segundo sexo e dígito terminal. Brasil, 1996 – 2015.



Fonte: DATASUS, 2018.

Figura 5: Tendência temporal dos indicadores de atração de dígito terminal no Brasil, 1996 – 2015.



Fonte: DATASUS, 2018.

CAPÍTULO 6

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 correção¹⁷.

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 pelos dígitos terminais "0" e "5", indicando uma preferência dos respondentes pelo relato de idades. Desta forma, a distribuição por idade sugere um nível de confiabilidade aceitável, sem a necessidade de ajustamentos ou suavizações. Portanto, tais os dados de mortalidade permitem a construção de tábuas de vida, importante instrumento de avaliação da mortalidade e esperança de vida, para a demografia. Estudo anterior¹⁸ 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^{19,20}, e não por declaração verbal dos familiares, uma prática comum em pesquisas domiciliares^{21,22}.

Randall e Coast²³ 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 do Sul da África. 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 Freitez²⁴ mostraram a

melhoria na qualidade da declaração de idade, ao comparar os Censos da década de 1990 com os Censos da década de 2000. Andrade et al²⁵, também analisando Censos da América Latina, constataram que em um contexto de estrutura etária rejuvenescida, independe a utilização do Índice de Whipple ou 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.

Estudos que utilizaram os IW e IM demonstraram que a preferência digital ocorreu mais frequentemente para o sexo feminino²⁶⁻²⁸, e pelos dígitos “0” e a preferência pelo “5”^{28,29}.

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 idade, em estudos, pode levar a inferências erradas. Embora haja rotina cuidadosa para obtenção dos dados, alguns problemas podem surgir; por exemplo, têm-se os chamados “erros sistemáticos”³⁰. Sabe-se que a aproximação da consciência etária se manifesta no fenômeno da acumulação de idade, para dados de idade autorreferidos ou por procuração^{20,31}. 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 avaliação incorreta das taxas demográficas, e por conseguinte interferir no planejamento de intervenções efetivas³⁰.

A precisão dos dados de idade em registros de saúde, finalmente, 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 garantia da precisão dos dados de idade.

A avaliação da qualidade dos dados de mortalidade é importante, pois para diversas situações de saúde, envolvendo uma série de agravos, os dados de óbito são a única informação disponível. Os resultados apontam para a boa qualidade dos dados, ainda que haja uma diferença entre sexos. Isto 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, por exemplo). Uma vez que o uso de ferramentas demográficas parece adequado para sistemas de informações em saúde. Adicionalmente, aponta-se a necessidade de aplicação em outras bases que não apenas a de mortalidade, bem como as bases de

informações sobre nascimentos, procedimentos ambulatoriais e internações hospitalares.

CAPÍTULO 7

CONSIDERAÇÕES FINAIS

O presente estudo teve como objetivo analisar a qualidade da informação da variável idade nos registros de óbito no Brasil, segundo sexo, de 1996 a 2015. A partir dos dados obtidos e analisados, pode-se perceber que a informação sobre a idade – que pode ser considerada uma proxy da qualidade do registro como um todo – é adequada, ressaltando o sucesso na rotina de controle de qualidade da informação.

A avaliação da qualidade dos dados de mortalidade é importante, pois para diversas situações de saúde, os dados de óbito são a única informação disponível para uma série de agravos. Neste sentido, o estudo possui como vantagem a utilização de uma técnica bastante utilizada em pesquisas populacionais, como é o caso do Censo Demográfico e a Pesquisa Nacional por Amostras de Domicílio (PNAD).

Ainda, o presente estudo aponta a diferença sistemática na qualidade da informação quando comparado o sexo. Neste sentido, há a necessidade de avaliação da qualidade em categorias específicas, como a causa do óbito, ou de acordo com o nível de desenvolvimento social, que se vê refletido, por exemplo, na divisão em Unidades Federativas. Desta forma, pode-se discriminar situações de vulnerabilidade que tornem clara a dificuldade na obtenção de dados com qualidade.

Finalmente, tais resultados apontam para um bom desempenho na opção do uso de técnicas demográficas para avaliar fontes de dados que são geridas pelo Ministério da Saúde, de forma que é interessante a coleta de dados e informações relacionados a outros Sistemas de Informação, como o e-SUS e os registros de pagamento, como o Sistema de Informações Ambulatoriais e de Internação Hospitalar.

REFERÊNCIAS

- AbouZahr C, Boerma T. Health information systems: the foundations of public health. *Bulletin of the World Health Organization*. 2005; 83(.8): 578-583. Agostinho CS, Queiróz BL. Estimativas da mortalidade adulta para o Brasil no período 1980/2000: uma abordagem metodológica comparativa. Disponível em <http://www.abep.org.br/publicacoes/index.php/anais/article/viewFile/1871/1830>. Acesso em 04 abri 2018.
- Agostinho CS, Queiroz BL. Estimativas da mortalidade adulta para o Brasil no período 1980/2000: uma abordagem metodológica comparativa. ABEP 2008
- Alves LA, Andrade PG, De Maria PF, Pereira ACR, Marins RL, Brusse GPL, Camargo KCM. Uma proposta de utilização do software R para a construção de algoritmos de avaliação da qualidade da declaração da idade. *Textos NEPO n. 73*, - Campinas, SP: Núcleo de Estudos de População "Elza Berquó" / Unicamp, 2016.
- Anderson BA, Silver BD. Issues of data quality in assessing mortality trends and levels in the New Independent States. In Bobadilla JL, Costello CA, Mitchell F.(org.), *Premature death in the new independent states – Washington, DC: National Academy Press*. (pp. 120-154).
- Andrade PG, Brusse GPL, Camargo KCM, Pereira ACR, Martins RL, 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." Disponível em <http://www.abep.org.br/publicacoes/index.php/anais/article/view/2707/2610>. Acesso em 04 abr 2018.
- Barbuscia DM, Rodrigues-Júnior AL. Completude da informação nas Declarações de Nascido Vivo e nas Declarações de Óbito, neonatal precoce e fetal, da região de Ribeirão Preto, São Paulo, Brasil, 2000-2007. *Cadernos de Saúde Pública*. 2011; 27(6): 1192-1200
- Bittar OJN, Biczuk M, Serinolli MI, Novaretti MCZ, Moura MMN. Sistemas de informação em saúde e sua complexidade. *Rev. Adm. Saúde*. 2018; 18(70): <http://dx.doi.org/10.23973/ras.70.77>
- Branco MAF. Sistemas de informação em saúde no nível local. *Cadernos de Saúde Pública*. 1996; 12(2): 267-270.

- Brito F. A transição demográfica no Brasil: as possibilidades e o desafio para a economia e a sociedade. Belo Horizonte: CEDEPLAR/UFMG; 2007.
- Calazans ATS. Qualidade da informação: conceitos e aplicações. *Transinformação*. 2008; 20(1): 29-45.
- Cavalcante RP, Pinheiro MM. Política nacional de informação e informática em saúde: avanços e limites atuais *Perspectivas em Gestão & Conhecimento*, João Pessoa 2011;1(2):106-119.
- Coale AJ, Kisker E. Mortality Crossovers: Reality or Bad Data?" *Population Studies*. 1986; 40:389-401.
- Coale AJ, Shaomin L. The effect of age misreporting in China on the calculation of mortality rates at very high ages. *Demography*. 1991; 28(.2): 293-301.
- Coeli CM, Carmargo KRCJ, Sanches KRB, Cascão AM. Sistemas de informação em saúde. In: Medronho RA, Carvalho DM, Bloch KV, Luiz RR, Werneck GL. *Epidemiologia*. São Paulo: Atheneu; 2009. p. 525-530.
- Correia LOS, Padilha BM, Vasconcelos SML. Métodos para avaliar a completude dos dados dos sistemas de informação em saúde do Brasil: uma revisão sistemática. *Ciência & Saúde Coletiva*. 2014; 19(11): 4467-4478.
- Del Popolo F. Los problemas en la declaración de la edad de la población adulta mayor en los censos. CEPAL - SERIE 8. *Población y desarrollo*. – Chile: CEPAL, 2000.
- Ewbank DC. Age misreporting and age-selective underenumeration: sources patterns and consequences for demographic analysis.– National Academy Press – Washington DC, 1981.
- FIBGE, Fundação Instituto Brasileiro de Geografia e Estatística. Estatísticas de Registro Civil. Disponível em: <http://www.sidra.ibge.gov.br/bda/acervo/acervo9.asp?e=c&p=RC&z=t&o=27>. Acesso em abril de 2018.
- Formiga MCC, Ramos PCF, Monteiro MFG. A qualidade dos dados censitários populacionais e sua associação com fatores socioeconômicos: um estudo para as mesorregiões do Estado do Rio Grande do Norte–Brasil. Disponível em <http://www.abep.org.br/publicacoes/index.php/anais/article/viewFile/1085/1049>. Acesso em 04 abril 2018.

- Frias PG, Szwarcwald CL, Souza Junior PRB, Almeida WS, Lira PIC. Correção de informações vitais: estimação da mortalidade infantil, Brasil, 2000-2009. *Rev Saúde Pública*. 2013; 47(6): 1048-58.
- Fundação Nacional de Saúde. Manual de procedimentos do sistema de informação sobre mortalidade. Brasília: Fundação Nacional de Saúde; 2001.
- Hakkert R. Fontes de Dados demográficos. - Belo Horizonte, Textos Didáticos – ABEP 1996. Disponível em <http://www.abep.nepo.unicamp.br/docs/outraspub/textosdidaticos/tdv03.pdf> Acesso em abril de 2018.
- Jannuzzi PM. Indicadores sociais no Brasil: conceitos, fontes de dados e aplicações para formulação e avaliação de políticas públicas, elaboração de estudos socioeconômicos. Alínea Editora, 2 ed, 2014.
- Latorre MRDO, Cardoso MRA. Análise de séries temporais em epidemiologia: uma introdução sobre os aspectos metodológicos. *Revista Brasileira de Epidemiologia*. 2001; 4(3): 145-152.
- 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. *Cadernos de Saúde Pública*. 2014; 30(8): 1721-1730.
- Luy M. Estimating Mortality Differences in Developed Countries From Survey Information on Maternal and Paternal Orphanhood. *Demography*. 2012; 49(2): 607-27.
- Magalhães LFA. Fontes de dados demográficos e estudos de população em Santa Catarina. *Revista NECAT*. 2015; 4(7): 23-37
- Moultrie TA , Tom A., et al. Tools for demographic estimation. International Union for the Scientific Study of Population, 2013.
- Nadalin SO. História e demografia: elementos para um diálogo. - Campinas: Associação Brasileira de Estudos Populacionais-ABEP, 2004.
- Noubissi A. L'indice de Whipple modifié: une application aux données du Cameroun, de la Suède et de la Belgique. *Population (french edition)*. 1992; 47(4): 1038-1041.

- ONU. 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. 76p. (Population Studies, n.23)
- Paes NA, Albuquerque MEE. Avaliação da qualidade dos dados populacionais e cobertura dos registros de óbitos para as regiões brasileiras. *Revista de Saúde Pública*. 1999; 33(1):43.
- Paes NA. Avaliação da cobertura dos registros de óbitos dos estados brasileiros em 2000. *Revista de Saúde Pública*. 2005; 39(6): 882-890.
- Paes NA. Qualidade das estatísticas de óbitos por causas desconhecidas dos Estados brasileiros." *Revista de Saúde Pública*. 2007; 41(3): 436-445
- Pinto AV. El pensamiento crítico en demografía. Santiago: Centro Latinoamericano de Demografía (CELADE), 1973.
- Preston SH, Elo IT. Effects of age misreporting on mortality estimates at older ages. *Population studies*. 1999; 53(2): 165-177
- Risi Júnior JB. Informação em saúde no Brasil: a contribuição da RIPSA. *Ciência & Saúde Coletiva*. 2006; 11 (4): 1049-1053.
- Romero D, 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. Disponível em http://www.alapop.org/alap/SerieInvestigaciones/InvestigacionesSI1aSi9/AdultoMayor_Partel-4.pdf. Acesso em 04 abril 2018.
- Siegel J, Swanson D. The methods and materials of demography. - San Diego: Elsevier Academic Press, 2004.
- Yi Z, Vaupel JW. Functional capacity and self-evaluation of health and life of oldest old in China. *Journal of Social Issues*. 2002; 58(4):733-748.