

# Evidências para não Utilizar o Sistema de Informações Hospitalares em Estimativas da Ocorrência de Agravos

Evidences for not Using the Hospital Information System to Estimate the Occurrence of Injuries

Sérgio Eduardo Soares Fernandes<sup>1</sup>

ORCID: 0000-0002-2511-400X

Derek Chaves Lopes<sup>2</sup>

ORCID: 0000-0002-4382-2528

André Vinícius Martignago<sup>3</sup>

ORCID: 0000-0002-2466-0500

Leila Bernarda Donato Göttems<sup>4</sup>

ORCID: 0000-0002-2675-8085

Fábio Ferreira Amorim<sup>5</sup>

ORCID: 0000-0003-0929-5733

Edgar Merchán-Hamann<sup>1</sup>

ORCID: 0000-0001-6775-9466

<sup>1</sup>Programa de Pós Graduação em Saúde Coletiva, Universidade de Brasília, Campus Darcy Ribeiro, Brasília, Distrito Federal, Brasil

<sup>2</sup>Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto - FMRP USP

<sup>3</sup>Universidade Federal de Santa Catarina, Florianópolis, Santa Catarina, Brasil

<sup>4</sup>Coordenação de Pós Graduação *Stricto Sensu*, Escola Superior de Ciências da Saúde, Brasília, Distrito Federal, Brasil

<sup>5</sup>Programa de Pós Graduação em Ciências da Saúde, Universidade de Brasília, Campus Darcy Ribeiro, Brasília, Distrito Federal, Brasil

**Autor Correspondente:** Derek Chaves Lopes. Endereço: SMHN Quadra 03, conjunto A, Bloco 1 Edifício FEPECS. CEP 70.710-907 – Brasília, DF. E-mail: derek.lopes5@gmail.com

## RESUMO

**Objetivo:** avaliar a confiabilidade do Sistema de Informações Hospitalares (SIH) para estimar a ocorrência de agravos para a população geral ou para internações hospitalares comparando-o com o Sistema de Informação sobre Mortalidade (SIM). **Métodos:** foram coletados conjuntos de dados equiparáveis entre o SIH e o SIM acerca dos óbitos ocorridos em hospitais entre 2008 e 2020 a partir da interface TABNET do Ministério da Saúde. **Resultados:** foram obtidos dados de 10.872.624 mortes intra-hospitalares registradas no SIM e 5.991.342 no SIH entre os anos de 2008 e 2020, o que corresponde a uma acurácia de 55,1% (IC 95%: 55,08% a 55,13%). O padrão é repetido em níveis diferentes com significância estatística ( $p < 0,001$ ). Análises por grupos de CID10 mostraram diferenças variáveis. **Conclusão:** apesar do uso regular de dados do SIH para inferir ocorrências populacionais, esse método é desaconselhável devido ao amplo subregistro.

**Palavras-chave:** Sistemas de Informação Hospitalar; Registros de Mortalidade; Mortalidade Hospitalar; Confiabilidade dos Dados; Registros Públicos de Dados de Cuidados de Saúde.

## ABSTRACT

**Objective:** to evaluate the reliability of the Hospital Information System (SIH) in estimating the occurrence of health problems in the general population and hospital admissions, by comparing its records with those of the Mortality Information System (SIM). **Methods:** comparable datasets from SIH and SIM were obtained for hospital deaths occurring between 2008 and 2020, using the Ministry of Health's TABNET interface. **Results:** between 2008 and 2020, SIM recorded 10,872,624 in-hospital deaths, whereas SIH recorded 5,991,342, corresponding to an overall accuracy of 55.1% (95% CI: 55.08%–55.13%). This discrepancy was consistent across different analytical levels and statistically significant ( $p < 0.001$ ). Analyses by ICD-10 code groups revealed varying degrees of difference. **Conclusion:** although SIH data are routinely used to estimate population-level health occurrences, this approach is not recommended due to substantial underreporting.

**Keywords:** Hospital Information Systems; Mortality Registries; Hospital Mortality; Data Accuracy; Public Reporting of Healthcare Data.

## INTRODUÇÃO

Sistemas de informação em saúde são instrumentos essenciais para a avaliação e monitoramento das condições de saúde e dão suporte à tomada de decisão pelos gestores de um sistema de saúde<sup>1-2</sup>. O Brasil possui importantes e abrangentes sistemas de informações em saúde, disponíveis de forma transparente e estruturada com detalhes para todas as unidades da federação<sup>1-2</sup>.

O Departamento de Informática do Sistema único de Saúde (DATASUS) dispõe de diferentes bases de dados de informações acerca de morbidade ou mortalidade: o Sistema de Informações Hospitalares (SIH), o Sistema de Informações de Mortalidade (SIM), o Sistema de Informações Ambulatoriais (SIA) e o Sistema de Informações de Agravos de Notificação (SINAN)<sup>1-2</sup>. Essas bases são alimentadas por fontes como formulários de atendimento, declarações e fichas de notificação<sup>1-2</sup>.

O SIH foi idealizado na década de 1970 e nomeado assim em 1991 com o objetivo de gerenciar o pagamento dos procedimentos relacionados às internações e facilitar o controle e auditoria hospitalar<sup>1-2</sup>. Sua base de dados contém informações parametrizadas relevantes sobre a internação hospitalar, com registros de códigos da classificação internacional de doenças (CID10) e, portanto, é frequentemente usado por pesquisadores em estudos epidemiológicos de áreas muito diversas como: queimaduras<sup>3</sup>, infectologia<sup>4</sup>, obstetrícia<sup>5</sup>, cardiologia<sup>6</sup>, pneumologia<sup>7</sup> e ortopedia<sup>8</sup>, entre outras.

O SIM, por conseguinte, foi instituído oficialmente pelo Ministério da Saúde (MS) em 1975, para registro de dados sobre as mortes ocorridas no país e só foi informatizado em 1979. Durante o processo de criação do Sistema Único de Saúde (SUS), após 1988, teve a responsabilidade da coleta e registro de dados transferida para as secretarias de saúde das unidades da federação (UF) e dos municípios<sup>1-2</sup>. O SIM disponibiliza publicamente informações epidemiológicas relevantes acerca dos óbitos ocorridos no Brasil que podem ser utilizadas pelos diferentes níveis de gestão em saúde para elaboração de políticas públicas ajustadas à realidade local<sup>1-2</sup>.

A fonte de dados do SIM é a declaração de óbito (DO), um documento padronizado pelo MS preenchido pelo médico que verificou o óbito pessoalmente ou prestou assistência ao paciente<sup>1-2</sup>. Constituem exceções as situações em que tal profissional não esteja disponível ou em que seja indicada necropsia ou outros tipos de

verificação médico-legal<sup>3-4</sup>. Ainda assim, os dados mínimos da DO serão preenchidos e farão parte da base de dados<sup>9-10</sup>.

Tanto o SIH como o SIM são frequentemente utilizados para inferências de agravos com base populacional no Brasil, contudo, embora a base do SIH contenha informações relevantes e parametrizadas, ela pode conter expressivos problemas de representatividade para esse fim, por não ser de uso universal<sup>11</sup>. Isso ocorre porque esse sistema só é utilizado por instituições que prestam serviço ao SUS, omitindo informações relativas aos serviços do sistema de saúde suplementar, assim como as relacionadas aos eventos que ocorrem em instituições de saúde de menor porte<sup>1-2</sup>.

O SIM, por outro lado, foi desenhado para captar integralmente os eventos que se propõe registrar<sup>1-2</sup>. Seu processo de coleta e registro de dados da DO é considerado seguro, com poucos riscos de enviesamento, já que antes da disponibilização da base, o MS realiza procedimentos de validação dos formulários<sup>12</sup>. Dessa forma, o SIM representa a totalidade dos óbitos oficialmente registrados, mesmo que ocorridos em serviços de saúde suplementar ou fora de estabelecimentos de saúde<sup>1-2</sup>. Esses procedimentos costumam gerar uma latência na disponibilização dos dados, que pode superar um ano<sup>12</sup>.

O estabelecimento de parâmetros de validade e confiança do SIH para inferências populacionais é essencial durante a avaliação crítica de estudos que o utilizam como fonte de dados e servem de referência metodológica para a literatura em saúde pública no país. Assim, esse estudo tem como objetivo avaliar a confiabilidade do SIH para estimar a ocorrência de agravos para a população geral ou para internações hospitalares, visto que essa proposta é encontrada na literatura, comparando-o com a base de dados do SIM.

## MÉTODO

Este é um estudo de confiabilidade do SIH em relação ao SIM, utilizando dados de óbitos de internações hospitalares de ambas as bases. Foram coletados conjuntos de dados de óbitos ocorridos em hospitais entre 2008 e 2020, período disponível comum às duas bases, estratificados por unidade da federação, mês, ano e grupos específicos de

Códigos CID10 para a causa *mortis*, a partir da interface TABNET<sup>13</sup> do Ministério da Saúde para ambas as bases.

A base de dados final do estudo foi compilada em uma tabela contendo como registros a quantidade de mortes ocorridas em cada unidade da federação em cada mês do período para cada sistema, para cada grupo do CID10. A tabela foi completada com campos calculados de taxa de mortalidade por 100.000 habitantes, tendo sido utilizado para o cálculo as populações oficiais do Ministério da Saúde validadas pelo IBGE<sup>14</sup>. O número total de registros na tabela correspondeu a doze meses de observação, multiplicados por 13 anos e, em seguida, por 27 unidades da federação, o que constituiu uma amostra de 4.212 registros. As buscas nas bases foram ajustadas para referenciar o local de ocorrência do óbito, com o intuito de minimizar os riscos de falhas de preenchimento e sua maior confiabilidade, visto se tratar de um dado obrigatório da declaração de óbito.

Os grupos selecionados foram: Doenças Infecciosas e Parasitárias (CID10 A00 a B99); Neoplasias (CID10 C00 a D48); Pneumonias (CID10 J12 a J18); Miocardiopatias Isquêmicas (CID10 I20 a I25); Insuficiência Cardíaca (CID10 I50); Doença Pulmonar Obstrutiva Crônica ou DPOC (CID10 J40 a J44); Doenças Cerebrovasculares (CID10 I60 ao I64); e Causas Externas (CID10 constantes no Capítulo XIX e XX).

A concordância entre os dados de ambos os sistemas foi representada em um gráfico de dispersão com uma linha de não diferença representando um resultado de igualdade entre os sistemas para efeito de comparação entre os eixos. A análise foi feita com cálculo do coeficiente de correlação de Pearson e a utilização de um gráfico de Bland-Altman com limites de 1,96 desvios padrão para cima ou para baixo em relação à média para se obter um intervalo de confiança de 95% (IC95%).

Para dar significância estatística à diferença entre as observações foi realizado um teste T para amostra única em comparação com o valor zero para todas as análises. Os testes de hipótese admitiram erro  $\alpha$  de 5% e os índices de correlação foram considerados fortes quando mostraram valor absoluto superior a 0,8.

Quando os resultados assumiram um padrão visual linear, foi calculada uma função de regressão linear simples e o coeficiente de determinação ( $R^2$ ) pelo método dos mínimos quadrados, assim como o coeficiente angular da curva de regressão a partir

da tangente da projeção sobre o eixo das abscissas, padronizado para representar o SIM como base de comparação.

As análises de dispersão foram feitas com: o número absoluto de óbitos; a taxa de mortalidade por cem mil habitantes para as cinco regiões do país; e para os anos do período observado, com o objetivo de estimar o viés de temporalidade. Nos casos em que os gráficos mostraram grupos de pontos de dispersão desalinhados com padrão do restante do país, foi produzido um novo gráfico de dispersão, estratificado pelas Unidades da Federação (UF) da região que mostrou diferença, para identificação dos padrões que explicavam a disparidade.

Foi ainda produzida uma representação da série histórica anual de registro de mortes hospitalares pelo SIM e SIH e feito o cálculo de regressão linear simples com o R2 para os anos de 2008 a 2019. O ano de 2020 foi excluído desta regressão pelo padrão atípico relacionado à Pandemia de COVID-19.

Por fim, foram construídas representações da série histórica anual de registros de mortes hospitalares para cada um dos grupos de CID10.

Todos os elementos gráficos foram produzidos com o *software* Microsoft Excel® 2014, utilizado também para os cálculos estatísticos, conferidos pelo *software* Python 3.10. Este estudo foi dispensado de avaliação por comitê de ética em pesquisa por ter utilizado exclusivamente dados disponíveis publicamente, nos termos da resolução N° 674/2022 do Conselho Nacional de Saúde.

## RESULTADOS

Ao todo, foram registradas 10.872.624 mortes intra-hospitalares no SIM e 5.991.342 no SIH entre os anos de 2008 e 2020, o que corresponde a uma acurácia de apenas 55,1% (IC95%: 55,08% a 55,13%) em relação ao SIM. O gráfico de concordância entre os óbitos registrados mensalmente por estado brasileiro em cada sistema demonstra um desvio da distribuição para o eixo do SIM com incremento aparentemente regular do distanciamento à medida que o número de mortes aumenta (Figura 1a).

Notavelmente, a região Sudeste se destaca das demais, à custa dos estados de São Paulo, pelo seu elevado número absoluto de óbitos e, principalmente, do Rio de

Janeiro, que mostra claramente uma baixa concordância entre as bases além de perder o padrão de distribuição linear, comum a todas as outras UF (Figuras 1b e 1d).

O coeficiente de correlação de Pearson demonstrou associação positiva pelo menos moderada para todos os estados. Os menores índices ocorreram no Rio de Janeiro (0,71), Acre (0,75), Tocantins (0,68) e Distrito Federal (0,79) (dados não apresentados em tabelas). Em todos os estados houve significância estatística com p-valor < 0,001. O teste T para uma amostra comparando com o valor zero da diferença entre as duas bases mostrou p-valor < 0,001.

A análise de Bland-Altman mostrou um desvio médio de 1.160 mortes do SIH em relação ao SIM e um padrão de distribuição atípico (Figura 1b) não normal da diferença entre as bases com p-valor < 0,001 pelo teste de Shappiro-Wilk.

As mesmas análises com as taxas de mortalidade por 100.000 habitantes mostraram padrões similares aos da análise com números absolutos, porém com maior dispersão (Figura 2). A comparação entre as figuras 1d e 2d - quando há ajuste dos dados de São Paulo pela população - demonstra uma mudança visualmente importante, aproximando a representação gráfica dos demais estados da região. Embora com índices de correlação de Pearson menores, a associação entre os sistemas persiste na maioria das UF, sendo menor no Rio de Janeiro (0,56), Acre (0,62), Amapá (0,75), Tocantins (0,58), Paraíba (0,79), Sergipe (0,78) e Distrito Federal (0,65). Todas as UF tiveram associações estatisticamente significativas ( $p < 0,001$ ). Assim como para os valores absolutos, o teste T para uma amostra comparando com o valor zero da diferença mostrou p-valor < 0,001.

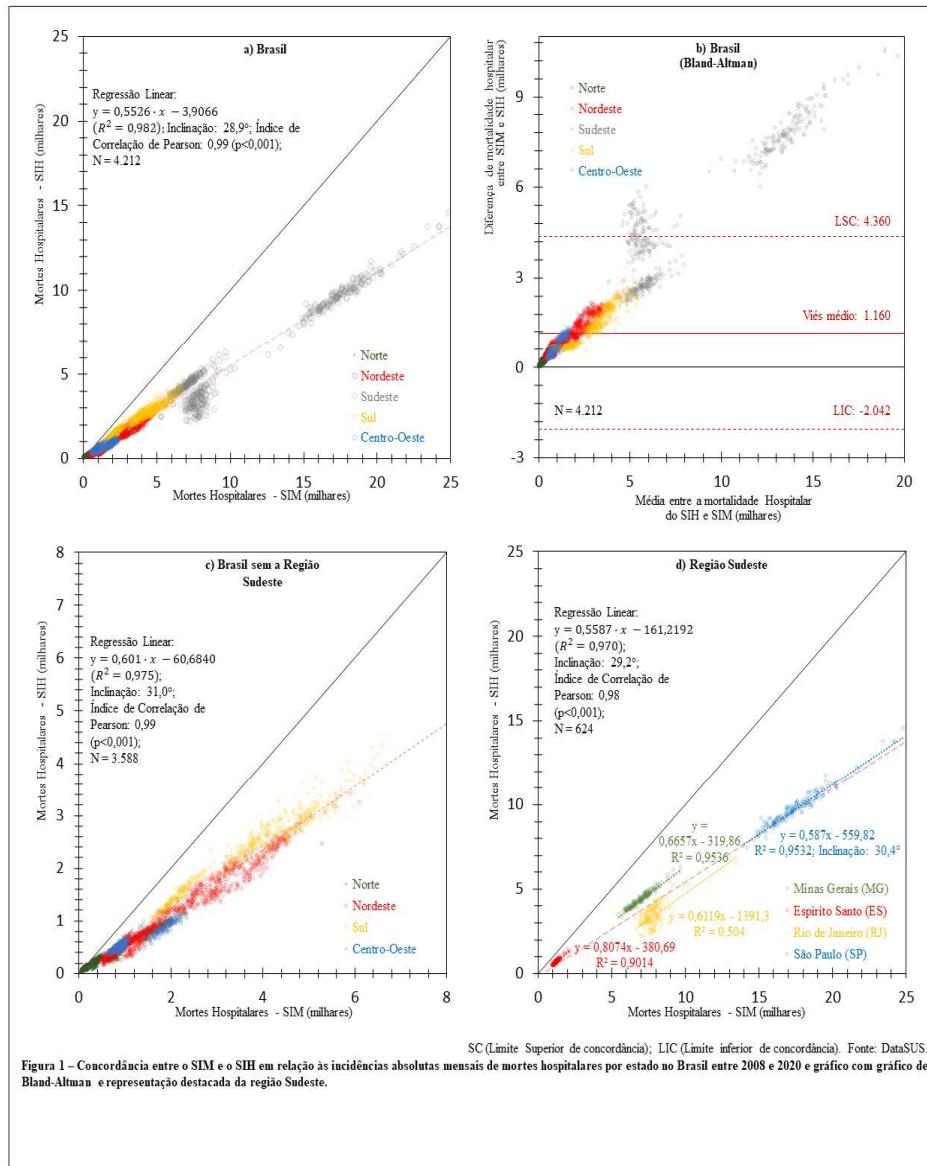


Figura 1 – Concordância entre o SIM e o SIH em relação às incidências absolutas mensais de mortes hospitalares por estado no Brasil entre 2008 e 2020 e gráfico com gráfico de Bland-Altman e representação destacada da região Sudeste. Fonte: DataSUS.

A análise de Bland-Altman para a mortalidade por 100.000 habitantes, assim como na Figura 1, apresentou um padrão atípico não normal de distribuição ( $p$ -valor  $< 0,001$  pelo teste de Shapiro-Wilk) com grande quantidade de valores fora dos limites do IC95%. O desvio médio foi de 13,7 mortes por 100.000 habitantes (Figura 2.b).

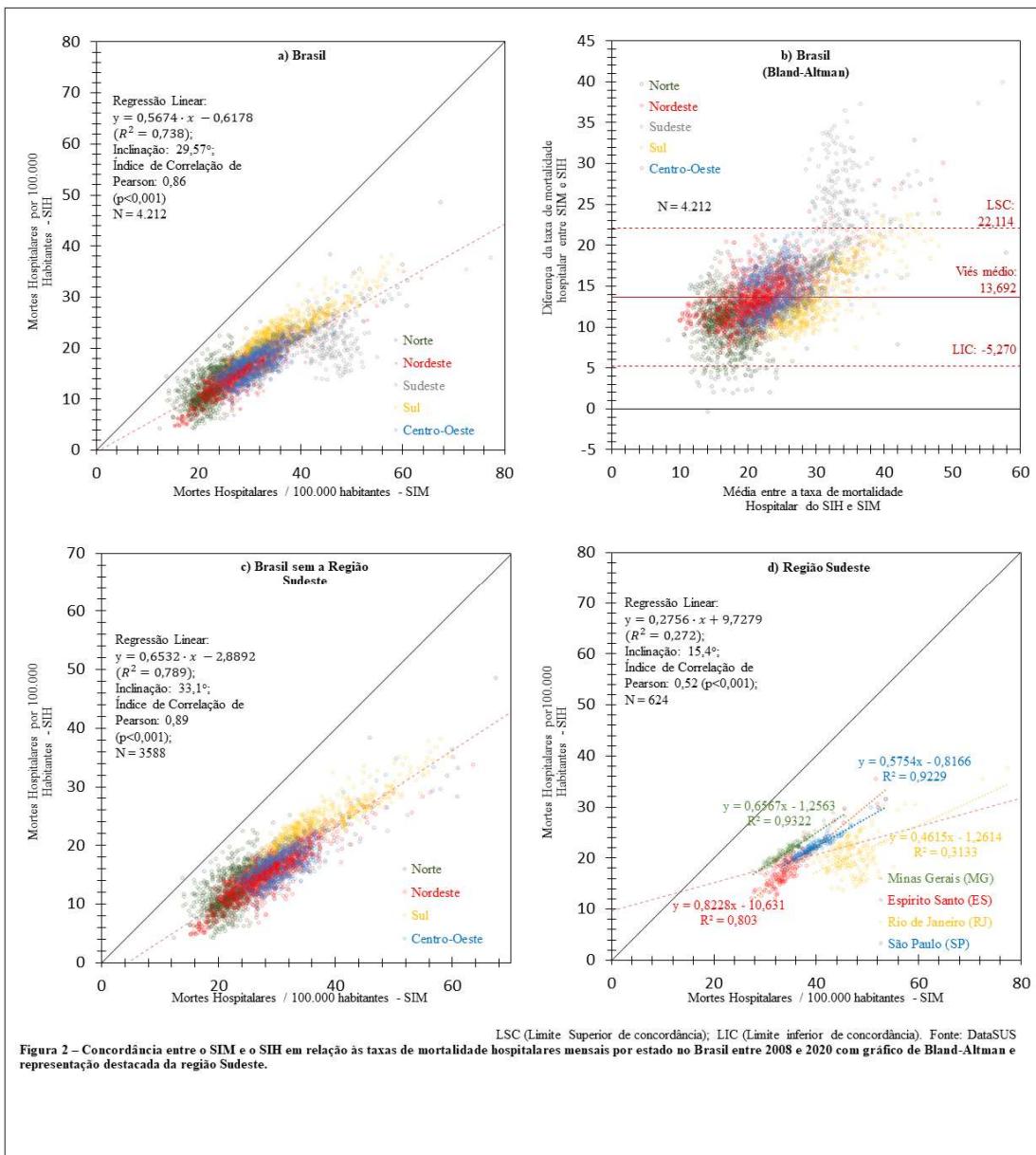
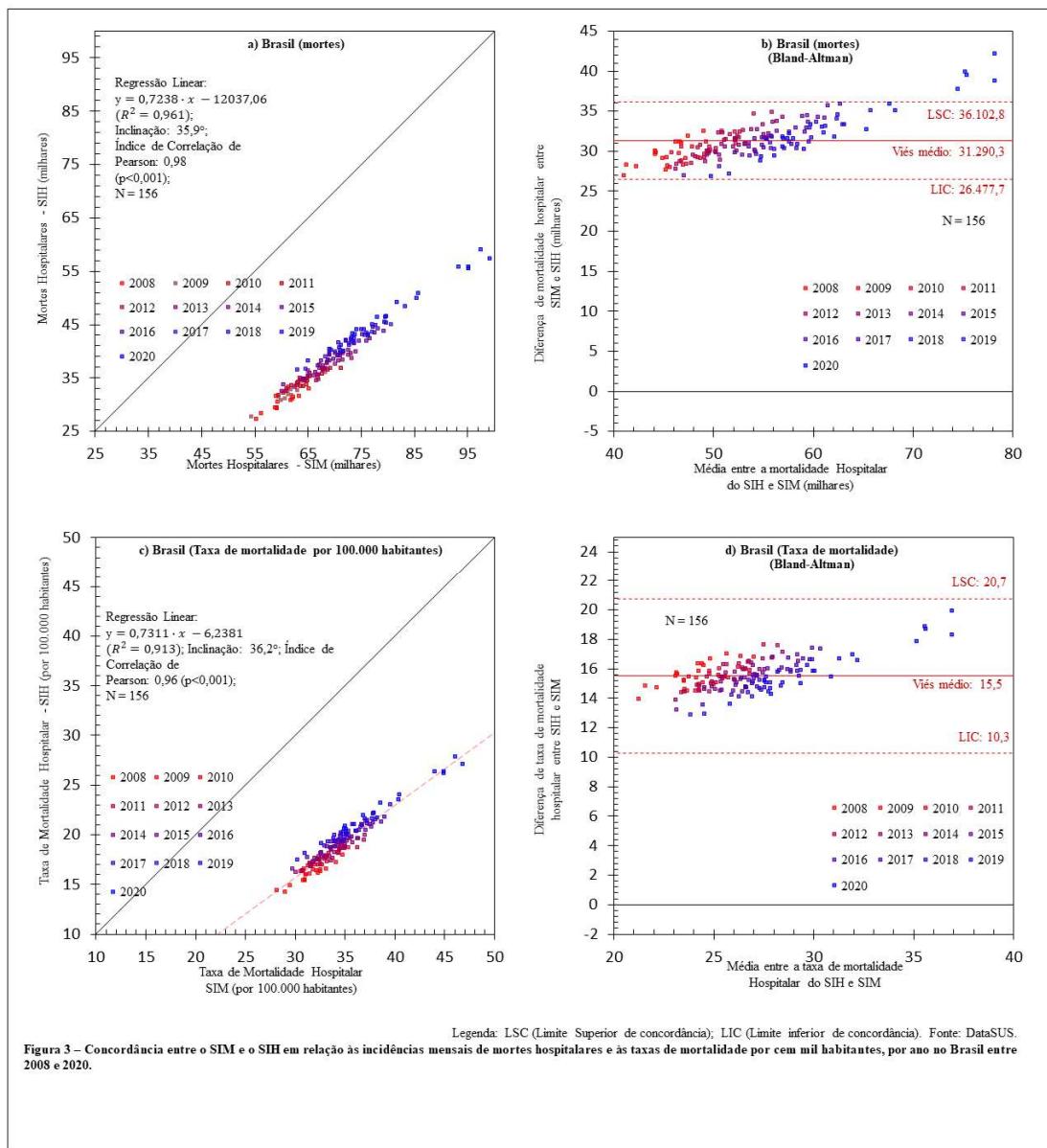


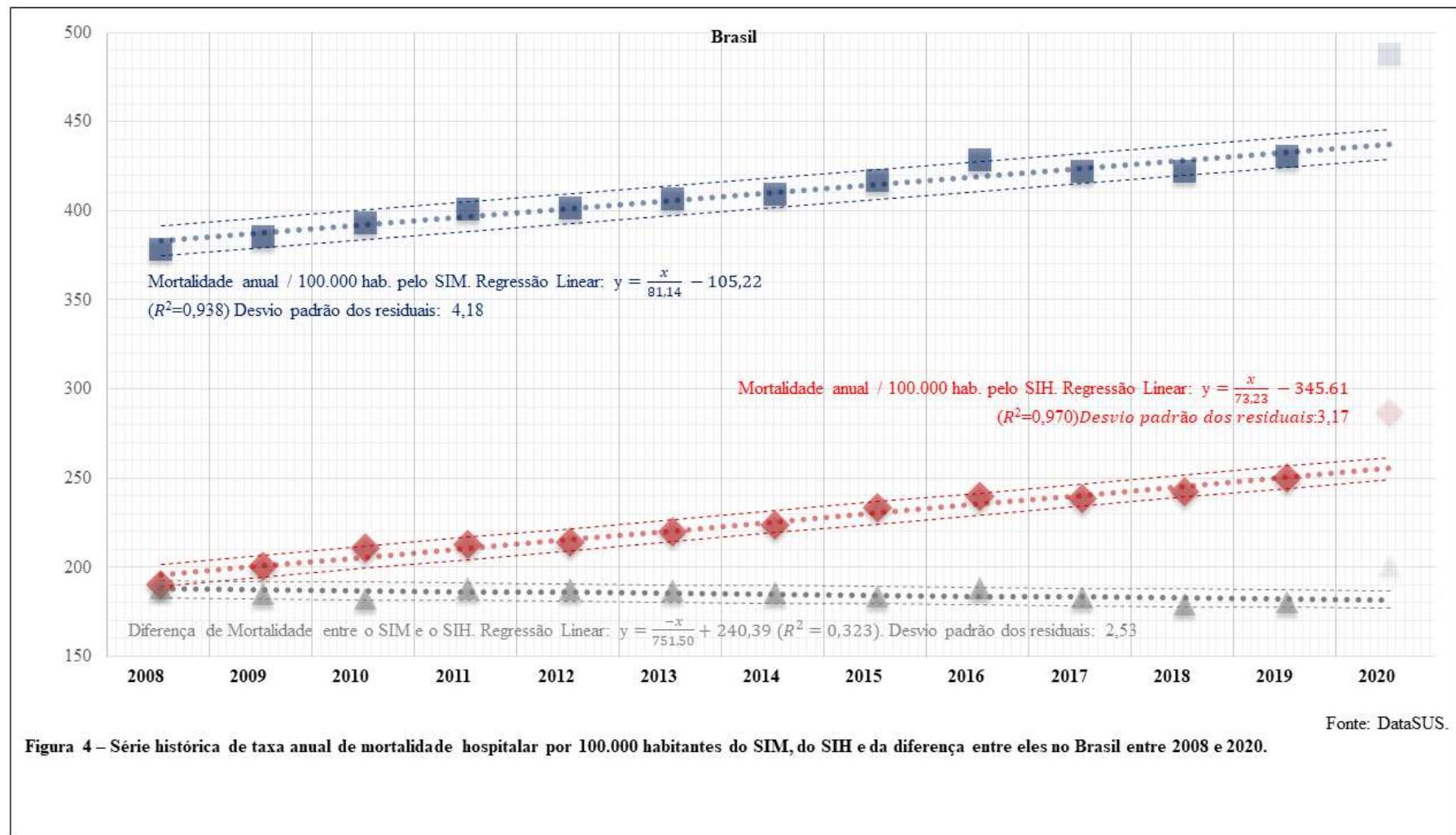
Figura 2 – Concordância entre o SIM e o SIH em relação às taxas de mortalidade hospitalares mensais por estado no Brasil entre 2008 e 2020 com gráfico de Bland-Altman e representação destacada da região Sudeste.

A distribuição temporal para o Brasil demonstra um padrão linear mais homogêneo, entretanto com tendência de crescimento tanto em números absolutos quanto por 100.000 habitantes, perceptível pela diferença na posição entre triângulos e quadrados na Figura 3. Embora a análise de Bland-Altman mostre uma distribuição mais contida pelos limites, o viés médio estava consistentemente distante de zero nas duas situações e o teste T para uma amostra comparando com o valor zero da diferença também mostrou  $p$ -valor  $< 0,001$ .

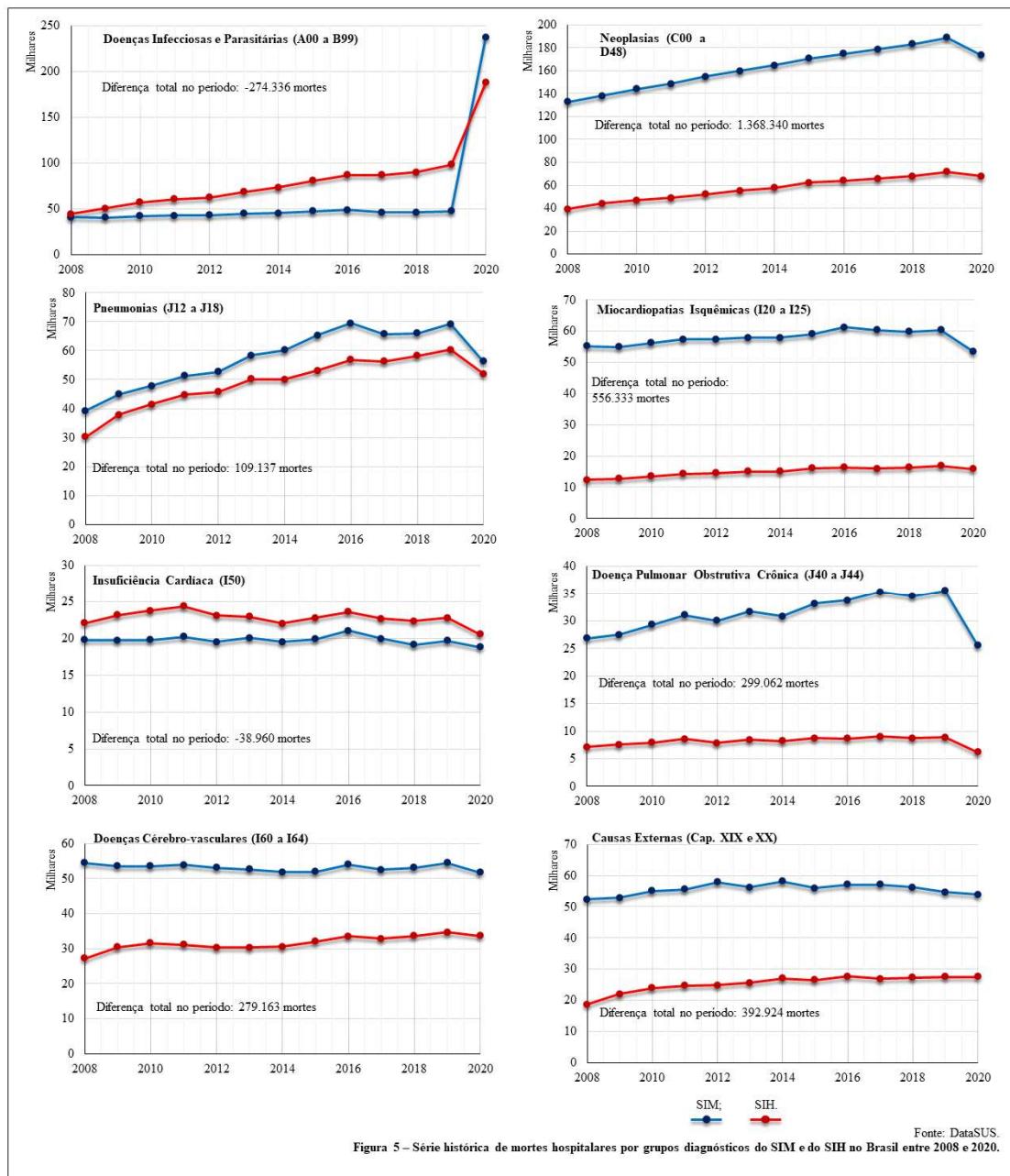


A tendência de crescimento do número mortes por 100.000 habitantes é ainda mais facilmente percebida na Figura 4. Com os dados observados para o período de 2008 a 2019, o ritmo de crescimento da mortalidade pelo SIH pareceu maior que o obtido pelo SIM. Na situação hipotética em que as tendências do período se mantivessem indefinidamente, as curvas só se encontrariam em maio do ano 2439.

Percebe-se ainda que as diferenças entre as bases de dados se comportam de forma heterogênea quando se comparam os vários grupos de CID estudados (Figura 5). Grupos como o de doenças infecciosas e o de insuficiência cardíaca mostram registro de mortes do SIH superior ao do SIM de forma persistente (Figura 5).



Também chamam a atenção, as mudanças de tendência ocorridas em 2020, mais evidentes nos grupos: doenças infecciosas e parasitárias, neoplasias, pneumonias, miocardiopatias isquêmicas e DPOC (Figura 5).



A diferença total observada entre os dois sistemas somou 4.881.282 mortes no período estudado, o que corresponde a apenas 44,9% (IC95%: 44,87% a 44,92%) dos 10.872.624 óbitos registrados pelo SIM.

## DISCUSSÃO

A distribuição linear no gráfico de dispersão entre os óbitos hospitalares pelo SIH e SIM sugere haver um padrão de comportamento com alguma previsibilidade, que aumenta a perda de registros pelo SIH conforme aumentam os registros do SIM. Esse efeito pode ser decorrente de múltiplas causas. Entre elas, pode estar o fato de os óbitos ocorridos em serviços privados não serem registrados no SIH<sup>15-16</sup>. Considerando que a cobertura média brasileira de saúde suplementar no período foi de 23,07%, conforme a Agência Nacional de Saúde Suplementar (ANS), e que o registro do SIH mostrou uma diferença comparativa ao SIM que corresponde a 44,9% dos registros, é razoável supor que essa explicação não é suficiente como única hipótese<sup>17</sup>.

É possível ainda que a diferença entre os óbitos ocorridos em pronto-socorros e prontos atendimentos não sejam incluídos nos registros do SIH<sup>15-16,18</sup>, uma vez que os pacientes atendidos nesses locais ficam em observação por até 48 horas, antes da solicitação da Autorização de Internação Hospitalar (AIH) e sua inserção no sistema<sup>18</sup>. Consequentemente, óbitos nesse período não são registrados como internações e não são registrados pelo SIH, mesmo tendo ocorrido efetivamente dentro de um hospital ou serviço de saúde<sup>15-16</sup>.

Além disso, os gestores locais credenciam seus hospitais junto ao MS para diferentes procedimentos e os contratos podem variar conforme as necessidades locais do SUS, a vocação ou especialidade do hospital, e sua capacidade produtiva, entre outros fatores. Dessa forma, a seleção dos códigos de doenças para registro nas AIH pode ser direcionada para atender a contratação efetuada<sup>18</sup>.

Ademais, o código CID10 de saída da internação registrado no SIH nem sempre reflete a causa *mortis* registrada na declaração de óbito, já que em internações com múltiplas morbidades pode haver a tendência de indicar uma causa mais relevante e não exatamente a causa da morte. Essa hipótese é capaz de explicar satisfatoriamente os casos em que o registro de óbitos no SIH foi superior ao do SIM como os observados com doenças infecciosas e parasitárias e com insuficiência cardíaca. A própria organização da CID-10 também pode contribuir com esse efeito, já que os códigos de doenças infecciosas e parasitárias não estão apenas nos capítulos A e B, mas também espalhados em vários outros capítulos conforme a pertinência da localização anatômica<sup>19</sup>.

Uma outra hipótese plausível guarda relação com o teto de remuneração estabelecido pelo MS para a produção hospitalar. Isso significa que, quando ultrapassado o limite de produção contratado, pode não haver incentivo para o preenchimento adequado das AIH pelo hospital, uma vez que essa atividade administrativa implica em custos sem o retorno proporcional de recursos<sup>15-16</sup>.

A diferença entre os sistemas é visível nos gráficos de dispersão das Figuras 1, 2 e 3 e, com a análise de Bland-Altman, foi possível quantificar o desvio do SIH em relação ao SIM. Uma diferença tão grande para um estrato tão significativo das internações hospitalares suporta a inferência de que os registros de internações com outros desfechos podem ser, de forma muito semelhante, imprecisos e, portanto, o SIH não deveria ser utilizado para estimativa de incidências e prevalências de agravos, principalmente em estados nos quais há maior cobertura do sistema de saúde privado, onde este efeito pode ser ainda maior.

Diversos estudos corroboram essa hipótese nas bases indexadas. A pesquisa realizada por Melo (2004)<sup>20</sup> encontrou um número de registros de óbitos por infarto agudo do miocárdio 5,5 vezes maior no SIM que no SIH e atribuiu esse efeito a guias de AIH não emitidas, escolhas de outros códigos para registro nas AIH e subnotificação de óbitos hospitalares<sup>20</sup>.

Um estudo transversal realizado em Londrina e Maringá por Tomimatsu et al. (2009)<sup>21</sup> com uma amostra de 6.388 pacientes internados por causas externas, comparando com os registros do SIH, mostrou uma capacidade de identificação de apenas 69% dos casos levantados pelo inquérito<sup>21</sup>.

Maia-Elkhoury (2007)<sup>22</sup> utilizou o método de captura e recaptura para comparar a mortalidade por leishmaniose visceral entre três das bases de dados oficiais Brasileiras: SIM, SIH e o SINAN para óbitos. O estudo chega à conclusão de que todos os sistemas de registro possuem erros relevantes para o desfecho estudado, sendo o SINAN-óbitos o sistema mais preciso (detectou 62,2% dos 9.481 casos levantados pelos três sistemas no período estudado)<sup>22</sup>. O SIM mostrou uma detecção de 4,1% contra 33,7% do SIH<sup>22</sup>. Esse resultado parece coerente com o gráfico de doenças infecciosas e parasitárias apresentado na Figura 5, em que o SIH aparenta detectar mais mortes com esses códigos que o SIM.

Os achados de Maia-Elkhoury (2007)<sup>22</sup> provocam a reflexão de que é muito difícil estabelecer a causa *mortis* com precisão, principalmente em situações complexas com grandes cadeias de morbidades e comorbidades. Com efeito, este viés do SIM é citado na literatura, sobretudo quando são utilizados métodos de *linkage*<sup>11,18,23-26</sup>. Entretanto, os métodos de validação e qualificação da informação, tem proporcionado segurança em vários conjuntos de causas e há muito vem sendo utilizado como referência em estudos de mortalidade<sup>27-28</sup>, por possuir indicadores de subnotificação<sup>9,29-30</sup>.

Independentemente do viés de registro da causa *mortis*, o número total de óbitos é bastante confiável, uma vez que possui um processo de qualificação baseado em múltiplos processos, não incluindo apenas os óbitos cujo corpo não possa ser encontrado. Por conseguinte, o número absoluto de mortes intra-hospitalares foi capaz de oferecer uma base de comparação confiável para o propósito deste estudo.

O efeito de aumento da perda de registros do SIH com o aumento do número de mortes hospitalares pelo SIM pareceu ser bastante consistente uma vez que se repete quando a observação é o número mensal absoluto ou o relativo de óbitos tanto para as UF, quanto para o país, com significância estatística em todos os casos. Já a diferença entre os dois sistemas pareceu apresentar uma tendência estável no período que não deverá ter mudanças significativas mesmo em longos prazos, a não ser que haja uma ação específica nesse sentido, capaz de melhorar o registro no SIH de forma mais contundente.

Uma das UF se destacou por perder a distribuição linear assumida pelas demais, tanto nos registros em número absoluto quanto relativo. O Rio de Janeiro parece apresentar um comportamento de registro de óbitos muito destoante por ser pouco previsível e manter médias de registro do SIH em relação ao SIM inferiores às do restante do país. Isso sugere que o Estado pode não possuir informações confiáveis para tomadas de decisão acerca de políticas públicas de saúde e, dessa forma, possui grandes oportunidades de melhoria, caso consiga controlar o processo de registro.

Os achados discutidos reforçam a necessidade de adotar estratégias para o uso de dados do SIH em outros estudos. Entre as opções, destaca-se a disponibilização facilitada de acesso às bases com *linkage* - atualmente é necessário realizar uma solicitação online através da Rede Nacional de Dados em Saúde (RNDS). Ainda, é

preciso realizar uma série de passos metodológicos, por vezes não familiares a pesquisadores menos habituados ao trabalho com grandes bases de dados<sup>31</sup>. A criação de um padrão único de identificação adotado com a centralização de múltiplas bases é uma alternativa interessante, como realizado no modelo de saúde digital dinamarquês<sup>32</sup>. Todavia, isso requer um esforço longitudinal e de diversos atores, dificultado em um país de dimensões continentais como o Brasil<sup>33</sup>.

O efeito da COVID-19 pode ser observado na Figura 5 com a elevação dos óbitos de doenças infecciosas e parasitárias em 2020. Embora o número absoluto de todas as mortes tenha aumentado em 2020, também foi observado um efeito compensatório pela redução das mortes por doença pulmonar obstrutiva crônica, pneumonias, neoplasias, entre outras. Embora essas doenças possam ser ligadas à cadeia de morbidades que contribuiu com a morte, a causa *mortis* provavelmente foi registrada como COVID-19.

Este estudo possui limitações inerentes ao método empregado. Conforme já discutido, embora a base de dados do SIM, usada para comparação, seja confiável não é possível afastar completamente viés de seleção por erros de registro de CID10. Contudo, a probabilidade de o tamanho desse efeito ser significativo é baixa, uma vez que, para o objetivo principal, foi utilizado o número total de óbitos. Em contrapartida, pode não ser seguro generalizar esses resultados para análises estratificadas pelos códigos de doenças conforme sugere a Figura 5.

Outro ponto que precisa ficar claro é que esta análise levou em conta apenas os casos comuns entre o SIH e o SIM: os óbitos intra-hospitalares. Portanto, não é possível confiar que o comportamento dos casos que não evoluíram para óbito segundo o SIH seja similar aos analisados nesse estudo. De forma semelhante, ressalta-se que os dados foram analisados com as bases de dados não ligadas por chaves de identificação e, portanto, não foi possível compreender as diferenças de intersecção e exclusividade de cada uma. A opção por não realizar o *linkage* de bases, contudo, foi justamente pelo amplo uso delas dessa maneira em outros trabalhos publicados. Contudo, não se espera que os resultados seriam diferentes para atingir os objetivos propostos, uma vez que a estimativa de dados exclusivos do SIH é baixa em decorrência das análises realizadas com todos os óbitos.

## CONSIDERAÇÕES FINAIS

Os achados sugerem que apesar do uso regular de dados do SIH para inferir ocorrências populacionais, esse método é desaconselhável devido ao amplo sub registro, principalmente em regiões com menor cobertura populacional exclusiva pelo SUS. Esses resultados não se aplicam ao uso do SIH para estudos de produção na saúde pública.

Ainda, na prática, demonstram a importância de utilizar formas de ligar diferentes bases de dados em análises ou pesquisas de informações em múltiplas bases, o que por vezes não é feito. A opção por pesquisas originais, com dados primários, pode ser a melhor opção em situações nas quais a confiabilidade elevada é necessária. Por fim, reforça a importância da cautela na interpretação de dados dos SIH e do SIM, sendo necessário amplo conhecimento de sua abrangência, vantagens e limitações para o uso de pesquisas científicas.

## REFERÊNCIAS

1. Brasil, Ministério da Saúde, Organização Pan-Americana da Saúde, Fundação Oswaldo Cruz. A experiência brasileira em sistemas de informação em saúde. 1aed. Ministério da Saúde, organizador. Vol. 1. Brasília, DF; 2009. 148 p. Disponível em: [https://bvsms.saude.gov.br/bvs/publicacoes/experiencia\\_brasileira\\_sistemas\\_saude\\_volumel.pdf](https://bvsms.saude.gov.br/bvs/publicacoes/experiencia_brasileira_sistemas_saude_volumel.pdf).
2. Brasil, Ministério da Saúde, Organização Pan-Americana da Saúde, Fundação Oswaldo Cruz. A experiência brasileira em sistemas de informação em saúde. Volume 2. 1aed. Ministério da Saúde, organizador. Vol. 2. Brasília, DF; 2009. 148 p. Disponível em: [https://bvsms.saude.gov.br/bvs/publicacoes/experiencia\\_brasileira\\_sistemas\\_saude\\_volume2.pdf](https://bvsms.saude.gov.br/bvs/publicacoes/experiencia_brasileira_sistemas_saude_volume2.pdf).
3. Citron I, Amundson J, Saluja S, Guilloux A, Jenny H, Scheffer M, et al. Assessing burn care in Brazil: An epidemiologic, cross-sectional, nationwide study. *Surgery*. 2017;163(5):1165–72. DOI: 10.1016/j.surg.2017.11.023.
4. Tenório M do P, Beraldi G. Iniciação científica no Brasil e nos cursos de medicina. *RevAssocMed Bras.* 2010;56(4):390–3. DOI: 10.1590/s0104-42302010000400007.

5. Silva N de Á, Errico L de SP de. Incidência do near miss materno entre as mulheres atendidas no sistema público de saúde do Brasil no ano de 2019. Universidade Federal de Minas Gerais; 2021. Disponível em: <https://repositorio.ufmg.br/handle/1843/36624>

6. Franco PYS, Rocha MV, Guedes BR, Almeida GP, Afonso GV, Ferreira BA, et al. Incidência do infarto agudo do miocárdio em idosos no norte de Minas Gerais no período de 2008 a 2018. *Brazilian Journal of Health Review*. 2021;4(4):15423–32. DOI: 10.34119/bjhrv4n4-082.

7. Vieira SC, Gualdi LP. Incidência de doenças respiratórias na região Nordeste do Brasil. Universidade Federal do Rio Grande do Norte; 2019. Disponível em: <https://repositorio.ufrn.br/items/66754294-5045-48e4-89f2-4a49268cc94a>.

8. de Andrade JP, Silva DZ da, Patrício DS. Incidência dos casos de fratura de fêmur no Brasil entre 2015 e 2020 através de dados epidemiológicos do DATASUS: faixa etária e gênero. *ScientiaGeneralis*. 2020;1(3):84–91. Disponível em: <https://scientiageneralis.com.br/index.php/SG/article/view/v1n3a9>.

9. Brasil, Ministério Da Saúde, Fundação Nacional de Saúde. Manual de Procedimento do Sistema de Informações sobre Mortalidade. 1o ed. Ministério da Saúde, organizador. Vol. Volume 1, Manual de Procedimentos do Sistema de Informações sobre Mortalidade. Brasília, DF; 2001. 36 p. Disponível em: [https://bvsms.saude.gov.br/bvs/publicacoes/sis\\_mortalidade.pdf](https://bvsms.saude.gov.br/bvs/publicacoes/sis_mortalidade.pdf).

10. Ministério da Saúde. Fundação Nacional de Saúde. Manual de Instruções para o Preenchimento da Declaração de Óbito. 4a. Ministério da Saúde, organizador. Brasília, DF; 2011. 54 p. Disponível em: [https://svs.aids.gov.br/download/manuais/Manual\\_Instr\\_Preench\\_DO\\_2011\\_jan.pdf](https://svs.aids.gov.br/download/manuais/Manual_Instr_Preench_DO_2011_jan.pdf).

11. Drumond E de F, Machado CJ, Vasconcelos M do R, França E. Utilização de dados secundários do SIM, Sinasc e SIH na produção científica brasileira de 1990 a 2006. *Rev Bras Estud Popul*. 2009;26(1):7–19. DOI: 10.1590/S0102-30982009000100002.

12. Brasil, Ministério da Saúde, Secretaria de Vigilância em Saúde. Sistema de Informação sobre Mortalidade (SIM) - Nota Informativa [Internet]. 2022 [citado 2022

jun 6]. Available from: <http://svs.aids.gov.br/dantps/centrais-de-conteudos/dados-abertos/sim/>.

13. Brasil, Ministério da Saúde. DATASUS - Ministério da Saúde - Transferência de Arquivos [Internet]. [citado 2022 abr 25]. Available from: <https://datasus.saude.gov.br/transferencia-de-arquivos/>.

14. Brasil, Ministério da Saúde. População residente - Estudo de estimativas populacionais por município, idade e sexo 2000-2021 - Brasil [Internet]. [citado 2022 abr 25]. Available from: <http://tabnet.datasus.gov.br/cgi/deftohtm.exe?sinannet/cnv/leishvdf.def>.

15. Brasil, Ministério da Saúde. Manual Técnico e Operacional do Sistema de Informação Hospitalar - orientações técnicas. Versão 1. Ministério da Saúde, organizador. Secretaria de Atenção à Saúde. Brasília, DF; 2012. 110 p. Disponível em: [https://bvsms.saude.gov.br/bvs/publicacoes/manual\\_tecnico\\_sistema\\_informacao\\_hospitalar\\_sus.pdf](https://bvsms.saude.gov.br/bvs/publicacoes/manual_tecnico_sistema_informacao_hospitalar_sus.pdf).

16. Brasil, Ministério da Saúde, Secretaria de Atenção à Saúde, Departamento de Regulação Avaliação e Controle. Manual técnico do Sistema de Informação Hospitalar. 1a. Editora do Ministério da Saúde, organizador. Brasília, DF; 2007. 198 p. Disponível em: [https://bvsms.saude.gov.br/bvs/publicacoes/07\\_0066\\_M.pdf](https://bvsms.saude.gov.br/bvs/publicacoes/07_0066_M.pdf).

17. Brasil, Ministério da Saúde. Agência Nacional de Saúde Suplementar. Dados e Indicadores do Setor. 2022. Disponível em: <https://www.gov.br/ans/pt-br/acesso-a-informacao/perfil-do-setor/dados-e-indicadores-do-setor>

18. Escola Politécnica de Saúde Joaquim Venâncio /Fiocruz, Instituto de Estudos em Saúde Coletiva UFRJ, Instituto de Medicina Núcleo Social/Uerj, Núcleo de Assessoria T e E em S /UFJF. Pesquisa “Uso integrado de Bases de Dados na Avaliação em Saúde”. 1a. Fiocruz, organizador. Rio de Janeiro; 2008. 62 p. Disponível em: <https://www.epsjv.fiocruz.br/publicacao/livro/uso-integrado-de-bases-de-dados-na-avaliacao-em-saude>

19. Brasil, Ministério da Saúde. Morbidade Hospitalar do SUS Lista de Tabulação para Morbidade [Internet]. 2022 [citado 2022 jun 5]. Available from: <http://tabnet.datasus.gov.br/cgi/sih/mxcid10lm.htm>.

20. Melo ECP, Travassos C, Carvalho MS. Qualidade dos dados sobre óbitos por infarto agudo do miocárdio, Rio de Janeiro. *RevSaude Publica*. 2004;38(3):385–91. DOI: 10.1590/S0034-89102004000300008.

21. Tomimatsu MFAI, Andrade SM de, Soares DA, Mathias TA de F, Sapata M da PM, Soares DFP de P, et al. Qualidade da informação sobre causas externas no Sistema de Informações Hospitalares. *RevSaude Publica*. 2009;43(3):413–20. DOI: 10.1590/s0034-89102009005000019.

22. Maia-Elkhoury ANS, Carmo EH, Sousa-Gomes ML, Mota E. Análise dos registros de leishmaniose visceral pelo método de captura-recaptura. *Medicina (B Aires)*. 2007;41(6):931–7. DOI: 10.1590/S0034-89102007000600007.

23. Rafael RAA, Ribeiro VS, Cavalcante MCV, dos Santos AM, Simões VMF. Relacionamento probabilístico: recuperação de informações de óbitos infantis e natimortos em localidade no Maranhão, Brasil. *CadSaude Publica*. 2011;27(7):1371–9. DOI: 10.1590/s0102-311x2011000700012.

24. Carmo RA do, Policena GM, Alencar GP, França EB, Bierrenbach AL. Subnotificação de óbitos por AIDS no Brasil: *linkage* dos registros hospitalares com dados de declaração de óbito. *Cien Saude Colet*. 2021;26:1299–310. DOI: 10.1590/1413-81232021264.15922019.

25. Lima CR de A, Schramm JM de A, Coeli CM, da Silva MEM. Revisão das dimensões de qualidade dos dados e métodos aplicados na avaliação dos sistemas de informação em saúde. *Cad Saude Publica*. 2009;25(10):2095–109. DOI: 10.1590/s0102-311x2009001000002.

26. Sanni Ali M, Ichihara MY, Lopes LC, Barbosa GCG, Pita R, Carreiro RP, et al. Administrative data linkage in Brazil: Potentials for health technology assessment. *Front Pharmacol*. 2019;10(SEP):1–20. DOI: 10.3389/fphar.2019.00984.

27. Paixão ES, Campbell OMR, Rodrigues LC, Teixeira MG, Costa M da CN, Brickley EB, et al. Validating linkage of multiple population-based administrative databases in Brazil. *PLoSOne*. 2019;14(3):1–11. DOI: 10.1371/journal.pone.0214050.

28. Costa LFL, Montenegro M de MS, Rabello Neto D de L, Oliveira ATR de, Trindade JE de O, Adair T, et al. Estimating completeness of national and subnational

death reporting in Brazil: application of record linkage methods. *BMC*. 2020;18(22):1–12. DOI: <https://doi.org/10.1186/s12963-020-00223-2>.

29. Brasil, Ministério da Saúde. Indicadores de cobertura que utilizam a metodologia do Busca Ativa [Internet]. Ministério da Saúde. 2022 [citado 2022 jun 6]. Available from: <http://svs.aids.gov.br/dantps/acesso-a-informacao/acoes-e-programas/busca-ativa/indicadores-de-saude/mortalidade/>.

30. Silva PH dos S, Cirilo SSV, Silva LSSBF. Déficit e ocupação de leitos de unidade de terapia intensiva adulto do Sistema Único de Saúde no estado do Piauí sob a ótica da COVID-19. *Vigilância Sanitária em Debate: Sociedade, Ciência & Tecnologia*. 2020;8(3):61–9. DOI: <https://doi.org/10.22239/2317-269x.01606>.

31. Garcia KKS, Miranda CB, de Sousa FN e F. Procedimentos para vinculação de dados da saúde: aplicações na vigilância em saúde. *Epidemiol. Serv. Saúde* [Internet]. 2022 [citado 2025 Ago 11] ; 31( 3 ): e20211272. Disponível em: [http://scielo.iec.gov.br/scielo.php?script=sci\\_arttext&pid=S1679-49742022000301100&lng=pt](http://scielo.iec.gov.br/scielo.php?script=sci_arttext&pid=S1679-49742022000301100&lng=pt). Epub 29-Set-2022. <http://dx.doi.org/10.1590/s2237-96222022000300004>.

32. Rachid R, Fornazin M, Castro L, Gonçalves LH, Penteado BE. Saúde digital e a plataformaização do Estado brasileiro. *Ciência & Saúde Coletiva*, 2023; 23:2143-2153. DOI: <https://doi.org/10.1590/1413-81232023287.14302022>.

33. Jensen TB, Thorseng, AA. Building National Healthcare Infrastructure: The Case of the Danishe-Health Portal. In: Aanestad, M., Grisot, M., Hanseth, O., Vassilakopoulou, P. (eds) *Information Infrastructures within European Health Care*. *Health Informatics*. Springer, Cham. [https://doi.org/10.1007/978-3-319-51020-0\\_13](https://doi.org/10.1007/978-3-319-51020-0_13)