

Avaliação dos dados de mortes por COVID-19 nas bases dos cartórios do RC-Arpen, SIVEP-Gripe e SIM no Brasil em 2020

Evaluation of data on deaths due to COVID-19 from the databases of the Civil Registry (RC-Arpen), SIVEP-Gripe, and SIM in Brazil in 2020

Evaluación de datos de muertes por COVID-19 en las bases de datos del Registro Civil-Arpen, SIVEP-Gripe y SIM en Brasil en 2020

Ricardo Guedes ¹
Gilson José Dutra ¹
Cecilia Machado ¹
Marina Aguiar Palma ¹

doi: 10.1590/0102-311XPT077222

Resumo

Este estudo compara os registros de óbitos por COVID-19 em 2020 para todo o território nacional. Utilizamos três bases distintas: Registro Civil (RC-Arpen), Sistema de Informação sobre Mortalidade (SIM) e Sistema de Informação da Vigilância Epidemiológica da Gripe (SIVEP-Gripe). Há discordâncias entre os números de mortes por COVID-19 divulgados pelas diversas bases e essas diferenças variam em cada Unidade da Federação. A base do RC-Arpen é atualizada mais rapidamente que as outras duas bases do Departamento de Informática do Sistema Único de Saúde – DATASUS (SIM e SIVEP-Gripe), além de ser mais indicada para monitoramentos e pesquisas que abrangem períodos mais recentes. Apesar da atualização mais lenta, as bases do DATASUS apresentam números geograficamente similares e divulgam dados mais detalhados sobre as mortes. Esse detalhamento das informações torna as bases do DATASUS mais adequadas para pesquisas que demandam mais informações sobre o paciente e o tratamento.

COVID-19; Registro Civil; Sistemas de Informação; Síndrome Respiratória Aguda Grave

Correspondência

R. Guedes
FGV EPGE Escola Brasileira de Economia e Finanças.
Praia de Botafogo 190, 4º andar, sala 414, Rio de Janeiro, RJ
22250-900, Brasil.
ricardobg@gmail.com

¹ FGV EPGE Escola Brasileira de Economia e Finanças, Rio de Janeiro, Brasil.



Introdução

O Brasil é o segundo país com maior número de mortes por COVID-19 em 2020¹. A pandemia se espalhou por todo o território nacional em menos de um ano. De acordo com o Sistema de Informação sobre Mortalidade (SIM), em 2020 ocorreram 210 mil mortes a mais do que em 2019, um crescimento anual acima de 15%. Normalmente, esse crescimento inesperado da mortalidade dificulta uma rápida organização e contabilização dessa informação, mas isso não ocorreu em 2021. A divulgação dos dados e das estatísticas foi mais célere que em outros anos. Devido às dimensões continentais e à infraestrutura precária em diversas localidades, contabilizar todas essas fatalidades demanda um complexo trabalho de compilação e análise de dados. No entanto, a gravidade da situação exigiu pesquisas e análises urgentes, que utilizassem a informação disponível no momento. Passado mais de um ano do primeiro caso de COVID-19 no Brasil, podemos melhor avaliar as vantagens e desvantagens apresentadas pelos bancos de dados públicos com informações sobre a pandemia.

Diversas análises^{2,3,4} optaram por utilizar os dados do Registro Civil (RC-Arpen) divulgados pela Associação Nacional dos Registradores de Pessoas Naturais (Arpen Brasil)⁵. Durante o ano de 2020, essas informações foram atualizadas diariamente em todo o território nacional. Segundo a Arpen Brasil, os óbitos obedecem a um prazo legal de até 14 dias a partir da data de falecimento para serem computados, mas essa base é preliminar e pode haver registros posteriores a esse prazo. A metodologia utilizada considera o relato tal como aparece na Declaração de Óbito (DO), incluindo até casos suspeitos no momento do falecimento⁶. Ademais, as informações do Portal de Transparência do Registro Civil (<https://transparencia.registrocivil.org.br>) excluem as causas externas de morte, ou seja, as causas não naturais. Outras duas novas bases foram divulgadas pelo Departamento de Informática do SUS (DATASUS) em 2021, apresentando com detalhes os dados das internações por síndrome respiratória aguda grave (SRAG) e do SIM. Os dados de SRAG foram coletados por meio do Sistema de Informação da Vigilância Epidemiológica da Gripe (SIVEP-Gripe), mas apresentaram uma defasagem de alguns meses na atualização. Os dados do SIM são analisados pelo Ministério da Saúde, mas ainda são preliminares. O SIM, assim como o RC-Arpen, utiliza a DO como fonte primária, porém essas informações passam por diferentes processos antes de serem utilizadas pela Arpen Brasil e pelo SIVEP-Gripe. Além disso, uma parcela da coleta de dados feita pelos cartórios não é informatizada e nem todas as secretarias de saúde do município têm processos dinâmicos e eficientes⁷.

Este artigo pretende comparar três fontes de informação distintas e mapear as diferenças entre as estatísticas divulgadas para todo o território nacional. Compilamos também uma sequência de versões da base de dados de SRAG (SIVEP-Gripe) extraídas em diferentes datas para avaliar a atualização da base de dados. Considerando apenas o ano de 2020, o trabalho aponta as principais características da base do RC-Arpen, comparando essas informações com as divulgadas por dois sistemas do DATASUS: o SIM e o SIVEP-Gripe. O objetivo é sinalizar as qualidades e as deficiências de cada fonte de dados e promover o uso consciente dessa informação para pesquisas futuras.

Métodos

Contabilizamos neste trabalho as mortes por COVID-19 no ano de 2020 em todo o território nacional, usando três bases de dados distintas: RC-Arpen, SRAG (SIVEP-Gripe) e SIM. Essas três fontes têm muitas semelhanças, mas também algumas diferenças relativas à classificação da causa de morte e à identificação do local do óbito. O objetivo foi tornar as bases de dados comparáveis, para que as divergências metodológicas tivessem o menor impacto possível na comparação.

O RC-Arpen documenta todas as mortes em território nacional usando como fonte as DOs registradas nos cartórios. As DOs registram informações do falecido como idade, sexo, local, data e causas do óbito. A Arpen Brasil organiza as informações provenientes de todos os cartórios brasileiros na Central de Informações do Registro Civil (CRC). A base de dados inclui a data e o local da morte, organizado em três variáveis: local de morte (hospital, casa, público ou outros), Unidade da Federação (UF) e cidade do cartório do registro. A base também abrange variáveis de sexo, grupo etário e causas de morte. A Arpen Brasil classificou a causa da morte como COVID-19 para todas as DOs com menção de COVID-19, coronavírus e novo coronavírus. Nesse contexto, os dados utilizados de óbito por

COVID-19 foram agrupados segundo a UF de ocorrência e consideramos apenas as mortes ocorridas no ano de 2020 na base extraída em 5 de maio de 2021. Ressalta-se que essa metodologia inclui casos suspeitos sem confirmação e casos em que a COVID-19 foi mencionada na DO, mas também outras doenças relacionadas à COVID-19, como cardiopatias ou acidentes vasculares cerebrais, foram apontadas como causas básicas do óbito.

O Ministério da Saúde faz o controle da SRAG no Brasil desde 2009, quando ocorreu a pandemia de influenza. O SIVEP-Gripe divulga a base de casos de SRAG ⁸, que, baseada nos diagnósticos laboratoriais, clínicos e de imagens (tomografias), classifica os pacientes infectados por COVID-19, dando preferência aos resultados laboratoriais. Consideramos neste artigo o local de internação do paciente para determinar a localização do óbito. Essa base de dados é atualizada semanalmente e apresenta informações de sexo, idade, local de óbito, local de residência, notificação e internação do paciente, assim como detalhes sobre atendimento, diagnóstico e evolução dos casos. Verificamos a informação de classificação final dos casos para encontrar as internações de SRAG diagnosticadas com COVID-19; a informação de evolução dos pacientes para identificar se o quadro clínico evoluiu para cura ou óbito; e a data de evolução para determinar a data do falecimento. Todas as bases incluem apenas as mortes ocorridas em 2020 por COVID-19, mas sofrem modificações significativas ao longo do tempo. Por essa razão, extraímos versões mensais para avaliar a defasagem e o ritmo de atualização dos dados e comparar a evolução do percentual de dados faltantes ao longo do ano. A primeira versão foi extraída em 1^o de fevereiro, mas optamos por usar a versão do dia 3 de maio em todas as comparações, pois essa data está mais próxima das datas de atualização das bases do RC-Arpen e do SIM.

O SIM considera os dados da DO para contabilizar as mortes por COVID-19. Entretanto, diferentemente do RC-Arpen, essas declarações de óbito são classificadas e revisadas por técnicos do Ministério da Saúde, o que torna a fonte mais confiável. Essa base apresenta todos os códigos da 10^a revisão da Classificação Internacional de Doenças (CID-10) informados na DO. Para identificar as mortes causadas por COVID-19, consideramos somente os casos com o código CID-10 para infecção por coronavírus (B34.2) registrado na causa básica do óbito. Utilizamos uma versão preliminar da base divulgada pelo DATASUS em 12 de maio, que inclui informações sobre local da morte, local de residência, sexo e idade.

A confrontação das bases de dados necessita de alguns ajustes com o objetivo de tornar seus respectivos números o mais comparáveis possível. Trabalhamos apenas com os óbitos ocorridos em 2020 no território nacional para que a análise considerasse o mesmo período e o mesmo limite geográfico para as três fontes de dados. Todas as bases de dados foram tratadas e comparadas no software R Studio (<https://rstudio.com/>). Marinho ⁹ fez uma análise detalhada das informações divulgadas nessas três bases e suas diferenças metodológicas. Este artigo tem outro foco, pois é baseado nos resultados estatísticos das mortes em 2020 e nas comparações dos registros de óbitos na pandemia ao longo do ano e entre diferentes localidades. Estamos considerando apenas o ano de 2020 para reduzir os efeitos de atualizações dos números nos meses mais recentes. O RC-Arpen apresenta a informação da cidade do cartório em que o óbito foi registrado, que usualmente é a mesma cidade do falecimento. A base de dados do SIM disponibiliza a informação da cidade de residência e cidade de ocorrência do óbito. Dessa forma, optou-se por utilizar a informação referente aos locais de ocorrência, que representam o local da morte. Já a base de SRAG apresenta três locais distintos, mas consideramos que o local da internação deve ser o mesmo do local da morte e, por isso, optamos por fazer nossas análises usando essa informação como referência.

Realizamos uma comparação do número de mortes entre todas as bases ao longo do tempo, com dados diários. Devido à sazonalidade semanal dos registros, decidimos usar uma média móvel de sete dias a fim de excluir esse efeito e suavizar a curva de óbitos. Ademais, agregamos os dados de 2020 e comparamos os números de mortes totais por localidade de ocorrência. Ao comparar as estatísticas por cidades, observamos que há diferenças significativas entre a quantidade de mortes por local de internação e por residência dos pacientes, mas essas diferenças ficam irrelevantes quando agregamos os dados por estado. Ressaltamos que este estudo escolheu apresentar análises geográficas por UF, que são partições maiores do território nacional, o que minimiza as divergências de informações sobre o local da morte nas bases de dados.

Resultados

A análise temporal diária dos números agregados apresenta diferenças significativas entre as três fontes de dados durante alguns períodos do ano de 2020. A Figura 1 mostra as médias móveis de mortes por COVID-19 dos últimos sete dias das bases de dados do RC-Arpen, SIM e SIVEP-Gripe no ano de 2020, extraídas no mês de maio de 2021. No primeiro pico de contágio, em maio, o RC-Arpen apresentou uma média móvel de 200 mortes diárias a menos que os números registrados pelos dados de SRAG (SIVEP-Gripe) e SIM. Entretanto, essa diferença foi diminuindo ao longo dos meses de junho e julho, e o RC-Arpen passou a contabilizar entre 20 e 60 mortes a mais que as outras duas séries. No mês de dezembro, a diferença de mortes diárias das médias móveis entre o RC-Arpen e a base de SRAG chega a quase 100 óbitos.

A Figura 2 apresenta o percentual de casos de COVID-19 com observações faltantes da variável de classificação final na base de SRAG. Esse campo da base de dados identifica se o paciente tem alguma síndrome respiratória aguda grave, como a COVID-19. A melhora temporal nos dados da SRAG pode ser vista na Figura 2, que exhibe os dados de quatro bases extraídas nos primeiros meses de 2021. A comparação indica que em todos os meses de 2020 houve atualizações e que elas reduzem o número de observações faltantes.

Mesmo utilizando uma base de dados atualizada em 3 de maio, observamos na Figura 2 que, nos últimos cinco meses do ano, mais de 5% das observações apresentam preenchimento incompleto, sendo que esse percentual alcança 9,4% em dezembro na variável de “classificação final dos casos”. Adicionalmente, nota-se que o percentual de dados faltantes é ainda maior para a variável “evolução dos casos”, chegando a mais de 10% nos últimos cinco meses do ano e a 17% no mês de dezembro de 2020. Essa falta de informação pode gerar subnotificação no fim do ano para a base de SRAG. Nossa análise aponta uma atualização relevante da base de dados nos meses seguintes às notificações da doença, em especial nos seis primeiros meses subsequentes.

Figura 1

Médias móveis (sete dias) de mortes por COVID-19 no Brasil registradas nas bases do Registro Civil da Associação Nacional dos Registradores de Pessoas Naturais (RC-Arpen), de síndrome respiratória aguda grave (SRAG) do Sistema de Informação da Vigilância Epidemiológica da Gripe (SIVEP-Gripe) e do Sistema de Informação sobre Mortalidade (SIM). Brasil, 2020.

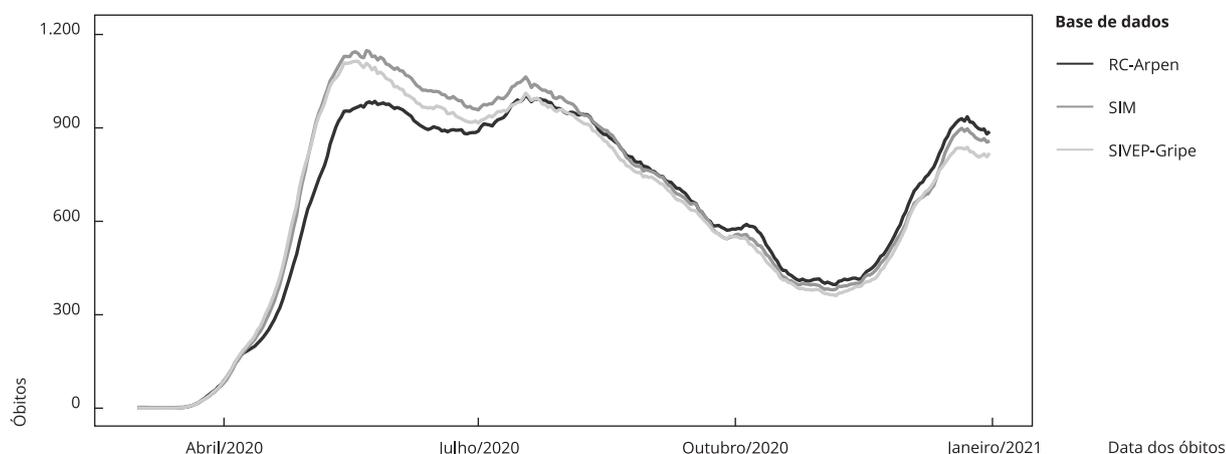
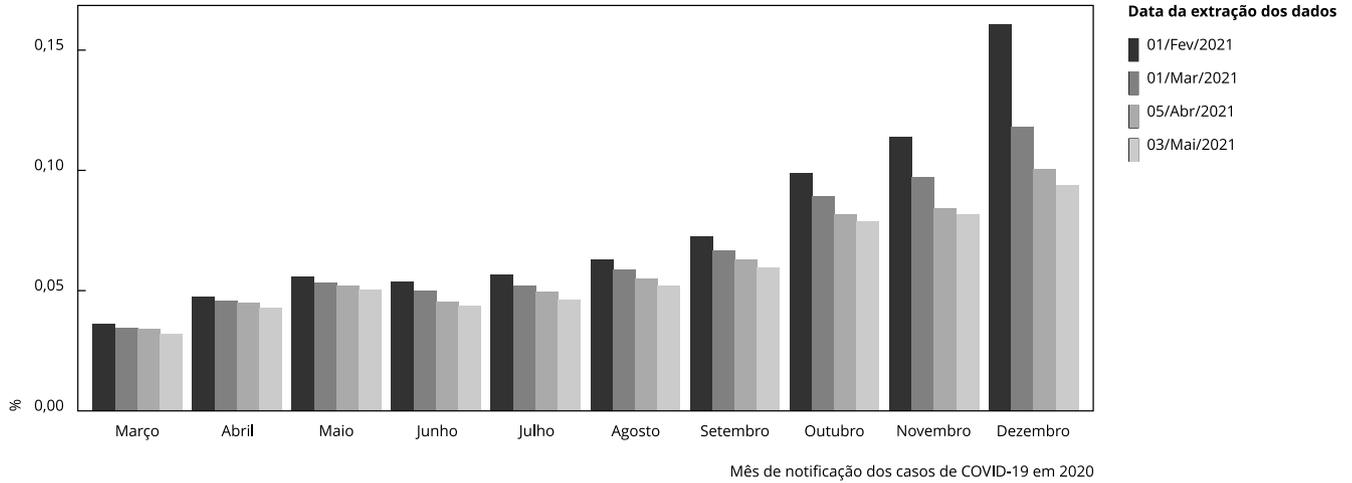


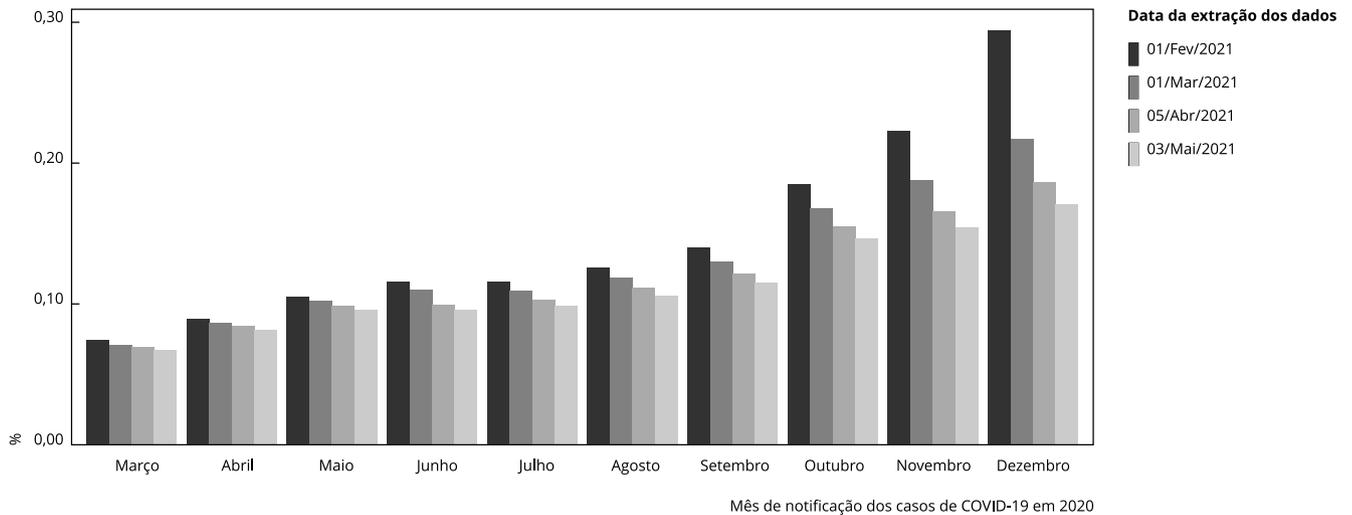
Figura 2

Percentual de casos de síndrome respiratória aguda grave (SRAG) sem informação de diagnóstico e sem informação de evolução dos pacientes na base de dados do Sistema de Informação da Vigilância Epidemiológica da Gripe (SIVEP-Gripe). Brasil, março a dezembro de 2020.

2a) Dados faltantes na classificação final dos casos



2b) Dados faltantes na evolução dos casos



Nota: casos de SRAG do SIVEP-Gripe extraídos nas datas de 01/Fev/2021, 01/Mar/2021, 05/Abr/2021 e 03/Mai/2021.

A Tabela 1 expõe os percentuais de dados faltantes para o mês de dezembro de 2020 por UF na base de SRAG do SIVEP-Gripe. Alguns estados do Norte e Nordeste – como Acre, Alagoas, Ceará, Maranhão, Pará, Pernambuco, Rondônia e Tocantins – apresentaram um percentual de dados faltantes acima de 10% na base extraída em 3 de maio. Por outro lado, no Estado de Roraima, não há dados faltantes para dezembro de 2020, assim como em outros meses do ano, no entanto, é o estado com menor número de casos e menor população de todo território nacional. A maioria dos estados do Sudeste, com exceção de São Paulo, também têm proporções altas de dados ausentes.

Análise das UFs

A Figura 3 apresenta algumas discrepâncias por UF entre as bases de dados do RC-Arpen, do SIVEP-Gripe e do SIM. As mortes assinaladas no RC-Arpen somam menos de 90% daquelas contabilizadas pela base de SRAG em nove estados, todos eles localizados nas regiões Norte e Nordeste. A Região Norte registrou em cartório apenas 75% das mortes contabilizadas pela base de SRAG, enquanto a Região Nordeste registrou 88%. Por outro lado, na Figura 3a, observamos que todos os estados do Sul, Sudeste e Centro-oeste apresentaram número semelhante ou superior de mortes no RC-Arpen, em especial os estados do Espírito Santo e Mato Grosso, que registraram um aumento de quase 50% dos óbitos, se comparados aos dados de SRAG.

Na Figura 3b, os números do RC-Arpen e do SIM exibiram diferenças expressivas na análise por UF. Os estados mais ao norte do país registraram menos mortes no RC-Arpen, o que também foi observado por Costa et al.¹⁰, utilizando dados do SIM para os anos de 2015 e 2016, enquanto os estados ao sul apontaram menos mortes no SIM. Na Figura 3c, a base de dados do SIVEP-Gripe apresentou números parecidos aos do SIM na maior parte do território nacional. Porém, houve mais registros de óbitos na base de SRAG do que na do SIM em alguns estados, apesar de a primeira base contabilizar apenas óbitos decorrentes de internações e a segunda incluir todas as mortes relacionadas à COVID-19.

Tabela 1

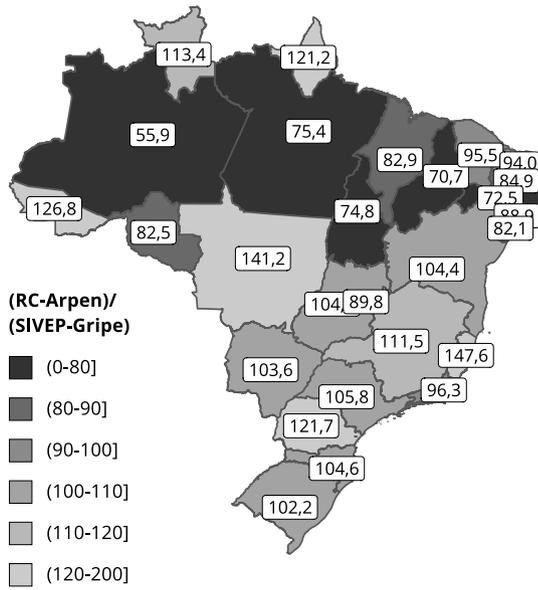
Percentual de dados faltantes da base de síndrome respiratória aguda grave (SRAG) do Sistema de Informação da Vigilância Epidemiológica da Gripe (SIVEP-Gripe) em dezembro de 2020 por Unidade da Federação (UF), Brasil.

UF	Data da extração dos dados			
	01/Fev/2021	01/Mar/2021	05/Abr/2021	03/Mai/2021
Acre	61,5	25,5	24,4	23,9
Alagoas	33,1	18,5	10,8	10,7
Amapá	0,7	0,3	0,0	0,0
Amazonas	9,7	4,1	2,2	1,4
Bahia	9,9	6,2	3,3	3,2
Ceará	31,7	26,8	25,8	25,1
Distrito Federal	10,5	9,6	8,1	6,8
Espírito Santo	24,2	19,8	19,0	18,6
Goiás	15,6	11,4	9,3	8,9
Maranhão	13,3	12,6	12,9	12,0
Mato Grosso	23,3	18,6	18,3	18,2
Mato Grosso do Sul	5,8	4,7	4,0	2,6
Minas Gerais	22,1	14,2	11,7	10,8
Pará	17,4	14,0	13,2	12,6
Paraíba	9,5	7,8	7,1	7,0
Paraná	16,7	10,5	9,9	9,1
Pernambuco	33,7	30,4	28,6	27,4
Piauí	14,4	10,5	9,6	8,6
Rio de Janeiro	26,4	19,2	15,6	13,4
Rio Grande do Norte	11,5	8,9	7,2	6,1
Rio Grande do Sul	2,5	1,7	1,4	1,3
Rondônia	31,5	23,9	15,0	14,9
Roraima	0,0	0,0	0,0	0,0
Santa Catarina	3,5	2,5	2,1	2,0
São Paulo	13,1	10,0	8,3	8,1
Sergipe	0,8	0,7	0,6	0,6
Tocantins	22,3	20,9	19,6	19,5
Brasil	16,1	11,8	10,0	9,4

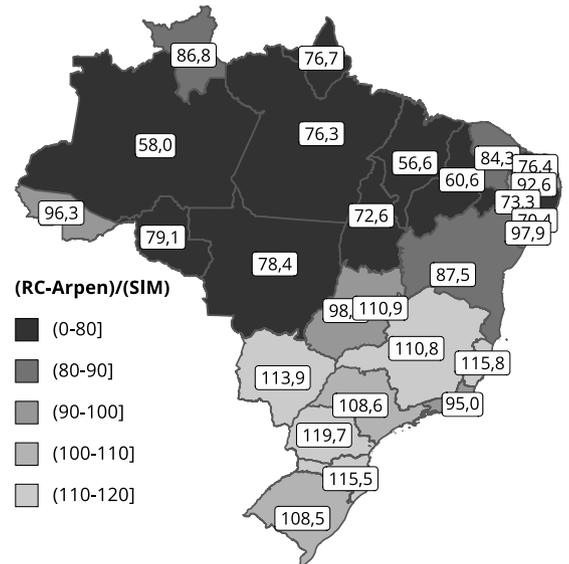
Figura 3

Mortes por COVID-19 nas bases do Registro Civil da Associação Nacional dos Registradores de Pessoas Naturais (RC-Arpen), de síndrome respiratória aguda grave (SRAG) do Sistema de Informação da Vigilância Epidemiológica da Gripe (SIVEP-Gripe) e do Sistema de Informação sobre Mortalidade (SIM) por Unidade da Federação. Brasil, 2020.

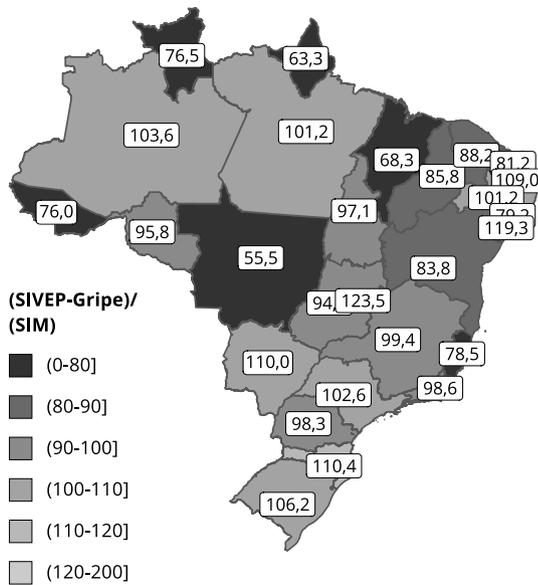
3a) Comparação entre as bases RC-Arpen e SIVEP-Gripe



3b) Comparação entre as bases RC-Arpen e SIM



3c) Comparação entre as bases SIVEP-Gripe e SIM



Fonte: dados das bases SRAG (SIVEP-Gripe), RC-Arpen e SIM.

Destacamos algumas UFs que apresentaram séries temporais de mortes por COVID-19 com comportamentos distintos ao longo do ano de 2020. No Brasil, houve várias ondas da pandemia e encontramos diversos padrões de evolução dos casos, que variam temporal e geograficamente ^{11,12,13,14,15,16}, além da heterogeneidade das políticas locais para o combate à disseminação da doença ^{17,18}. Seleccionamos alguns exemplos representativos de estados que apresentaram muita discrepância (Amazonas e Pará), alta compatibilidade (Minas Gerais, São Paulo, Rio de Janeiro e Rio Grande do Sul) ou um padrão variável ao longo do tempo (Ceará e Pernambuco) em relação a compatibilidade entre as bases de dados.

A Figura 4 mostra a média móvel de sete dias das mortes registradas no RC-Arpen (preto), na base de SRAG (cinza claro) e a base do SIM (cinza) para os estados do Amazonas, Ceará, Minas Gerais, Pará, Pernambuco, Rio de Janeiro, Rio Grande do Sul e São Paulo. Estados como Pará e Amazonas registraram uma quantidade de óbitos em cartório abaixo da divulgada pelo SIVEP-Gripe e SIM ao longo de todo o ano. Amazonas é o caso mais preocupante, pois apresentou um percentual entre 50% e 80% durante todo o ano de 2020, enquanto o Pará demonstrou um baixo número de mortes no cartório até junho e depois de novembro de 2020. O Estado de Pernambuco mostrou um comportamento dinâmico ao longo do tempo, iniciando a pandemia com registros no RC-Arpen menores que a meta-de dos casos fatais contabilizados nas bases SRAG e SIM, mas terminando o ano com notificações no RC-Arpen acima dos registros nos dados de SRAG, porém ainda abaixo dos números divulgados pelo SIM. Alguns estados mostraram uma alta compatibilidade entre as três bases de dados, como é o caso do Rio Grande do Sul. Minas Gerais exibiu alta compatibilidade entre os dados do DATASUS, mas uma ligeira sobrenotificação do RC-Arpen. Esse padrão com número de óbitos registrados na base da Arpen Brasil acima dos dados de SRAG e do SIM se repete em alguns estados, como em São Paulo após junho de 2020. Apesar dos casos de sobrenotificações do RC-Arpen, os três estados mais populosos, São Paulo, Rio de Janeiro e Minas Gerais, têm demonstrado alta compatibilidade entre as bases de dados e uma dinâmica muito semelhante ao longo do ano.

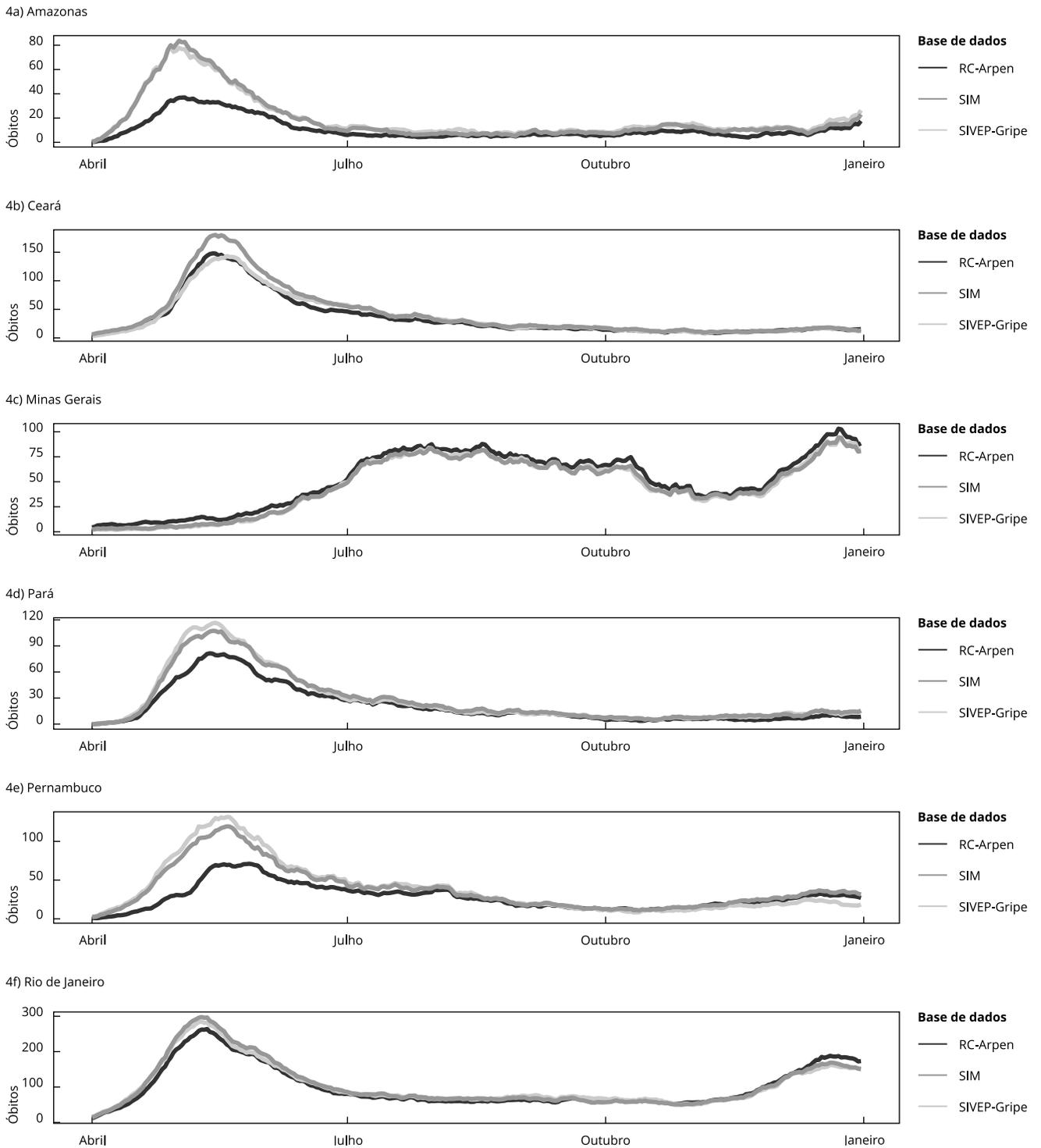
Discussão

A base de dados de SRAG do SIVEP-Gripe passa por um escrutínio mais detalhado, pois arquiva diversas etapas do atendimento médico do paciente antes do falecimento e descarta casos suspeitos de COVID-19. Por essa razão, é possível ter mais precisão nos diagnósticos e nas causas das mortes. Contudo, a atualização desses dados é lenta, o que gera uma subnotificação nos períodos mais recentes. Mesmo considerando a base atualizada em maio de 2021, o percentual de dados faltantes nos últimos meses do ano é ainda maior que o registrado no começo do ano. Avaliando uma sequência mensal de atualizações da base do SIVEP-Gripe, observamos que os primeiros meses do ano registraram poucas variações no percentual de valores faltantes. Por outro lado, o mês de dezembro registrou reduções significativas na proporção de dados ausentes, o que caracteriza uma instabilidade na divulgação de mortes nos meses subsequentes. A partir dessas informações, concluímos que são necessários ao menos seis meses para uma estabilização na proporção dessa informação ausente. Entretanto, todo o histórico da base segue sendo atualizado constantemente e ocorreram pequenas revisões em todos os meses de 2020. Na Tabela 1, verifica-se também uma heterogeneidade geográfica no preenchimento da base de SRAG pelo SIVEP-Gripe, o que pode gerar diferenças na contabilidade de casos e óbitos das bases mais recentes. O alto percentual de dados faltantes nos estados do Rio de Janeiro e Minas Gerais indica que a ausência de informação nem sempre está correlacionada às condições socioeconômicas do local. Outra hipótese seria que a quantidade de casos poderia ter algum impacto na gestão e divulgação dos dados pela SIVEP-Gripe, mas o alto percentual de preenchimento das informações nos estados de São Paulo e Amazonas nos impede de afirmar que essa tese é verdadeira. Ademais, notamos que o Estado de São Paulo adotou uma base unificada de óbitos ¹⁹.

Os números extraídos da base do SIM consideram somente os casos de 2020 em que a COVID-19 foi identificada como causa básica na DO. Apesar de seguir uma fonte oficial, ainda existe a possibilidade de a causa básica não ter sido verificada por testes laboratoriais até o momento do registro, ou ter sido diagnosticada com base em evidências fracas. Em razão das bases do SIM e de SRAG serem mais detalhadas e apresentarem uma coerência entre si, recomenda-se usar essas fontes para estudos

Figura 4

Média móvel (sete dias) de mortes por COVID-19 em 2020 nas bases do Registro Civil da Associação Nacional dos Registradores de Pessoas Naturais (RC-Arpen), de síndrome respiratória aguda grave (SRAG) do Sistema de Informação da Vigilância Epidemiológica da Gripe (SIVEP-Gripe) e do Sistema de Informação sobre Mortalidade (SIM) para os estados do Amazonas, Ceará, Minas Gerais, Pará, Pernambuco, Rio de Janeiro, Rio Grande do Sul e São Paulo, Brasil.

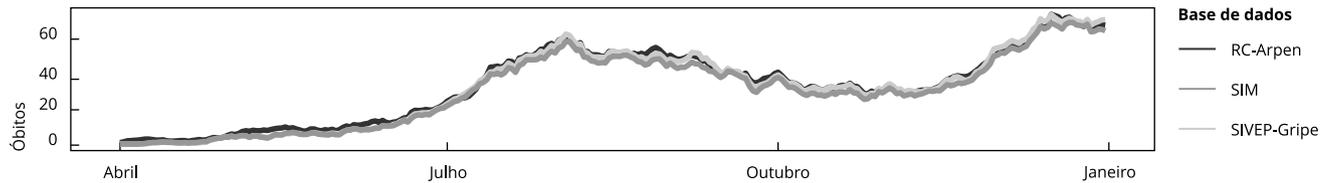


(continua)

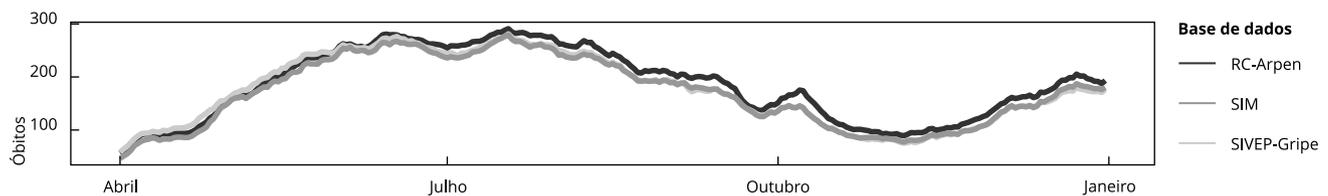
Figura 4

Média móvel (sete dias) de mortes por COVID-19 em 2020 nas bases do Registro Civil da Associação Nacional dos Registradores de Pessoas Naturais (RC-Arpen), de síndrome respiratória aguda grave (SRAG) do Sistema de Informação da Vigilância Epidemiológica da Gripe (SIVEP-Gripe) e do Sistema de Informação sobre Mortalidade (SIM) para os estados do Amazonas, Ceará, Minas Gerais, Pará, Pernambuco, Rio de Janeiro, Rio Grande do Sul e São Paulo, Brasil.

4g) Rio Grande do Sul



4h) São Paulo



que demandem variáveis relativas à morbidade prévia dos pacientes ou a informações socioeconômicas e que não necessitem considerar as notificações mais recentes, isto é, óbitos e casos ocorridos nos últimos meses.

Os dados de cartório (RC-Arpen) contabilizados pela Arpen Brasil têm um critério mais amplo, pois consideram como mortes por COVID-19 todas as DOs que fazem menção à doença, incluindo os casos suspeitos. No início da pandemia, houve escassez de testes e em muitos casos a declaração de óbito foi feita sem a realização de testes laboratoriais, em especial na população vulnerável²⁰. Logo, há espaço para que essa metodologia superestime ou subestime o número de mortes, devido à imprecisão do diagnóstico, embora não seja necessariamente um erro de contagem. Após exames e diagnósticos mais precisos, existe a possibilidade de alterar a certidão de óbito por meio de uma averbação. Entretanto, tal procedimento não é automático e deve partir da pessoa interessada, que precisa fornecer uma comprovação por meio de documentos emitidos pelos órgãos competentes²¹.

Portanto, estados que tiveram número de mortes registradas em cartório muito acima da base do SIVEP-Gripe e do SIM podem estar considerando muitos casos suspeitos, não confirmados. Outra possível explicação seria a inclusão de mortes que não foram diagnosticadas como SRAG na base de dados do RC-Arpen. Entretanto, em vários desses estados há também mais notificações na base do RC-Arpen do que na base do SIM, que abrange esses casos não diagnosticados como síndrome respiratória. Por último, também haveria a possibilidade de esses pacientes não terem buscado atendimento médico. Entretanto, a maioria dos estados em que a base do SIVEP-Gripe identificou menos óbitos do que a base de RC-Arpen apresenta desenvolvimento socioeconômico relativamente mais alto. De fato, a razão de óbitos (RC-Arpen/SRAG) evidenciou uma correlação positiva de 0,41 com o Índice de Desenvolvimento Humano (IDH) das UFs²². De maneira consistente, o RC-Arpen exibiu número maior de óbitos que o SIM para estados com IDH elevado, pois a correlação entre a razão desses registros (RC-Arpen/SIM) e o IDH estadual é positiva (0,73). Portanto, a tese de que estados mais ricos tendem a ter melhores dados advindos de cartórios é apoiada pelas evidências aqui mostradas.

Diógenes et al.²³ observaram pequenas diferenças espaciais entre os dados do RC-Arpen e do SIM para o ano de 2010, entretanto, as diferenças calculadas neste artigo são consideravelmente maiores em todas as regiões. Em nosso estudo, observamos um registro bem maior de mortes na base do SIM para as regiões Norte e Nordeste em comparação com as diferenças calculadas por Diógenes et al. Ademais, os dados publicados no artigo citado apresentaram números quase iguais de mortes nas duas bases para as regiões Sul e Sudeste, em 2010, enquanto nosso trabalho revelou uma contagem maior de óbitos para a base do RC-Arpen. Esses resultados indicam que algumas divergências nos dados observadas no passado se intensificaram durante a pandemia.

O painel InfoGripe²⁴ é organizado pela Fundação Oswaldo Cruz (Fiocruz) e monitora os casos de SRAG notificados no SIVEP-Gripe. Segundo Alves et al.²⁵, que analisaram os dados do RC-Arpen e SRAG de janeiro a julho de 2020, existe uma subnotificação dos dados do sistema InfoGripe quando comparados aos do RC-Arpen no Estado de Minas Gerais, sexto maior IDH do país. Os autores verificaram que as notificações do RC-Arpen de óbitos por COVID-19 representaram mais que o dobro dos dados registrados no InfoGripe para o período analisado.

Em razão dos pontos já citados, os estados que apresentam números de óbitos registrados em cartório sistematicamente abaixo das bases de dados do Sistema Único de Saúde (SUS) são os que requerem maior cautela na análise dos dados de mortes por COVID-19. As bases de dados do SUS estão mais bem documentadas e dispõem de uma metodologia mais estrita, deixando pouco espaço para sobrenotificações. A base do SIVEP-Gripe, por conter mais informações sobre os pacientes, permite estudos mais detalhados sobre atendimento hospitalar²⁶; características étnicas^{27,28,29}, sociodemográficas, epidemiológicas e clínicas^{30,31}; e avaliação da transmissão da epidemia^{32,33}. A comparação com a base de SRAG é uma evidência de que há subcontagem de mortes do RC-Arpen para esses estados. Excluindo os estados que apresentam uma subnotificação sistemática em relação aos dados do DATASUS, a maioria deles no Norte e Nordeste, a base de dados da Arpen Brasil pode ser útil para estudos que necessitem de dados recentes atualizados e para o monitoramento da doença pelas autoridades de saúde, que precisam responder com medidas de controle em tempo hábil^{34,35}.

Outra comparação desses dados, feita por Costa et al.³⁶ para os dados do RC, do Sistema de Informações sobre Nascidos Vivos (SINASC) e do SIM, mostra que há uma alta taxa de pareamento entre os bancos de dados, mas aponta um percentual maior de registros não pareados no Nordeste e, principalmente, no Norte. Esse resultado é coerente com as nossas análises que apontaram uma discrepância de óbitos entre as bases do RC-Arpen e do SIM para as mesmas UFs. Essas bases também foram usadas por Silva et al.³⁷ e Bastos et al.³⁸ para identificar um aumento de casos de SRAG acima do esperado para o Brasil. A base de dados do SIM de 2020 serviu de fonte para estudar a importância de priorizar áreas socialmente vulneráveis durante a vacinação contra COVID-19³⁹. Outros trabalhos^{40,41,42,43,44,45} avaliaram o excesso de mortes nas capitais brasileiras, usando os óbitos por COVID-19 contabilizados nas bases do RC-Arpen e do SIM. Entretanto, alguns estudos mais recentes já utilizaram as bases do RC-Arpen e SIVEP-Gripe⁴⁶, encontrando diferenças entre os números divulgados para o Município de Manaus (Amazonas).

Neste artigo, avaliamos como os dados da Arpen Brasil (RC-Arpen) registram números diferentes de mortes por UF quando comparados às bases do SIM e de SRAG. As divergências são especialmente notáveis em alguns estados do Norte e Nordeste do país, onde observamos uma subnotificação de dados do RC-Arpen. Algumas dessas diferenças já haviam sido apontadas em outros anos, mas durante a pandemia identificamos que houve um aumento dessas divergências. Os resultados indicam que a base de SRAG necessita de pelo menos seis meses para que a variável de evolução do caso seja preenchida e para alcançar um percentual de dados faltantes em torno de 5% para todo o país. Ainda assim, tal percentual não é homogêneo e alguns estados registram mais de 20% de casos com informações faltantes. Portanto, para a utilização de estatísticas de óbito mais recentes, o RC-Arpen é mais indicado pela sua rápida atualização. Por outro lado, as bases de SRAG (SIVEP-Gripe) e do SIM, apesar de necessitarem de mais tempo para serem preenchidas e publicadas, são indicadas para pesquisas que precisem de um detalhamento maior dos casos de COVID-19, pois apresentam mais informações sobre os pacientes e os óbitos.

Colaboradores

R. Guedes contribuiu com o tratamento dos dados, análise dos resultados, redação e revisão do artigo; e aprovou a versão final do manuscrito. G. J. Dutra contribuiu com o tratamento dos dados, análise dos resultados, redação e revisão do artigo; e aprovou a versão final do manuscrito. C. Machado contribuiu com a concepção da pesquisa, análise dos resultados e revisão do artigo; e aprovou a versão final do manuscrito. M. A. Palma contribuiu com a análise dos resultados e revisão do artigo; e aprovou a versão final do manuscrito.

Informações adicionais

ORCID: Ricardo Guedes (0000-0003-2343-9967); Gilson José Dutra (0000-0002-3373-7485); Cecilia Machado (0000-0001-9400-4069); Marina Aguiar Palma (0000-0001-8134-8946).

Agradecimentos

Este trabalho foi realizado com o apoio da Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ). Agradecemos à Fundação Getúlio Vargas (FGV) pelo apoio e a Soraya Adiva Roman Eyzaguirre, Carlos Andres Charris Vizcaino e Valdemar Rodrigues de Pinho Neto pela ajuda e pelos comentários.

Referências

1. World Health Organization. Coronavirus disease (COVID-2019). Situation reports. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports> (acessado em 25/Abr/2022).
2. Brant LCC, Nascimento BR, Teixeira RA, Lopes MACQ, Malta DC, Oliveira GMM, et al. Excess of cardiovascular deaths during the COVID-19 pandemic in Brazilian capital cities. *Heart* 2020; 106:1898-905.
3. Sanchez M, Moura E, Moreira J, Lima R, Barreto I, Pereira C, et al. Mortalidade por COVID-19 no Brasil: uma análise do Registro Civil de óbitos de janeiro de 2020 a fevereiro de 2021. *SciELO Preprints* 2021; 22 mar. <https://preprints.scielo.org/index.php/scielo/preprint/view/2012>.
4. Pereira CCA, Marcelino AB, Machado CJ. O registro civil nos quatro primeiros meses da pandemia de Covid-19 no Brasil: estudo comparativo entre óbitos por mês de ocorrência e de registro. *Rev Bras Estud Popul* 2020; 37:e0127.
5. Oliveira M. Brazil-civil-registry-data. <https://github.com/capyvara> (acessado em 05/Mai/2021).
6. França EB, Ishitani LH, Teixeira RA, Abreu DMX, Corrêa PRL, Marinho F, et al. Óbitos por COVID-19 no Brasil: quantos e quais estamos identificando? *Rev Bras Epidemiol* 2020; 23:e200053.
7. Lima EEC, Gonzaga MR, Freire FHMA, Queiroz BL. Fontes alternativas de informação de óbitos no Brasil: contexto de pandemia Covid-19. Ottawa: Centre of Excellence for Civil Registration and Vital Statistics Systems; 2021.
8. Secretaria de Vigilância em Saúde, Ministério da Saúde. Guia de vigilância epidemiológica: emergência de saúde pública de importância nacional pela doença pelo coronavírus 2019 – COVID-19. Brasília: Ministério da Saúde; 2022.
9. Marinho MF. Semelhanças e diferenças na interpretação dos dados sobre SG, SRAG e COVID-19: SIM, SIVEP-Gripe e Cartórios de Registro Civil. In: Santos AO, Lopes LT, organizadores. Coleção Covid-19. v. 2: Planejamento e gestão. Brasília: Conselho Nacional de Secretários de Saúde; 2021. p. 112-25.
10. Costa LFL, Montenegro MMS, Rabello Neto DL, Oliveira ATR, Trindade JEO, Adair T, et al. Estimating completeness of national and sub-national death reporting in Brazil: application of record linkage methods. *Popul Health Metr* 2020; 18:22.
11. Candido DS, Claro IM, Jesus JG, Souza WM, Moreira FRR, Dellicour S, et al. Evolution and epidemic spread of SARS-CoV-2 in Brazil. *Science* 2020; 369:1255-60.
12. Castro MC, Kim S, Barberia L, Ribeiro AF, Gurzenda S, Ribeiro KB, et al. Spatiotemporal pattern of COVID-19 spread in Brazil. *Science* 2021; 372:821-6.

13. Hallal PC, Hartwig FP, Horta BL, Silveira MF, Struchiner CJ, Vidaletti LP, et al. SARS-CoV-2 antibody prevalence in Brazil: results from two successive nationwide serological household surveys. *Lancet Glob Health* 2020; 8:e1390-8.
14. Rocha R, Atun R, Massuda A, Rache B, Spinola P, Nunes L, et al. Effect of socioeconomic inequalities and vulnerabilities on health-system preparedness and response to COVID-19 in Brazil: a comprehensive analysis. *Lancet Glob Health* 2021; 9:e782-92.
15. Mellan TA, Hoeltgebaum HH, Mishra S, Whitaker C, Schnekenberg RP, Gandy A, et al. Report 21 – Estimating COVID-19 cases and reproduction number in Brazil. Londres: Imperial College London; 2020.
16. Costa GS, Cota W, Ferreira SC. Outbreak diversity in epidemic waves propagating through distinct geographical scales. *Phys Rev Res* 2020; 2:043306.
17. Pereira AK, Oliveira MS, Sampaio TS. Heterogeneidades das políticas estaduais de distanciamento social diante da COVID-19: aspectos políticos e técnico-administrativos. *Rev Adm Pública* 2020; 54:678-96.
18. Kerr L, Kendall C, Silva AAM, Aquino EML, Pescarini JM, Almeida RLF, et al. COVID-19 no Nordeste brasileiro: sucessos e limitações nas respostas dos governos dos estados. *Ciênc Saúde Colet* 2020; 25 Suppl 2:4099-120.
19. Waldvogel BC, Ferreira CEC, Camargo ABM, Jordani MS, Ortiz LP. Base unificada de nascimentos e óbitos no Estado de São Paulo: instrumento para aprimorar os indicadores de saúde. *São Paulo Perspect* 2008; 22:5-18.
20. Sousa TCM, Moreira NP, Krieger JE, Rosa ISC, Zamudio MM, Veras MASM, et al. Using RT-PCR testing to assess the effectiveness of outbreak control efforts in São Paulo State, the pandemic's epicenter in Brazil, according to socioeconomic vulnerabilities. *medRxiv* 2020; 4 dez. <https://www.medrxiv.org/content/10.1101/2020.10.29.20221960v3>.
21. Chossani FW. O registro do óbito em tempos de Covid-19. https://arpensp.org.br/index.php/noticia/94347#_ftn2 (acessado em 18/Ago/2022).
22. Atlas do Desenvolvimento Humano no Brasil. IDHM dos estados brasileiros. <http://www.atlasbrasil.org.br/ranking> (acessado em 15/Ago/2022).
23. Diogenes VHD, Pinto Junior EP, Gonzaga MR, Ichihara MYT, Queiroz BL, Lima EEC, et al. Diferenciais na cobertura das bases de dados de óbitos no Brasil em 2010. *OSF Preprints* 2021; 8 jun. <https://osf.io/ve6t7>.
24. Fundação Oswaldo Cruz. InfoGripe. <http://info.gripe.fiocruz.br/> (acessado em 25/Abr/2022).
25. Alves THE, Souza TA, Silva SA, Ramos NA, Oliveira SV. Underreporting of death by COVID-19 in Brazil's second most populous state. *Front Public Health* 2020; 8:578645.
26. Ranzani OT, Bastos LSL, Gelli JGM, Marchesi JF, Baiao F, Hamacher F, et al. Characterisation of the first 250,000 hospital admissions for COVID-19 in Brazil: a retrospective analysis of nationwide data. *Lancet Respir Med* 2021; 9:407-18.
27. Baqui P, Bica I, Marra V, Ercole A, Scharr M. Ethnic and regional variations in hospital mortality from COVID-19 in Brazil: a cross-sectional observational study. *Lancet Glob Health* 2020; 8:e1018-26.
28. Ribeiro H, Lima VM, Waldman EA. In the COVID-19 pandemic in Brazil, do brown lives matter? *Lancet Glob Health* 2020; 8:e976-7.
29. Peres IT, Bastos LSL, Gelli JGM, Marchesi JF, Dantas LF, Antunes BBP, et al. Sociodemographic factors associated with COVID-19 in-hospital mortality in Brazil. *Public Health* 2021; 192:15-20.
30. Niquini RP, Lana RM, Pacheco AG, Cruz OG, Coelho FC, Carvalho LM, et al. SRAG por COVID-19 no Brasil: descrição e comparação de características demográficas e comorbidades com SRAG por influenza e com a população geral. *Cad Saúde Pública* 2020; 36:e00149420.
31. de Souza WM, Buss LF, Candido DS, Carrera JP, Li S, Zarebski AE, et al. Epidemiological and clinical characteristics of the COVID-19 epidemic in Brazil. *Nat Hum Behav* 2020; 4:856-65.
32. Oliveira SB, Pôrto VBG, Ganem F, Mendes FM, Almiron M, Oliveira WK, et al. Monitoring social distancing and SARS-CoV-2 transmission in Brazil using cell phone mobility data. *medRxiv* 2020; 5 mai. <https://www.medrxiv.org/content/10.1101/2020.04.30.20082172v1>.
33. Ganem F, Mendes FM, Oliveira SB, Pôrto VBG, Araujo WN, Nakaya HI, et al. The impact of early social distancing at COVID-19 outbreak in the largest Metropolitan Area of Brazil. *medRxiv* 2020; 15 abr. <https://www.medrxiv.org/content/10.1101/2020.04.06.20055103v2>.
34. García-Basteiro AL, Chaccour C, Guinovart C, Lluçà A, Brew J, Trilla A, et al. Monitoring the COVID-19 epidemic in the context of widespread local transmission. *Lancet Respir Med* 2020; 8:440-2.
35. Carvalho CA, Carvalho VA, Campos MAG, Oliveira BLCA, Diniz EM, Santos AM, et al. Delay in death reporting affects timely monitoring and modeling of the COVID-19 pandemic. *Cad Saúde Pública* 2021; 37:e00292320.
36. Costa LFL, Trindade JEO, Oliveira ATR. Pareamento de dados das estatísticas do registro civil e das estatísticas vitais (Sistema de Informações sobre Nascidos Vivos e Mortalidade) – 2015. In: Oliveira ATR, organizador. *Sistemas de estatísticas vitais no Brasil: avanços, perspectivas e desafios*. Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística; 2018. p. 45-61.

37. Silva LV, Harb MPAA, Santos AMTB, Teixeira CAM, Gomes VHM, Cardoso EHS, et al. COVID-19 mortality underreporting in Brazil: analysis of data from government internet portals. *J Med Internet Res* 2020; 22:e21413.
38. Bastos LS, Niquini RP, Lana RM, Villela DAM, Cruz OG, Coelho FC, et al. COVID-19 e hospitalizações por SRAG no Brasil: uma comparação até a 12ª semana epidemiológica de 2020. *Cad Saúde Pública* 2020; 36:e00070120.
39. Passos VMA, Brant LCC, Pinheiro PC, Correa PRL, Machado IE, Santos MR, et al. Maior mortalidade durante a pandemia de COVID-19 em áreas socialmente vulneráveis em Belo Horizonte: implicações para a priorização da vacinação. *Rev Bras Epidemiol* 2021; 24:e210025.
40. Orellana JDY, Cunha GM, Marrero L, Moreira RI, Leite IC, Horta BL. Excesso de mortes durante a pandemia de COVID-19: subnotificação e desigualdades regionais no Brasil. *Cad Saúde Pública* 2020; 36:e00259120.
41. Orellana JDY, Marrero L, Horta BL. Excesso de mortes por causas respiratórias em oito metrópoles brasileiras durante os seis primeiros meses da pandemia de COVID-19. *Cad Saúde Pública* 2021; 37:e00328720.
42. Orellana JDY, Cunha GM, Marrero L, Horta BL, Leite IC. Explosão da mortalidade no epicentro amazônico da epidemia de COVID-19. *Cad Saúde Pública* 2020; 36:e00120020.
43. Silva GA, Jardim BC, Santos CVB. Excesso de mortalidade no Brasil em tempos de COVID-19. *Ciênc Saúde Colet* 2020; 25:3345-54.
44. Carvalho TA, Boschiero MN, Marson FAL. COVID-19 in Brazil: 150.000 deaths and the Brazilian underreporting. *Diagn Microbiol Infect Dis* 2021; 99:115258.
45. Lima EEC, Vilela EA, Peralta A, Rocha M, Queiroz BL, Gonzaga MR, et al. Investigating regional excess mortality during 2020 COVID-19 pandemic in selected Latin American countries. *Genus* 2021; 77:30.
46. Azevedo e Silva G, Jardim BC, Lotufo PA. Mortalidade por COVID-19 padronizada por idade nas capitais das diferentes regiões do Brasil. *Cad Saúde Pública* 2021; 37:e00039221.

Abstract

This study compares the death records due to COVID-19 in 2020 for Brazilian territory. Three distinct databases were used: Civil Registry (RC-Arpen), Brazilian Mortality Information System (SIM), and Influenza Epidemiological Surveillance Information System (SIVEP-Gripe). We identified differences between the numbers of deaths due to COVID-19 in these databases, varying in each federative unit. The RC database is updated faster than the other two databases (SIM and SIVEP-Gripe) and it is most suited for monitoring and for studies covering recent periods. Despite the slower update, Brazilian Health Informatics Department (DATASUS) databases present geographically similar numbers and disclose more detailed data on deaths. This detailed information improves the DATASUS databases for studies that require more information about the patient and treatment.

COVID-19; Civil Registration; Information Systems; Severe Acute Respiratory Syndrome

Resumen

Este estudio compara los registros de defunciones por COVID-19 en 2020 para todo el territorio brasileño. Se utilizaron tres bases diferentes: Registro Civil (RC-Arpen), Sistema de Información de Mortalidad (SIM) y Sistema de Información de Vigilancia Epidemiológica de la Gripe (SIVEP-Gripe). Existen discrepancias entre las cifras de muertes por COVID-19 comunicadas por las distintas bases de datos y estas diferencias varían en cada Unidad Federal. La base de datos RC se actualiza más rápidamente que las otras dos bases de datos del Departamento de Informática del Sistema Único de Salud – DATASUS (SIM y SIVEP-Gripe) y es más adecuada para el seguimiento y las encuestas que abarcan periodos más recientes. A pesar de que la actualización es más lenta, las bases de datos del DATASUS presentan cifras geográficamente similares y revelan datos más detallados sobre las muertes. Este detalle de la información hace que las bases de datos del DATASUS sean más adecuadas para investigaciones que requieren más información sobre el paciente y el tratamiento.

COVID-19; Registro Civil; Sistemas de Información; Síndrome Respiratorio Agudo Grave

Recebido em 29/Abr/2022
 Versão final reapresentada em 07/Nov/2022
 Aprovado em 12/Dez/2022