

Artigo submetido a 11 de Agosto 2021; versão final aceite a 21 de Setembro de 2021
Paper submitted on August 11, 2021; final version accepted on September 21, 2021

Variáveis Associadas à Mortalidade Por Covid-19 nos Municípios Brasileiros: Um Estudo Exploratório¹

Variables associated with Covid-19 mortality in Brazilian municipalities: an exploratory study

Horacio Nelson Hastenreiter Filho

hnhfilho@gmail.com

Professor associado da Escola de Administração da Universidade Federal da Bahia (UFBA)

Luiz Ricardo Cavalcante (autor correspondente/ corresponding author)

luiz.ricardo.teixeira.cavalcante@gmail.com

Consultor legislativo e professor do mestrado em administração pública do Instituto Brasileiro de Ensino, Desenvolvimento e Pesquisa (IDP)

Resumo

O objetivo deste trabalho é identificar variáveis significativamente correlacionadas com a mortalidade por covid-19 nos 5.570 municípios brasileiros. Os resultados das regressões usando dados em seções transversais indicam correlações positivas e significativas entre mortes por milhão de habitantes e as seguintes variáveis: *i*) população absoluta; *ii*) PIB *per capita*; *iii*) inclusão em região metropolitana; *iv*) percentual da população com 60 anos ou mais; *v*) percentual da população em áreas urbanas; e *vi*) percentual de votos em Jair Bolsonaro nas eleições de 2018. Os resultados se revelam robustos em diversos modelos, são consistentes com uma série de outros trabalhos já publicados e alcançam um período mais longo (mais de um ano de pandemia de covid-19 no Brasil).

Palavras-chave: Covid-19. Mortalidade. Municípios. Brasil.

Código JEL: R11, R15, R58

Abstract

The aim of this work is to identify variables significantly correlated with Covid-19 mortality in the 5,570 Brazilian municipalities. Results of cross section regressions indicate positive and significant correlations between deaths per million inhabitants and the following variables: *i*) total population; *ii*) per capita GDP; *iii*) inclusion in metropolitan region; *iv*) share of the population aged 60 years or over; *v*) share of the population living in urban areas; *vi*) Jair Bolsonaro's vote share in the 2018 presidential elections. Results are robust in several different models, consistent with a series of previous works on the subject and extended to a longer period of time (more than one year of Covid-19 pandemic in Brazil).

Keywords: Covid-19. Mortality. Municipalities. Brazil

Código JEL: R11, R15, R58

¹ Os autores agradecem ao grupo Núcleo de Operações e Inteligência em Saúde (NOIS) da PUC-RJ, a Lígia Gabrielli Fernandes, a Patrick Franco Alves e a Paula Maçaira.

1. INTRODUÇÃO

Declarada pandemia global pela Organização Mundial de Saúde (OMS) em março de 2020, a síndrome respiratória aguda coronavírus 2 (SARS-CoV-2) associada à covid-19 (acrônimo em inglês para *coronavirus disease 2019*) alcançou virtualmente todo o mundo em um intervalo bastante reduzido. Estima-se que os números de contaminações e de mortes chegaram, respectivamente, a 177 milhões e 3,8 milhões em junho de 2021. Observam-se, no entanto, enormes assimetrias entre continentes, países e regiões. Enquanto os australianos, por exemplo, se beneficiam de uma vida quase sem restrições e sem mortes associadas à doença, na América Latina, diversos países ainda convivem com ondas de contaminação. No tempo em que a rica Europa possui 14 países entre os 20 com maior taxa de óbitos por milhão de habitantes, na África, continente mais pobre do planeta, o país com pior taxa, a Tunísia, aparece apenas na 43ª posição no nefasto *ranking*.²

A posição do Brasil no *ranking* mundial das estatísticas de covid-19 é bastante desconfortável. Os dados relativos ao início de julho de 2021 colocam o país na segunda posição em número absoluto de mortes com um total de quase 530 mil, número superado somente pelos Estados Unidos. Considerando o número de mortes por milhão de habitantes (frequentemente usado em comparações internacionais e intrarregionais), o Brasil situa-se na sétima posição no *ranking*, atrás apenas de cinco países do leste europeu, todos com população inferior à da cidade de São Paulo, e do Peru. Dentro do país, há ainda grandes assimetrias entre as 27 unidades da federação. No início de julho, as estatísticas indicavam 1.278 mortes por milhão no Maranhão e 3.453 em Rondônia, que ocupavam as extremidades do *ranking*. Exercícios contrafactuais indicam, portanto, que o total de mortes no Brasil seria de 270 mil se as demais 26 unidades da federação repetissem a taxa do Maranhão e de 732 mil se fossem as taxas de Rondônia a se repetir no restante do país.

Ainda em 2020, a covid-19 já havia alcançado os 5.570 municípios do país.³ Também nesse caso há diferenças significativas entre as taxas de óbitos verificadas, muitas vezes na mesma unidade da federação. Discrepâncias desse tipo têm atraído a atenção de pesquisadores de múltiplas áreas do conhecimento em todo o mundo. Há uma vasta amplitude de variáveis explicativas e de modelos de análise adotados para esclarecer por que municípios e países se diferenciam positiva e negativamente das médias gerais de contaminação e mortes relacionadas à pandemia. Conforme se evidencia na seção 2 deste trabalho, diferenças internacionais foram objeto de trabalhos como os de Chaundry et al. (2020), que analisam o papel de fatores socioeconômicos, características da população, aspectos relativos aos sistemas de saúde e respostas dos governos nas taxas de mortalidade por covid-19. Padula e Davidson (2020) ressaltam que diferenças nas taxas de mortalidade têm associação com indicadores associados aos sistemas de saúde. No caso do Brasil, em particular, trabalhos como os de Raymundo et al (2021) procuram explicar a diferença entre a incidência de covid-19 nos municípios brasileiros a partir de dados sociodemográficos.

O presente trabalho insere-se nesse conjunto. Busca-se, em particular, identificar variáveis significativamente correlacionadas com o número de mortes por covid-19 acumuladas até 30/06/2021 por milhão de habitantes nos 5.570 municípios do país. Nesse sentido, estendem-se as análises a um período mais longo, uma vez que alcança mais de um ano de pandemia de covid-19 no Brasil. Além disso, algumas variáveis adicionais (como a inserção em regiões metropolitanas) são consideradas no conjunto das variáveis explicativas. Outra inovação do trabalho envolve o uso de regressões ponderadas pela população e de regressões específicas para as macrorregiões para validar os resultados. Para o atingimento desses objetivos, o artigo, além desta introdução, é composto de uma revisão bibliográfica sobre as variáveis que podem integrar o modelo, a descrição dos procedimentos metodológicos usados na coleta e na análise dos dados, os resultados obtidos e, finalmente, as considerações finais.

2. VARIÁVEIS ASSOCIADAS ÀS MORTES POR COVID-19

Os estudos voltados para a identificação de variáveis associadas à evolução da covid-19 cobrem um grande número de fatores e vêm recebendo enorme e constante destaque desde o início da pandemia,

² Os dados internacionais mais recentes mencionados nesta seção foram extraídos do site “Our World in Data”. Disponível em <https://cutt.ly/gmODIrS>. Acesso em 12 jul. 2021.

³ Disponível em <https://cutt.ly/Lne8hzL>. Acesso em 25 maio 2021.

não somente nas publicações da área de saúde, mas também em publicações relacionadas a outras áreas do conhecimento, como as ciências sociais. Nesta seção, em particular, analisam-se brevemente as publicações cujo foco recai sobre os fatores que, ao longo da pandemia, têm mostrado correlação com os números de casos e de mortes por covid-19 em países ou em jurisdições subnacionais. Tendo em vista o interesse que a pandemia tem despertado e a velocidade com que os resultados são divulgados, é evidente que qualquer revisão dessa natureza será limitada. Além disso, justamente por esses fatores, deve-se considerar que várias referências mencionadas nesta seção foram publicadas na forma de “texto para discussão” dos quais pode haver mais de uma versão à medida que os dados são atualizados. De qualquer modo, as referências indicadas a seguir parecem fornecer matéria-prima suficiente para a definição da metodologia e para a análise dos resultados obtidos.

É evidente que no âmbito das ciências de saúde há um enorme volume de publicações que analisam as taxas de mortalidade por covid-19 com base em dados dos indivíduos. Bhaskaran et al. (2021), por exemplo, buscaram identificar a relação da mortalidade com atributos como idade, sexo, obesidade, status em relação ao fumo, etnia e privação (medida de pobreza), tendo como contraponto a mortalidade por outras causas. O Serviço Nacional de Saúde do Reino Unido, por sua vez, relacionou grupos de indivíduos extremamente vulneráveis ao vírus (por exemplo, transplantados, pacientes em tratamentos radioterápicos e pacientes com alguns tipos específicos de tumores malignos).⁴ Mathur et al. (2021) estudaram o impacto das diferenças étnicas nas taxas de contaminação, hospitalização e morte por covid-19. Os autores concluem que *“some minority ethnic populations in England have excess risks of testing positive for SARS-CoV-2 and of adverse COVID-19 outcomes compared with the White population, even after accounting for differences in sociodemographic, clinical, and household characteristics”*. Todos esses trabalhos procuram identificar indivíduos mais ou menos vulneráveis a se contaminar, a desenvolver formas mais graves da doença ou de vir a óbito.

Essas análises foram rapidamente estendidas ao nível de países e de jurisdições subnacionais. Nesse caso, incorporam-se às variáveis demográficas (que refletem o perfil etário da população, por exemplo) outras variáveis que contribuam para explicar as discrepâncias nas taxas de mortalidade em países ou em jurisdições subnacionais. Portanto, as unidades de análise, nesses casos, são agregados populacionais dos quais é possível obter informações sobre casos e mortes por covid-19 e sobre as possíveis variáveis explicativas. Os primeiros trabalhos, ainda no início da pandemia, já tratavam das enormes discrepâncias nas taxas de contágio e de mortalidade em comparações internacionais. Ficou também rapidamente evidente que em alguns países, como a Itália e o Brasil, os elevados números de casos e óbitos não se distribuem uniformemente, havendo notáveis diferenças regionais a serem esclarecidas.

Ainda em maio de 2020, Chaundry et al. (2020) submeteram a um periódico de medicina clínica um artigo cujo foco recai sobre as taxas de mortalidade e sobre outras respostas (como número de casos críticos) em um amplo conjunto de países. No trabalho, os autores analisam fatores socioeconômicos, características da população, aspectos relativos aos sistemas de saúde e respostas dos governos (por exemplo, fechamento de fronteiras e restrições à circulação). Os autores concluem que *“low levels of national preparedness, scale of testing and population characteristics were associated with increased national case load and overall mortality”*. O foco de Padula e Davidson (2020) recai sobre o papel dos sistemas de saúde na redução das taxas de mortalidade. No Brasil, apenas dois meses após o registro do primeiro caso de covid-19, Hallal (2020) já se ocupava de apresentar diferenças entre taxas de mortalidade ao redor do mundo, correlacionando-as com a expectativa de vida nos diferentes países e conjecturando sobre a possível pertinência da testagem da população e das medidas de distanciamento social como variáveis explicativas.

Um surpreendente artigo, também com dados dos primeiros meses da pandemia, foi publicado por Aitken et al. (2020), estabelecendo associação positiva entre o Índice de Segurança de Saúde Global (ISSG) e os casos e óbitos por covid-19.⁵ Os autores especularam, então, que o índice que deveria, supostamente, refletir a capacidade de um país para lidar com epidemias ou pandemias, relacionava-se de forma oposta ao imaginado. Em um cenário de maior globalização, países mais desenvolvidos, e, portanto, com maior ISSG, têm maior exposição a estrangeiros que viajam para atividades turísticas e de negócios. Além disso, são países com baixo histórico de enfrentamento de pandemias e podem ter se

⁴ Disponível em <https://cutt.ly/TEbRXcF>. Acesso em 09 jul. 2021.

⁵ O ISSG foi desenvolvido em 2014 com o objetivo de medir a capacidade dos países de lidar com surtos de doenças infecciosas (Aitken et al., 2020).

iludido com a falsa garantia de que um alto ISSG poderia levá-los mais facilmente à adoção de mecanismos de controle da doença.

No Brasil, já há evidências de que a vulnerabilidade social tem forte correlação com a incidência de casos e com as taxas de mortalidade por covid-19. Cestari et al. (2021), por exemplo, analisam a distribuição espacial da incidência de casos de covid-19 em áreas distintas de Fortaleza. O trabalho foi submetido para publicação ainda em setembro de 2020. Rocha et al (2021) discutem, já em 2021, o efeito das desigualdades econômicas e das vulnerabilidades do sistema de saúde na capacidade de dar respostas no combate à pandemia.

Várias dessas considerações, no entanto, são aplicáveis a um período em que a covid-19 ainda se encontrava em seu processo inicial de disseminação. Após quase um ano e meio do registro do primeiro caso, as taxas de mortalidade se alteraram e outras variáveis parecem mais ajustadas à explicação das assimetrias verificadas. Os estudos mais atuais, quando o mundo já registra mais de 3,8 milhões de mortes e de 177 milhões de casos, vêm contemplando um amplo conjunto de fatores demográficos, econômicos e epidemiológicos, além de variáveis associadas à adoção de políticas públicas. A produção sobre o tema, mesmo que ainda recente, utiliza diversos métodos de análise. Além disso, em virtude do evidente interesse que a pandemia tem despertado em múltiplas áreas do conhecimento, as análises causais e explicativas não têm se limitado às tradicionais abordagens da área de saúde ou epidemiológica.

Raymundo et al (2021) empregam um amplo conjunto de dados sociodemográficos para explicar a diferença entre a incidência de covid-19 nos municípios brasileiros. Trabalhando com dados de mortalidade relativos ao período entre fevereiro e setembro de 2020 e usando múltiplos modelos de regressão, os autores concluem que a desigualdade de renda (aferida pelo índice de Gini) e o número de enfermeiros por mil habitantes estão positivamente correlacionados com a incidência da doença. Por outro lado, a mortalidade estaria negativamente correlacionada com a própria incidência da covid-19. Raymundo et al. (2021) argumentam que *“greater access to health services improved the diagnosis and notification of the disease and was associated with more cases in the municipalities”*. Eles demonstram também a presença de elevados índices de autocorrelação espacial e de multicolinearidade entre as variáveis. Isso os leva a excluir várias delas do modelo original e a restringir os fatores explicativos a um número reduzido. Isso explica por que aspectos como renda domiciliar (fortemente correlacionada, em nível municipal, com o número de enfermeiros por mil habitantes), por exemplo, estão ausentes de suas conclusões gerais.

No caso das políticas públicas, em particular, aspectos relacionados tanto à adoção de medidas não-farmacológicas (distanciamento social e uso de máscaras, por exemplo) pelos governos como à assimilação dessas medidas pela população estão evidentemente relacionados à disseminação da covid-19 e, portanto, às taxas de mortalidade. Trata-se tanto do estímulo à adoção dessas medidas (na forma, por exemplo, de campanhas ou de manifestações de personalidades influentes) como de sanções em caso de descumprimento das orientações das autoridades sanitárias. Na prática, a adesão efetiva da população às medidas não-farmacológicas resulta da combinação de estímulos e sanções, da permeabilidade da população às campanhas de esclarecimento e de sua predisposição ao *compliance* da legislação e das normas infralegais.

Gelfand et al. (2020) desenvolveram um estudo procurando relacionar a força das normas sociais, avaliada como rigidez cultural, e as taxas de casos e de mortes por milhão de habitantes registradas em diversos países do mundo. Foi verificada uma relação negativa entre rigidez cultural e mortes por milhão de habitantes, indicando que países com normas sociais mais rígidas têm mais controle sobre a pandemia. Isso implica dizer que aspectos de natureza essencialmente política podem interferir na evolução da covid-19.

No Brasil, esses aspectos têm sido analisados em diversos trabalhos que buscam associar a evolução da covid-19 em diferentes jurisdições subnacionais às preferências políticas da população, aferidas, em geral, com base no percentual de votos atribuídos ao presidente Jair Bolsonaro nas eleições de 2018. O pressuposto é que regiões em que o presidente obteve maiores percentuais de votos seriam mais sensíveis a sua postura frequentemente contrária às medidas não farmacológicas de combate à disseminação do vírus. Matéria publicada no jornal Valor Econômico em 24/05/2021, por exemplo, aponta que o contágio por covid-19 seria maior em municípios em que o presidente foi mais bem votado nas eleições de 2018. A matéria traz dados agrupados por decis ou segmentados por municípios em que o presidente obteve mais ou menos de 50% dos votos. Entretanto, trata-se de uma análise limitada porque não leva em consideração outras variáveis. Variáveis omitidas eventualmente correlacionadas com o perfil dos

eleitores evidentemente podem concorrer para explicar a evolução da covid-19. Porém, mesmo antes da publicação dessa matéria, alguns trabalhos acadêmicos que envolvem controles mais rigorosos buscaram avaliar a influência do perfil eleitoral na evolução da covid-19.

Ajzenman, Cavalcanti e da Mata (2020), ainda no início da pandemia, identificaram que, após o presidente do Brasil rejeitar pública e enfaticamente os riscos associados à pandemia e desaconselhar a adoção de medidas de restrição à circulação, as medidas de distanciamento social dos cidadãos em localidades favoráveis ao presidente reduziram-se muito mais significativamente do que em lugares em que o seu apoio é mais fraco. O estudo mostrou-se robusto para diferentes especificações do modelo empírico e para diferentes métricas de apoio político. Complementarmente, encontraram evidências sugestivas de que esse impacto é impulsionado em localidades com níveis relativamente mais altos de penetração da mídia. Da mesma forma, Leone (2021) usa dados de geolocalização de 60 milhões de aparelhos celulares para medir o distanciamento social e reafirma uma associação estatisticamente significativa entre essa variável e o apoio ao presidente Bolsonaro.

Cabral, Ito e Pongeluppe (2021) corroboram a visão de que as lideranças políticas podem exercer um papel crítico no combate à pandemia, sobretudo a partir da forma como influenciam o comportamento social. Em estudo baseado nas taxas de contágio e nos percentuais de votos no presidente na eleição de 2018 nos 5.570 municípios brasileiros, os pesquisadores discutem a influência das lideranças negacionistas nos diferentes resultados em relação à transmissão do vírus, observando com especial atenção a correlação entre o crescimento na contaminação e óbitos no país e um conjunto de discursos do presidente proferidos na televisão e nas rádios em março e em abril de 2020 em que se minimizava a gravidade da pandemia de covid-19. A conclusão é que em municípios com maior adesão ao presidente, os habitantes são mais propensos a assumir comportamentos de risco, descuidando-se do uso de máscaras e das medidas de isolamento, desrespeitando o regramento estabelecido pelas medidas de distanciamento e facilitando, assim, a maior circulação do vírus.

De modo geral, os diferentes estudos aqui citados consideram, como variáveis dependentes, os números de casos e de mortes por covid-19. Ocasionalmente, indicadores indiretamente relacionados com essas variáveis (por exemplo, indicadores de distanciamento social aferidos com base em dados obtidos de telefones celulares) são também usados como variáveis dependentes. Os estudos usam um amplo espectro de variáveis explicativas, que vão de características demográficas (por exemplo, perfil etário da população), passam por vários indicadores socioeconômicos (por exemplo, PIB *per capita*) e chegam a medidas de preferências políticas. Há também uma ampla variedade de modelos de análise, que vão de correlações simples a modelos mais complexos. Em seu conjunto, esses trabalhos amparam a escolha da metodologia adotada neste artigo (envolvendo a seleção das variáveis empregadas e a escolha do modelo de análise, que são objeto da próxima seção) e a interpretação dos resultados reportados na seção 4.

3. METODOLOGIA

Nesta seção, descrevem-se os procedimentos metodológicos adotados para identificar variáveis significativamente correlacionadas com o número de mortes por covid-19 por milhão de habitantes nos municípios brasileiros. Inicialmente, com base na revisão bibliográfica da seção precedente, foram selecionadas variáveis para as quais se dispõe de informações em nível municipal, que foram então tabuladas para os 5.570 municípios do país. As fontes de dados empregadas estão indicadas em detalhes na seção 3.1. Em seguida, definiram-se modelos de análise estatística para verificar a existência de correlação entre as variáveis selecionadas e a mortalidade por covid-19.

3.1 Seleção de variáveis e fontes de dados

Há um volume razoavelmente grande de dados em escala municipal relacionados à covid-19. O quadro 1 registra os dados empregados neste trabalho segmentando-os nos seguintes blocos:

- Informações geográficas: relacionadas com a localização do município (nome, códigos de seis e de sete dígitos, unidade da federação, região, e eventual inserção em região metropolitana).⁶ Essas

⁶ Embora regressões de dados municipais frequentemente empreguem matrizes de distâncias entre os municípios ou outros dados espaciais como variáveis explicativas, optou-se, nesse caso, por simplesmente usar uma *dummy* para a inserção ou não do município em regiões

informações foram extraídas da base “Produto Interno Bruto dos Municípios” do Instituto Brasileiro de Geografia e Estatística (IBGE).

- Informações relacionadas à covid-19 (números de casos e de mortes). Dados relativos aos 5.570 municípios do país foram extraídos do repositório de dados públicos “Brasil.io”. Uma vez que o modelo adotado usa dados transversais, empregaram-se apenas os dados acumulados até a data de referência. Na prática, filtraram-se os registros relativos a 30/06/2021 e extraíram-se, para essa data, os dados dos municípios.

- Informações sociodemográficas: relacionadas com o perfil da população do município (população residente, densidade demográfica, população com 60 anos ou mais, sexo, cor ou raça e população urbana). Os dados relativos à população residente em 2020 foram extraídos da base “Estimativas da População” também do IBGE, assim como as áreas territoriais dos municípios (usadas para calcular a densidade demográfica). Estimativas de população residente por faixa etária e por sexo relativas a 2020 foram extraídas do DataSUS.⁷ Os percentuais da população em áreas urbanas e de cor/raça branca foram extraídos do censo 2010 do IBGE.

- Informações econômicas: relacionadas com indicadores de renda (PIB e PIB *per capita*) e de distribuição de renda (índice de Gini): as primeiras foram extraídas da base “Produto Interno Bruto dos Municípios” do IBGE. Nesses casos, os dados mais recentes referem-se a 2018. Já os dados sobre o índice de Gini da renda domiciliar *per capita* dos municípios relativos a 2010 (último ano disponível) foram extraídos do DataSUS.

- Informações de atendimento de saúde e de educação: relacionadas com indicadores de saúde (número de leitos e de médicos por 100 mil habitantes) e de educação (população com nível superior). Os indicadores de saúde provieram de dados consolidados pelo IBGE na base Geo Hub covid-19, ao passo que os de educação foram extraídos do censo 2010 do IBGE.

- Informações políticas: votos nas eleições presidenciais de 2018 no primeiro e no segundo turno. Esses dados já em formato de planilha foram extraídos do site “Dados Finos”. Optou-se por trabalhar com os percentuais de votos no primeiro e no segundo turno obtidos pelos dois candidatos mais bem colocados. Como se trata de variáveis negativamente correlacionadas no primeiro turno e perfeitamente negativamente correlacionadas no segundo, nas regressões optou-se por estimar o coeficiente de uma ou de outra, mas não de ambas simultaneamente.

O quadro 1, sintetiza o conjunto de variáveis empregadas, suas descrições e suas fontes.

Naturalmente, diversas outras variáveis poderiam ter sido selecionadas com base nas referências indicadas na seção 2 deste trabalho. Buscou-se limitar esse conjunto para evitar uma espécie de inflação de variáveis que mantêm entre si elevados níveis de correlação. Por exemplo, o acesso ao auxílio emergencial, que foi utilizado por Pereira et al (2021) para indicar vulnerabilidade social, é evidentemente correlacionado com indicadores econômicos e de desigualdade. Da mesma forma, o número de enfermeiros por mil habitantes – positivamente correlacionado com a incidência da covid-19 no modelo de Raymundo et al. (2021) – mantém uma forte associação com a renda domiciliar *per capita* (ausente de seu modelo). Optou-se, dessa forma, por selecionar um conjunto relativamente limitado de variáveis capazes de capturar, direta ou indiretamente, as dimensões da análise. Além disso, buscou-se privilegiar as variáveis mais “gerais” (por exemplo, *renda per capita*) sempre que havia multicolinearidade com variáveis mais “específicas” (por exemplo, acesso a programas de transferência de renda ou número de profissionais de saúde). Com isso, buscou-se evitar resultados que, à primeira vista, poderiam parecer contraintuitivos (por exemplo, uma correlação positiva entre mortes e extensão da rede de saúde, em grande medida explicada pela renda). As informações políticas, em particular, visam a capturar algum tipo de permeabilidade maior ou menor ao discurso contrário, de alguma forma, às medidas não-farmacológicas de combate à disseminação do vírus. Não foram usados dados referentes à vacinação nos municípios em virtude do percentual relativamente pequeno da população imunizada até o final do primeiro semestre de 2021. Mesmo no caso do município de Serrana (onde a vacinação foi feita de forma mais acelerada), optou-se por não fazer um tratamento destacado do restante da base.

metropolitanas. A razão para isso é que se assume que a disseminação da Covid-19 está mais relacionada a algum tipo de conurbação dos municípios do que propriamente à distância física entre eles. Em outras palavras: a influência de municípios relativamente próximos, mas que não estabelecem interações econômicas e sociais relevantes é menor do que a influência que municípios que pertencem a uma mesma aglomeração urbana exercem uns sobre os outros.

⁷ Disponíveis em <https://cutt.ly/snhvolm> e em <https://cutt.ly/pnkeNIZ>. Acesso em 01 jun. 2021.

Quadro 1: variáveis utilizadas nos modelos

| Variáveis | Descrição | Fontes |
|--|--|--|
| Informações geográficas (relacionadas com a localização no município) | | |
| COD7, COD6 e NOME | Códigos (seis e sete dígitos) e nome do município | IBGE (https://cutt.ly/Inhyr4W) |
| UF, N, NE, CO, SE, S e RM | Unidade da federação, região, e eventual inserção em região metropolitana | IBGE (https://cutt.ly/Inhyr4W) |
| Informações relacionadas à covid-19 | | |
| CASOS MORTES | Número de casos e número de mortes por covid-19 acumulados até 30/06/2021 por município (boletins epidemiológicos das 27 Secretarias Estaduais de Saúde) | Brasil.io (https://cutt.ly/In5kdUJ) |
| Informações sociodemográficas (perfil da população) | | |
| POP | População residente (2020) | IBGE (https://cutt.ly/Gn5aaT1) |
| DENS_DEM | Densidade demográfica (2020) calculada com base na população residente e na área dos municípios | IBGE (https://cutt.ly/enhPo2D) |
| PERC60MAIS | Percentual da população com 60 anos ou mais (2020, estimado) | DataSUS (https://cutt.ly/Jn5d3wF) |
| PERC_MASC | Percentual da população do sexo masculino (2020, estimado) | DataSUS (https://cutt.ly/Jn5d3wF) |
| PERC_URB | Percentual da população em área urbana (2010) | IBGE (https://cutt.ly/Tn5kHOc) |
| PERC_BRANCOS | Percentual da população de cor / raça branca (2010) | IBGE (https://cutt.ly/Tn5kHOc) |
| Informações econômicas (indicadores de renda e de distribuição de renda) | | |
| PIB e PIB_P_CAP | PIB (2018) e PIB <i>per capita</i> (2018) | IBGE (https://cutt.ly/Inhyr4W) |
| GINI | Índice de Gini da renda domiciliar per capita dos municípios (2010) | DataSUS (https://cutt.ly/pnkeNIZ) |
| Informações de atendimento de saúde e de educação (médicos, leitos e educação superior) | | |
| MEDICOS_100MIL | Quantidade de postos de trabalho de profissionais médicos, no SUS e em hospitais privados (2019) de acordo com o Cadastro Nacional de Estabelecimentos de Saúde (CNES) | IBGE (Geo Hub Covid-19) (https://cutt.ly/vn5hTB3) |
| LEITOS_100MIL | Quantidade de leitos hospitalares, exceto os leitos psiquiátricos, no SUS e em hospitais privados (dezembro de 2019) de acordo com o Cadastro Nacional de Estabelecimentos de Saúde (CNES) | IBGE (Geo Hub Covid-19) (https://cutt.ly/vn5hTB3) |
| PERC_POP_EDUC_SUP | Percentual da população com pelo menos nível superior de graduação concluído (2010) | IBGE (https://cutt.ly/un5jorG) (https://cutt.ly/0n5jfsb) |
| Informações políticas (votos nas eleições presidenciais de 2018) | | |
| PERCVOTOS17_1T PERCVOTOS17_2T PERCVOTOS13_1T PERCVOTOS13_2T | Percentuais de votos nos dois candidatos mais bem colocados primeiro e no segundo turno das eleições presidenciais de 2018 | “Dados Finos” (https://cutt.ly/vnktTWL) |

Fonte: elaboração dos autores.

3.2 Modelo de análise

O modelo geral de estimação utilizado neste trabalho associa o vetor de variáveis dependentes Y_i (números de casos e número de mortes por covid-19) aos vetores formados pelas variáveis geográficas

(G_i), sociodemográficas (S_i), econômicas (E_i), atendimento de saúde e de educação (SE_i) e políticas (P_i), considerando ainda um termo de erro ϵ_i .

$$Y_i = \beta_0 + \beta_1 G_i + \beta_2 S_i + \beta_3 E_i + \beta_4 SE_i + \beta_5 P_i + \epsilon_i (1)$$

O grande número de variáveis explicativas possíveis, muitas das quais associadas entre si, implica dizer que um modelo que considerasse todo seu conjunto exibiria elevados níveis de multicolinearidade. Isso pode ocorrer, inclusive, com variáveis classificadas em diferentes blocos (por exemplo, variáveis econômicas estão obviamente correlacionadas com variáveis sociodemográficas e de atendimento de saúde e de educação). Por essa razão, optou-se por modelos mais parcimoniosos, mas buscou-se, sempre que possível, incluir todas as dimensões (embora nem todas as variáveis ao mesmo tempo). Testes de multicolinearidade foram empregados para validar os modelos e para descartar variáveis redundantes.

Com relação à escolha do modelo de regressão, adotou-se o seguinte procedimento:

- Inicialmente, testaram-se vários modelos de regressão usando mínimos quadrados ordinários (MQO). Embora menos robustos no caso em análise (em função da não-normalidade da distribuição de variáveis como o número de mortes por milhão de habitantes), esses modelos acabam gerando, no caso específico, resultados semelhantes aos de modelos mais sofisticados (ver, por exemplo, Roubaud et al., 2020).

- Uma vez selecionadas as variáveis a serem empregadas, novas regressões foram executadas com base nos modelos lineares generalizados (MLG). Nesse caso, optou-se por distribuições de Poisson e binomial negativa. Em linha com resultados previamente reportados por Roubaud et al (2020), as variáveis selecionadas mantiveram sua significância mesmo com base em modelos mais robustos.

Analísaram-se, quando aderentes aos modelos adotados, *i*) as hipóteses de normalidade e de ausência de multicolinearidade; e *ii*) os resíduos das regressões. Alternativamente, foram estimados os coeficientes de regressões ponderadas pelas populações dos municípios. Esse procedimento atribui peso proporcionalmente maior a municípios mais populosos nas regressões, em lugar de tratar da mesma forma municípios com alguns milhares e com muitos milhões de habitantes. Por fim, regressões específicas para cada uma das cinco macrorregiões brasileiras foram também executadas.

Na prática, portanto, estimaram-se os coeficientes de um conjunto de regressões usando dados em seções transversais nas quais o número de mortes por milhão de habitantes no município i foi a variável dependente e variáveis como população total, *PIB per capita* e outras foram usadas como variáveis explicativas. De modo geral, adotou-se o logaritmo natural das variáveis (inclusive do número de mortes por milhão de habitantes), exceto quando se tratava de porcentagens (por exemplo, população residente com 60 anos ou mais) ou de *dummies* (por exemplo, inserção em regiões metropolitanas). Embora dados sobre número de casos estivessem também disponíveis, optou-se por trabalhar com o número de mortes porque se assume que pode haver grande subnotificação de casos e que a variância dessa subnotificação pode ser igualmente elevada. Nesse sentido, o número de mortes parece refletir mais adequadamente a disseminação da covid-19 no Brasil. Essa opção é consistente com vários artigos citados na seção 2 deste trabalho. Para o número de mortes por milhão de habitantes, usou-se o logaritmo natural, no caso de modelos de regressão mais simples (mínimos quadrados), e contagem, no caso dos modelos generalizados.

4. RESULTADOS

Inicialmente, foram testadas diversas combinações de variáveis explicativas nos modelos MQO. De modo geral, buscou-se selecionar as variáveis que exibissem elevada significância estatística e que agregassem poder explicativo ao modelo (isto é, que elevassem o R^2 ajustado). Algumas variáveis se mostraram significativas apenas em combinações específicas; esse foi o caso, por exemplo, do índice de Gini. Nesses casos, optou-se por excluí-las do modelo. Como resultado, privilegiaram-se variáveis que se revelaram recorrentemente significativas. A inclusão de *dummies* para macrorregiões nem sempre se revelou significativa, embora em alguns casos os dados indiquem uma correlação positiva e significativa

com a inserção na região Norte, por exemplo. Isso explica a opção por excluir essas *dummies* das regressões e a opção por regressões específicas para cada uma dessas macrorregiões.

De modo geral, municípios mais populosos, com maior PIB *per capita*, situados em regiões metropolitanas, com maior presença de pessoas com 60 anos ou mais, com maior percentual de população urbana e maiores percentuais de votos no presidente Jair Bolsonaro nas eleições presidenciais de 2018 tendem a apresentar maiores níveis de mortalidade. Os indicadores de atendimento de saúde e de educação nem sempre se revelaram significativos (possivelmente em virtude de sua correlação com os indicadores econômicos já incluídos no modelo) e por essa razão não foram incluídos.

Desse modo, o modelo ajustado nas regressões MQO foi:

$$Y_i = \beta_0 + \beta_1 LN_POP_i + \beta_2 LN_PIB_P_CAP_i + \beta_3 RM_i + \beta_4 PERC60MAIS_i + \beta_5 PERC_URB_i + \beta_7 PERCVOTOS17_1T_i + \epsilon_i \quad (2)$$

O valor obtido para o R^2 ajustado no modelo (0, 3085) revela sua limitada capacidade preditiva (uma vez que apenas pouco mais de 30% das variações do número de mortes por milhão de habitantes são explicados pelas variáveis inseridas). Esse é um resultado esperado e consistente com outros exercícios dessa natureza, uma vez que diversos outros fatores podem concorrer para explicar o número de mortes por milhão de habitantes nos municípios brasileiros. O quadro 2 registra os resultados obtidos para esse modelo e registra os baixíssimos p-valores observados em todos os casos.

Quadro 2: resultados do modelo MQO

| | Coefficiente | Erro padrão | t-valor | p-valor |
|----------------|--------------|-------------|---------|---------------------------|
| Intercepto | 4.479765 | 0.148179 | 30.232 | $< 2 \times 10^{-16}$ *** |
| LN_POP | 0.037629 | 0.007414 | 5.075 | 4×10^{-07} *** |
| LN_PIB_P_CAP | 0.155134 | 0.014724 | 10.536 | $< 2 \times 10^{-16}$ *** |
| RM | 0.053384 | 0.017278 | 3.090 | 0.00201 *** |
| PERC60MAIS | 1.813985 | 0.179963 | 10.080 | $< 2 \times 10^{-16}$ *** |
| PERC_URB | 0.686934 | 0.040642 | 16.902 | $< 2 \times 10^{-16}$ *** |
| PERCVOTOS17_1T | 0.665757 | 0.057749 | 11.529 | $< 2 \times 10^{-16}$ *** |

Residual standard error: 0.534 on 5489 degrees of freedom (74 observations deleted due to missingness),

Multiple R-squared: 0.3093.

Adjusted R-squared: 0.3085,

F-statistic: 409.6 on 6 and 5489 DF, p-value:

$< 2.2e-16$. Significância: 0 '***'; 0.001 '**'; 0.01 '*';

0.05 '.'; 0.1 ' ';

Fonte: elaboração dos autores.

A análise dos resíduos do modelo (indicada na figura 1), por sua vez, sugere um ajuste adequado dos dados, uma vez que não exibem um padrão definido, são razoavelmente alinhados à linha dos quantis teóricos e são homocedásticos. Além disso, a presença de *outliers* é reduzida.⁸

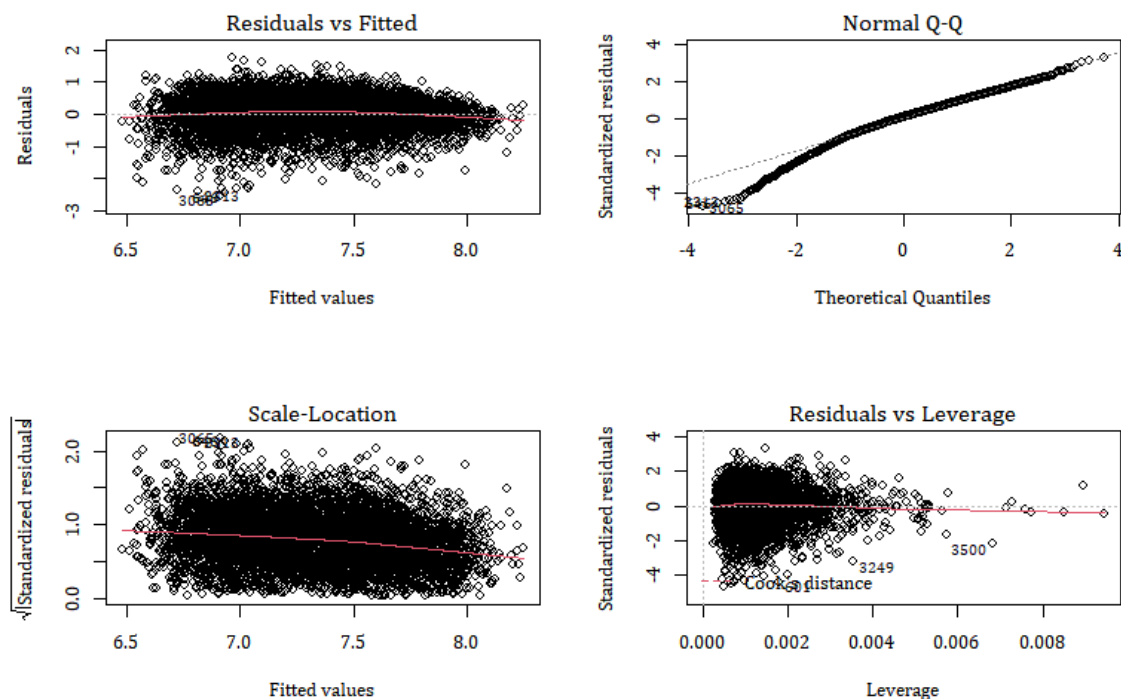
Testes de multicolinearidade exibiram, para esse modelo, resultados satisfatórios (revelando que as variáveis explicativas podem estar simultaneamente presentes na regressão). Contudo, o modelo tem limitações relacionadas à autocorrelação entre os resíduos e à normalidade das variáveis. Esse é, por exemplo, o caso das mortes por milhão de habitantes, cujo histograma está mostrado na figura 2.

A distribuição não normal é bastante evidente no gráfico. Além disso, percebe-se uma certa “inflação de zeros” (que leva, em um modelo em que se usa o logaritmo das mortes por milhão de habitantes, à eliminação de alguns registros).

Por essa razão, recorreu-se, após a seleção das variáveis a serem empregadas, às regressões com base em MLG assumindo distribuições de Poisson e binomial negativa. Os resultados da regressão MLG assumindo a distribuição de Poisson estão indicados no quadro 3.

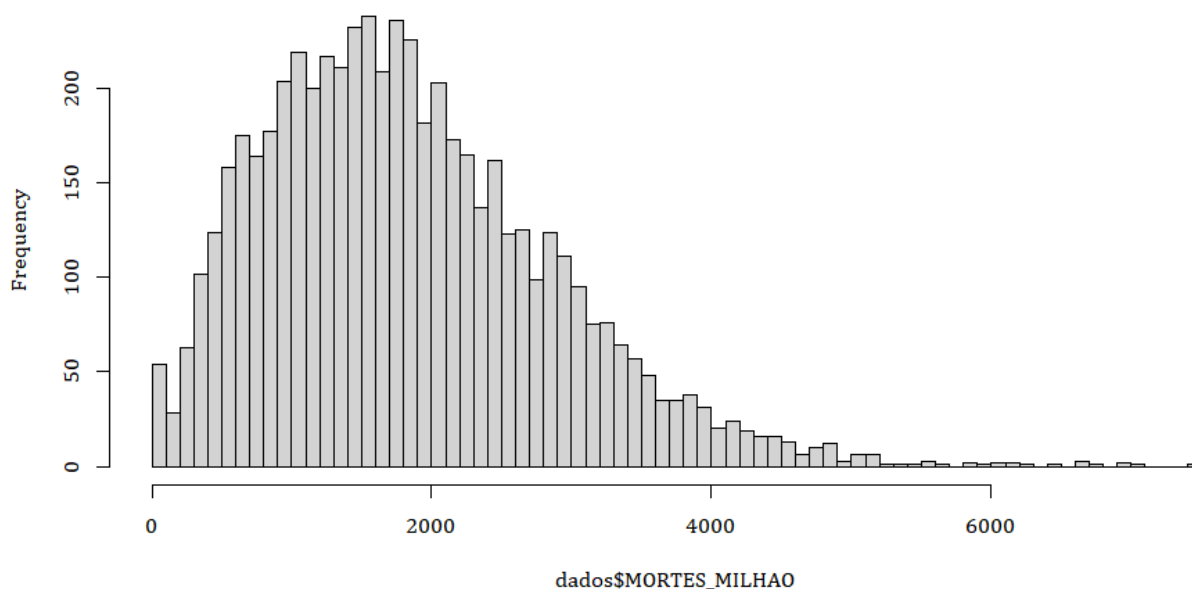
⁸ Disponível em <https://cutt.ly/4ma4rKO>. Acesso em 30/06/2021.

Figura 1: resíduos do modelo



Fonte: elaboração dos autores.

Figura 2: histograma do logaritmo dos óbitos por covid-19 acumulados até 30/06/2021 por milhão de habitantes, municípios brasileiros



Fonte: elaboração dos autores.

Quadro 3: resultados do modelo (Poisson)

| | Coefficiente | Erro padrão | t-valor | p-valor |
|----------------|---------------------|--------------------|----------------|---------------------------|
| Intercepto | 5,2446385 | 0,0061169 | 857,40 | $< 2 \times 10^{-16}$ *** |
| LN_POP | 0,0201800 | 0,0003071 | 65,72 | $< 2 \times 10^{-16}$ *** |
| LN_PIB_P_CAP | 0,1166727 | 0,0006090 | 191,59 | $< 2 \times 10^{-16}$ *** |
| RM | 0,0180132 | 0,0007342 | 24,54 | $< 2 \times 10^{-16}$ *** |
| PERC60MAIS | 1,6145908 | 0,0076845 | 210,11 | $< 2 \times 10^{-16}$ *** |
| PERC_URB | 0,5875367 | 0,0018354 | 320,11 | $< 2 \times 10^{-16}$ *** |
| PERCVOTOS17_1T | 0,6710454 | 0,0024963 | 268,81 | $< 2 \times 10^{-16}$ *** |

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 3214311 on 5542 degrees of freedom

Residual deviance: 2291048 on 5536 degrees of freedom

(27 observations deleted due to missingness)

AIC: 2341539

Significância: 0 '***'; 0.001 '**'; 0.01 '*'; 0.05 '.'; 0.1 ' ';

Fonte: elaboração dos autores.

Já os resultados da regressão MLG assumindo a distribuição binomial negativa estão indicados no quadro 4.

Tabela 4: resultados do modelo (binomial negativa)

| | Coefficiente | Erro padrão | z-valor | p-valor |
|----------------|---------------------|--------------------|----------------|---------------------------|
| Intercepto | 4,976099 | 0,157664 | 31,561 | $< 2 \times 10^{-16}$ *** |
| LN_POP | 0,018785 | 0,007868 | 2,387 | 0,0203 * |
| LN_PIB_P_CAP | 0,142541 | 0,015634 | 9,117 | $< 2 \times 10^{-16}$ *** |
| RM | 0,042773 | 0,018434 | 2,320 | 0,0170 * |
| PERC60MAIS | 1,711069 | 0,191312 | 8,944 | $< 2 \times 10^{-16}$ *** |
| PERC_URB | 0,609155 | 0,043266 | 14,079 | $< 2 \times 10^{-16}$ *** |
| PERCVOTOS17_1T | 0,647746 | 0,061513 | 10,530 | $< 2 \times 10^{-16}$ *** |

(Dispersion parameter for Negative Binomial (2.2326) family taken to be 1)

Null deviance: 7751.0 on 5542 degrees of freedom

Residual deviance: 6077,1 on 5536 degrees of freedom

(27 observations deleted due to missingness)

AIC: 91088

Number of Fisher Scoring iterations: 1

Theta: 3.0719

Std. Err.: 0.0572

2 x log-likelihood: -91071.8720

Significância: 0 '***'; 0.001 '**'; 0.01 '*'; 0.05 '.'; 0.1 ' ';

Fonte: elaboração dos autores.

Os resultados indicados nas tabelas 2 e 3 são mais robustos do que aqueles obtidos com base no modelo MQO (em função da não-normalidade da distribuição de variáveis como o número de mortes por milhão), mas são essencialmente idênticos e reafirmam a significância das variáveis selecionadas. Em síntese, os resultados reafirmam as correlações positivas e significativas entre mortes por milhão de habitantes até 30/06/2021 e as seguintes variáveis: *i*) população absoluta; *ii*) PIB *per capita*; *iii*) inclusão em região metropolitana; *iv*) percentual da população com 60 anos ou mais; *v*) percentual da população em áreas urbanas; e *vi*) percentual de votos em Jair Bolsonaro no primeiro turno das eleições de 2018.

A relação positiva entre população e mortes por covid-19 reflete o fato de que maiores centros urbanos tendem a concentrar (pelo menos no período analisado) os focos de disseminação da doença. O PIB *per capita* revelou-se positivamente correlacionado com as mortes por covid-19 possivelmente porque o vírus tende a circular mais onde há mais atividade econômica (controlando por outras variáveis). A inserção em região metropolitana revelou-se significativa porque, no caso de municípios idênticos, aqueles que estão mais próximos de centros urbanos maiores com os quais mantêm vínculos econômicos e sociais tendem a estar mais expostos à circulação do vírus. Conforme esperado, o perfil etário mantém uma forte correlação com o número de mortes por covid-19, assim como o grau de urbanização do

município. Finalmente, os resultados reafirmam a correlação positiva e significativa entre as taxas de mortalidade e o percentual de votos destinados a Jair Bolsonaro no primeiro turno das eleições presidenciais de 2018. De modo geral, os resultados são consistentes com uma série de outros trabalhos já publicados e os estendem a um período mais longo, uma vez que alcançam mais de um ano de pandemia de covid-19 no país.

Em todos os modelos foram também estimados os coeficientes de regressões ponderadas pelas populações dos municípios. Esse procedimento – que atribui peso proporcionalmente maior a municípios mais populosos nas regressões – não alterou os principais resultados obtidos. Os coeficientes mantiveram-se significativos (embora com eventuais mudanças de escala, em função da introdução dos pesos de ponderação). No caso da regressão MQO, o R^2 ajustado alcançou 0,5184, indicando uma elevação considerável do poder explicativo do modelo.

Finalmente, regressões em que se aplicam filtros para as macrorregiões também reafirmaram, de modo geral, os resultados obtidos.

Tabela 4: resultados do modelo com filtros macrorregionais (binomial negativa)

| | Norte | Nordeste | Centro-Oeste | Sudeste | Sul |
|-----------------|---------------------------|---------------------------|---------------------------|---------------------------|----------------------------|
| Intercepto | 3,67696 | 3,84285 | 5,44444 | 4,42017 | 7,07356 |
| Pr(> z) | $1,21 \times 10^{-7} ***$ | $< 2 \times 10^{-16} ***$ | $< 2 \times 10^{-16} ***$ | $< 2 \times 10^{-16} ***$ | $< 2 \times 10^{-16} ***$ |
| RM | 0,11622 | 0,09570 | 0,01728 | -0,06514 | 0,08702 |
| Pr(> z) | 0,141632 | 0,00233 ** | 0,826 | 0,132 | 0,01111 * |
| LN_PIB_P_CAP | 0,23779 | 0,20359 | 0,06606 | 0,13204 | -0,01504 |
| Pr(> z) | 0,000352 *** | $3,5 \times 10^{-11} ***$ | 0,128 | $1,23 \times 10^{-5} ***$ | 0,69630 |
| LN_POP | 0,07272 | 0,04290 | 0,02128 | 0,02313 | 0,01862 |
| Pr(> z) | 0,010691 * | 0,00355 ** | 0,408 | 0,115 | 0,30894 |
| PERC60MAIS | 3,29692 | 3,98650 | 3,40469 | 2,54653 | 1,29327 |
| Pr(> z) | 0,002660 ** | $< 2 \times 10^{-16} ***$ | $4,43 \times 10^{-7} ***$ | $1,16 \times 10^{-8} ***$ | 0,00031 *** |
| PERC_URB | 0,34597 | 0,88204 | 0,62761 | 0,62506 | 0,56879 |
| Pr(> z) | 0,025703 * | $< 2 \times 10^{-16} ***$ | $5,17 \times 10^{-5} ***$ | $1,18 \times 10^{-9} ***$ | $1,02 \times 10^{-10} ***$ |
| PERCVO-TOS17_1T | 0,52584 | -0,08601 | 1,00801 | 1,46554 | -0,06546 |
| Pr(> z) | 0,021110 * | 0,65738 | $4,79 \times 10^{-5} ***$ | $< 2 \times 10^{-16} ***$ | 0,63845 |

Significância: 0 ****; 0,001 ***; 0,01 **; 0,05 *; 0,1 .;

Fonte: elaboração dos autores.

Conforme se observa, ainda que haja uma redução dos p-valores (inclusive em função do menor número de observações quando os filtros são adotados) e que ocasionalmente alguns coeficientes deixem de ser significativos em regiões específicas (nos casos das regiões Centro-Oeste e Sul, isso se observa com três variáveis; no caso da região Sudeste, com duas; no caso das regiões Norte e Nordeste, com uma), os resultados reafirmam as conclusões gerais a que se chegou com a análise agregada. Em regressões complementares com os filtros macrorregionais e ponderadas pela população total de cada município, todas as variáveis se mantêm significativas em todos os casos.

5. CONSIDERAÇÕES FINAIS

Neste trabalho, identificou-se um conjunto abrangente de variáveis significativamente correlacionadas com o número de mortes por covid-19 acumuladas até 30/06/2021 por milhão de habitantes nos 5.570 municípios do país. Os resultados obtidos agregam-se a esforços precedentes de identificação de variáveis associadas aos números de casos e de mortes por covid-19 em países e em jurisdições subnacionais e estendem as análises a um período mais longo, uma vez que alcançam mais de um ano de pandemia de covid-19 no Brasil. Novas variáveis foram consideradas nas regressões, com destaque para a inserção em regiões metropolitanas. Os resultados se mostraram robustos em diversos modelos de

análise. Regressões ponderadas pela população e regressões específicas para as macrorregiões foram usadas para validar os resultados.

Desse modo, embora diversos outros estudos com objetivos similares venham sendo desenvolvidos (conforme se evidenciou na seção 2), a extensão do período abrangido, a diversidade de aspectos em análise e os múltiplos exercícios de validação dos resultados parecem ser contribuições importantes para a necessidade persistente de se explicarem as diferenças territoriais na evolução da pandemia. Essas diferenças deverão manter seu caráter dinâmico nos próximos meses, uma vez que a pandemia recrudesce ou se torna mais amena em diferentes regiões e municípios do país, segundo aspectos que podem se renovar ao longo do tempo.

O processo de vacinação em curso no país, que se dá em velocidades diferentes nos diversos estados e municípios, certamente impactará as taxas de contaminações e óbitos também assimetricamente enquanto um conjunto relevante da população não estiver vacinada. É improvável que nos próximos modelos dedicados ao mesmo objeto do presente artigo o percentual de vacinação não se apresente como relevante variável explicativa. Obviamente, a compreensão mais abrangente das variações nas taxas de contaminações e de óbitos requer uma combinação de estudos populacionais com estudos que exigem um acompanhamento mais próximo e rigoroso de um segmento populacional, como os estudos de coorte e os de caso-controle, por exemplo. Trabalhos complementares com esses tipos de abordagens devem ser desenvolvidos, sobretudo, quando tiverem como perspectiva explicar as diferenças nas taxas de óbito entre vacinados e não vacinados, cada vez mais demandadas e relevantes.

Não se sabe ao certo, independentemente do sucesso do processo vacinal, por quanto tempo a humanidade continuará a conviver com a pandemia de covid-19. Além disso, como crescentemente consensuado entre os epidemiologistas do mundo, pandemias podem se tornar cada vez mais habituais. Assim sendo, a necessidade de compreensão de sua dinâmica e a identificação de fatores que interferem na sua capacidade de se espalhar e causar mortes seguramente permanecerão na ordem do dia.

BIBLIOGRAFIA

Aitken, T., Chin, K. L., Liew, D., & Ofori-Asenso, R. (2020). Rethinking pandemic preparation: Global Health Security Index (GHSI) is predictive of COVID-19 burden, but in the opposite direction. *Journal of Infection*, 81 (2), 318-356. <https://doi.org/10.1016/j.jinf.2020.05.001>.

Ajzenman, N., Cavalcanti, T. & Da Mata, D. (2020). More Than Words: Leaders' Speech and Risky Behavior during a Pandemic (Working paper, Cambridge-INET Working Paper Series No: 2020/19 / Cambridge Working Papers in Economics: 2034). Cambridge: Institute for New Economic Thinking. <https://doi.org/10.2139/ssrn.3582908>.

Bhaskaran, K., Bacon, S., Evans S., Bates, C. J., Rentsch, C. T., MacKenna, B., Tomlinson, L., Walker, A. J., Schultze, A., Morton, C. E., Grint, D., Mehrkar, A., Eggo, R. M., Inglesby, P., Douglas, I. J., McDonald, H. I., Cockburn, J., Williamson, E. J., Evans, D., Curtis, H. J., Hulme, W. J., Parry, J., Hester, F., Harper, S., Spiegelhalter, D., Smeeth, L., & Goldacre, B. (2021). Factors associated with deaths due to COVID-19 versus other causes: population-based cohort analysis of UK primary care data and linked national death registrations within the OpenSAFELY platform. *The Lancet Regional Health – Europe*, 6, 1-13. <https://doi.org/10.1016/j.lanepe.2021.100109>.

Cabral, S., Ito, N. C., & Pongeluppe, L. S. (2021). The Disastrous Effects of Leaders in Denial: Evidence from the COVID-19 Crisis in Brazil. Available at: <http://dx.doi.org/10.2139/ssrn.3836147>.

Cestari, V.R.F., & Moreira, T.M.M. (2020). Vulnerabilidade social e incidência de COVID-19 em uma metrópole brasileira. *Ciênc. saúde coletiva*, 26 (3), 1023-1033. <https://doi.org/10.1590/1413-81232021263.42372020>.

Chaudhry, R., Dranitsaris, G., Mubashir, T., Bartoszko, J., & Riazi, S. (2020). A country level analysis measuring the impact of government actions, country preparedness and socioeconomic factors on COVID-19 mortality and related health outcomes. *EClinicalMedicine*, 25, 100464. <https://doi.org/10.1016/j.eclinm.2020.100464>

Gelfand, M. J., Jackson, J. C., Pan, X., Nau, D., Pieper, D., Denison, E., Dagher, M., Van Lange, P. A. M., Chiu, C.-Y., & Wang, M. (2021). The relationship between cultural tightness–looseness and

COVID-19 cases and deaths: a global analysis. *Lancet Planet Health*, 5, e135-e144. [https://doi.org/10.1016/S2542-5196\(20\)30301-6](https://doi.org/10.1016/S2542-5196(20)30301-6).

Hallal, P.C. (2020). Worldwide differences in COVID-19-related mortality. *Ciênc. saúde coletiva*, 25 (suppl 1), 2403-2410. <https://doi.org/10.1590/1413-81232020256.1.11112020>.

Leone, T. (2021). The Harmful Effects of Denial: When Political Polarization meets COVID-19 Social Distancing. *Middle Atlantic Review of Latin American Studies*, 4 (3), 10-30. <https://doi.org/10.23870/marlas.327>.

Mathur, R., Rentsch, C. T., Morton, C. E., Hulme, W. J., Schultze, A., MacKenna, B., Eggo, R. M., Bhaskaran, K., Wong, A. Y. S., Williamson, E. J., Forbes, H., Wing, K., McDonald, H. I., Bates, C., Bacon, S., Walker, A. J., Evans, D., Inglesby, P., Mehrkar, A., ... Goldacre, B. (2021). Ethnic differences in SARS-CoV-2 infection and COVID-19-related hospitalisation, intensive care unit admission, and death in 17 million adults in England: An observational cohort study using the OpenSAFELY platform. *The Lancet*, 397(10286), 1711–1724. [https://doi.org/10.1016/S0140-6736\(21\)00634-6](https://doi.org/10.1016/S0140-6736(21)00634-6).

Padula, W. V., & Davidson, P. (2020). Countries with High Registered Nurse (RN) Concentrations Observe Reduced Mortality Rates of Coronavirus Disease 2019 (COVID-19). *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3566190>.

Pereira, M. D., Pereira, M. D., Amorim, L. P., Santos, G. K. N., & Bezerra, C. M. O. (2020). O isolamento social em tempos de covid-19 e seu paradoxo com os grupos em vulnerabilidade social. *Caderno de Graduação – Ciências Humanas e Sociais*, 6 (2), 239-249. <https://periodicos.set.edu.br/cadernohumanas/article/view/9034>.

Raymundo, C. E., Oliveira, M. C., Eleuterio, T. ., André, S. R.; Silva, M. G., Queiroz, E. R. S., & Medronho, R. A. (2021). Spatial analysis of COVID-19 incidence and the sociodemographic context in Brazil. *PLoS ONE*, 16 (3), 1-16. <https://doi.org/10.1371/journal.pone.0247794>.

Rocha, R., Atun, R., Massuda, A., Rache, B., Spínola, P., Nunes, L., Lago, M. & Castro, M. C. (2021). Effect of socioeconomic inequalities and vulnerabilities on health-system preparedness and response to COVID-19 in Brazil: a comprehensive analysis. *Lancet Glob Health*, 9 (6), e-782-e792. [https://doi.org/10.1016/S2214-109X\(21\)00081-4](https://doi.org/10.1016/S2214-109X(21)00081-4).

Roubaud, F., Razafindrakoto, M., Saboia, J., Castilho, M., & Pero, V. (2020). The municípios facing COVID-19 in Brazil: socioeconomic vulnerabilities, transmission mechanisms and public policies (Discussion Paper 032/2020). Rio de Janeiro: UFRJ / Instituto de Economia.