



**UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”
FACULDADE DE MEDICINA**

Gabriel Berg de Almeida

**Aplicação de modelos matemáticos em pandemias:
um estudo de comportamento epidemiológico a partir
da Covid-19.**

Tese apresentada à Faculdade de Medicina de Botucatu, Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus de Botucatu, para obtenção do título de Doutor em Doenças Tropicais.

**ORIENTADOR: Prof. Dr. Carlos Magno Castelo Branco Fortaleza
COORIENTADORA: Prof. Dra. Cláudia Pio Ferreira**

**Botucatu
2022**

Gabriel Berg de Almeida

**Aplicação de modelos matemáticos em pandemias:
um estudo de comportamento epidemiológico a
partir da Covid-19.**

Tese apresentada à Faculdade de Medicina de Botucatu, Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus de Botucatu, para obtenção do título de Doutor em Doenças Tropicais.

ORIENTADOR: Prof. Dr. Carlos Magno Castelo Branco Fortaleza

COORIENTADORA: Prof. Dra. Cláudia Pio Ferreira

Botucatu

2022

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO TÉC. AQUIS. TRATAMENTO DA INFORM.
DIVISÃO TÉCNICA DE BIBLIOTECA E DOCUMENTAÇÃO - CÂMPUS DE BOTUCATU - UNESP
BIBLIOTECÁRIA RESPONSÁVEL: ROSEMEIRE APARECIDA VICENTE-CRB 8/5651

Almeida, Gabriel Berg de.

Aplicação de modelos matemáticos em pandemias : um estudo de comportamento epidemiológico a partir da Covid-19 / Gabriel Berg de Almeida. - Botucatu, 2022

Tese (doutorado) - Universidade Estadual Paulista "Júlio de Mesquita Filho", Faculdade de Medicina de Botucatu

Orientador: Carlos Magno Castelo Branco Fortaleza

Coorientador: Cláudia Pio Ferreira

Capes: 40101096

1. Pandemias. 2. Covid-19. 3. Epidemiologia - Pesquisa.
4. Análise espacial (Estatística). 5. Modelos matemáticos.
6. SARS-CoV-2.

Palavras-chave: Análise espacial; Covid-19; Epidemiologia; Modelagem matemática; SARS-CoV-2.

Dedicatória

Dedico este trabalho a todos aqueles afetados da maneira mais drástica que uma epidemia pode afetar o ser humano, com a dor e com a morte. Dedico a todos aqueles que perderam entes queridos. Àqueles que ficaram com sequelas e ainda lutam pela recuperação. Dedico a todos que não tiveram a oportunidade e o privilégio de se prevenir, ou, ainda, não tiveram o acesso ao cuidado em saúde quando precisaram. Àqueles massacrados pela falta de políticas públicas responsáveis, pelo negacionismo, pelas *fake news* e pela miséria da situação em que foram colocados. Nosso trabalho sempre foi para minimizar o sofrimento e a perda.

Agradecimentos

Agradeço primeiramente ao meu orientador, professor Carlos Magno. A admiração pela pessoa e pelo trabalho é de longa data, desde os primeiros contatos ainda na graduação em medicina. Mas a parceria é mais recente e eu não posso deixar de agradecer a todos os ensinamentos. Se hoje eu posso buscar o título de epidemiologista, é porque tive escola. Muito obrigado.

À minha coorientadora, professora Cláudia, pela paciência extrema na imersão em modelagem matemática. As tarde e noites foram longas, mas valeram cada esforço. Minha admiração é gigante.

Ao Thomas, pela parceria em todos os artigos aqui publicados. Sua análise e seu companheirismo foram fundamentais.

À professora Cristiana, que embora não estivesse diretamente envolvida neste trabalho, foi essencial para meu entendimento em epidemiologia e modelagem matemática.

À Ângela, Michelle e Lorena, companheiras incansáveis de pesquisa.

Aos meus pais, Luiz Eduardo e Mary. Vocês são a base da minha formação. Obrigado.

À minha irmã, Maria Eduarda, maior motivo para que eu nunca desista.

Ao Pedro, que acompanhou todo o processo. Obrigado pela compreensão e incentivo.

Aos meus amigos. À minha família.

Obrigado.

Epígrafe

“A transmissibilidade do cólera não deve ser ocultada do povo, sob a alegação de que tal conhecimento geraria pânico ou implicaria o abandono dos enfermos”

John Snow,

em Sobre a Maneira da Transmissão do Cólera

Resumo

Modelos matemáticos, estatísticos e computacionais são extensamente utilizados para o estudo das doenças infecciosas, sua dispersão e seus fatores de risco. A identificação do SARS-CoV-2 e da doença causada por ele, a Covid-19, culminou em esforços globais para o entendimento da dinâmica de transmissão e dispersão em uma situação pandêmica. O presente estudo é fruto de um processo contínuo de pesquisa epidemiológica em saúde. Foi proposto um modelo matemático para o estudo da dinâmica de transmissibilidade do SARS-CoV-2 em pequenas e médias cidades no interior do Brasil, e o possível efeito de intervenções em saúde pública. Também foram propostos modelos geográficos, a partir de análise exploratória de dados secundários e visualização cartográfica: estudamos a evolução da epidemia no interior do Estado de São Paulo, relacionando o grau de conectividade entre municípios e as estruturas elementares espaciais que influenciam a dispersão da doença; e a distribuição espacial precoce da incidência de Covid-19 no Brasil, em conjunto com o aumento da capacidade laboratorial brasileira para o diagnóstico molecular da doença. Por fim, apresentamos um estudo de acompanhamento dos duzentos primeiros dias de epidemia em São Paulo, através de estimativas do número de reprodução (R_t) do SARS-CoV-2 e sua relação com as intervenções em saúde pública adotadas pelo governo do Estado. Nossos resultados em modelagem matemática demonstram que intervenções não-farmacológicas são necessárias para o controle de epidemia e que cada município deve ser estudado individualmente, levando-se em consideração fatores bióticos e abióticos que podem influenciar na

transmissão de doença. O modelo geográfico confirma a hipótese de que o SARS-CoV-2 se dissemina por contiguidade da metrópole para suas cidades vizinhas mais próximas e que também é observado um espalhamento à distância, um processo de dispersão hierárquica, na qual os municípios maiores se relacionam com a metrópole, através de rodovias, hidrovias, pontes aéreas e do fluxo de pessoas e comércio. Além disso, fica claro que a incidência precoce da Covid-19 foi mais importante em grandes metrópoles brasileiras, e avançou no sentido do interior em todas as regiões, acompanhado do importante aumento da capacidade de diagnóstico por laboratórios públicos credenciados. Por fim, demonstramos os resultados do acompanhamento do R_t ao longo dos duzentos dias subsequentes à ocorrência do primeiro caso de Covid-19 em São Paulo e como a variação em suas estimativas consegue ser relacionada às diferentes medidas restritivas vigentes no momento. Todos os estudos apresentados são complementares, buscando uma compreensão ecológica e, através de evidência científica, auxiliar as tomadas de decisão em gestão de saúde pública e, conseqüentemente, o enfrentamento racional à epidemia.

Abstract

Mathematical, statistical and computational models are widely used to study infectious diseases, their spread and their risk factors. Identifying SARS-CoV-2 and the disease it causes, Covid-19, has culminated in global efforts to understand the dynamics of transmission and dispersion in a pandemic situation. The present study results from a continuous process of epidemiological health research. A mathematical model was proposed to study the dynamics of SARS-CoV-2 transmissibility in small and medium-sized cities in the interior of Brazil and the possible effect of public health interventions. Geographical models were also proposed, based on exploratory secondary data analysis and cartographic visualization. We studied the evolution of the epidemic in the inner State of São Paulo, relating the degree of connectivity between municipalities and the elementary spatial structures that influence the spread of the disease; And the early spatial distribution of the incidence of Covid-19 in Brazil, together with the increase in Brazilian laboratory capacity for the molecular diagnosis of the disease. Finally, we present a follow-up study of the first two hundred days of the epidemic in São Paulo through estimates of the reproduction number (R_t) of SARS-CoV-2 and its relationship with public health interventions adopted by the state government. Our results in mathematical modelling demonstrate that non-pharmacological interventions are necessary for epidemic control and that each municipality should be studied individually, taking into account biotic and abiotic factors that can influence disease transmission. The geographic model confirms the hypothesis that SARS-CoV-2 spreads by contiguity from the metropolis to its closest neighbouring cities and that a

distant spread is also observed, a process of hierarchical dispersion, in which the largest municipalities relate to other cities, the metropolis, through highways, waterways, air bridges and the flow of people and commerce. In addition, it is clear that the early incidence of Covid-19 was more important in large Brazilian metropolises and advanced towards the interior in all regions, accompanied by a significant increase in diagnostic capacity by accredited public laboratories. Finally, we demonstrate the results of monitoring the R_t over the two hundred days following the first case of Covid-19 in São Paulo and how the variation in its estimates can be related to the different restrictive measures in force at the time. All the studies presented are complementary, seeking an ecological understanding and, through scientific evidence, helping decision-making in public health management and, consequently, the rational confrontation of the epidemic.

Lista de abreviaturas

CEP – Comitê de Ética em Pesquisa

CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico

DRS – Departamento Regional de Saúde

DOE – Diário Oficial do Estado

EUA – Estados Unidos da América

HCFMB – Hospital das Clínicas da Faculdade de Medicina de Botucatu

IC – Intervalo de confiança

OMS – Organização Mundial de Saúde

OPAS – Organização Pan-Americana da Saúde

R_0 – Número Básico de Reprodução

R_t – Número Efetivo de Reprodução

SIVEP-Gripe – Sistema de Informação de Vigilância Epidemiológica da Gripe

SRAG – Síndrome Respiratória Aguda Grave

UTI – Unidade de Terapia Intensiva

Sumário

INTRODUÇÃO.....	12
OBJETIVOS.....	17
ÉTICA EM PESQUISA.....	19
ARTIGO 1.....	20
ARTIGO 2.....	34
ARTIGO 3.....	48
ARTIGO 4.....	51
CONCLUSÃO.....	55
REFERÊNCIAS BIBLIOGRÁFICAS.....	60

Introdução

A pandemia de Covid-19, em pouco menos de dois anos, já se tornou o maior e mais importante evento epidêmico desde a gripe espanhola, no início do século XX. A identificação de um novo vírus respiratório, na cidade de Wuhan, na China, em dezembro de 2019, com potencial pandêmico, alarmou as autoridades em saúde pública de todo o mundo.¹ A Covid-19, doença causada pelo SARS-CoV-2 (até então denominado 2019-nCov), ficou conhecida mundialmente como uma pneumonia viral, capaz de levar indivíduos à falência respiratória e a desfechos graves como internação em leito de terapia intensiva (UTI) e óbito.

Modelos matemáticos, estatísticos e computacionais têm sido extensamente utilizados para o estudo das doenças infecciosas, sua dispersão e seus fatores de risco.^{2,3} Classicamente, a epidemiologia é definida como o estudo da associação entre as doenças e os fatores de risco individuais ou populacionais. Já a modelagem matemática e estatística, especialmente aquela que utiliza abordagem não autônoma e espacial, pode auxiliar no entendimento da dinâmica de transmissão e da dispersão espacial das doenças, bem como realizar projeções em curto e longo prazo.

Assim, a declaração de emergência em saúde pública de interesse internacional pela Organização Mundial de Saúde (OMS)⁴ em 30 de janeiro de 2020, culminou em um esforço global do entendimento não só da biologia do novo vírus e da doença capaz de provocar, mas também da possibilidade de identificação de padrões de dispersão nos diversos países e continentes e do acometimento de diferentes populações.

No Brasil, o primeiro caso de Covid-19 foi confirmado em 25 de fevereiro de 2020, na cidade de São Paulo, em um viajante internacional oriundo da Itália.⁵ Em conjunto, Itália e Estados Unidos (EUA), foram as principais rotas de importação do SARS-CoV-2, em decorrência do importante fluxo aeroviário em um momento de espalhamento da cepa original.⁶ Mas o que esperar da dinâmica de transmissão de um novo agente infeccioso em um país de dimensões continentais e com importantes disparidades regionais, como o Brasil, e quais intervenções que podem impactar nesse cenário?

O presente estudo é fruto de um processo contínuo do entendimento precoce de transmissão e dispersão da Covid-19 no Brasil. Inicialmente, as grandes metrópoles foram objeto de estudo de diversos grupos de pesquisa em epidemiologia e modelagem matemática, mas foram identificados poucos modelos que incluíssem pequenos e médios municípios no interior do país.^{7,8,9,10} O objetivo inicial então, foi desenvolver um modelo dinâmico de transmissão de doença infecciosa que contemplasse esses municípios e permitisse a avaliação de possíveis intervenções em saúde pública.

Em modelos compartimentais, indivíduos de uma população são divididos em subgrupos (compartimentos) e a dinâmica da infecção é estudada coletivamente.¹¹ Um modelo compartimental clássico SEIR (Susceptível-Exposto-Infectado-Recuperado) foi adaptado e proposto para a Covid-19. Considerando a possibilidade de isolamento e quarentena de infectados e infectantes, um novo compartimento foi considerado (Q), gerando um modelo SEIRQ.

Definida a estrutura do modelo matemático, é essencial a definição de parâmetros adequados para que o modelo gere resultados confiáveis. Inicialmente, os parâmetros fundamentais, como número básico de reprodução (R_0), período

infeccioso e taxa de hospitalizações e óbitos, por exemplo, foram obtidos a partir da literatura dos primeiros casos e surtos locais, principalmente China e Itália. Com o objetivo de incorporar características locais brasileiras para o modelo, foram escolhidas aleatoriamente 29 cidades pequenas e médias do interior de todas as regiões do país, que pudessem ser representativas e estudadas no que se refere a fatores bióticos e abióticos com potencial interferência no modelo.

Apesar do fato de que os resultados desse primeiro exercício de modelagem pudessem ajudar a entender a dinâmica de transmissão e auxiliar a gestão de saúde na tomada de decisão das medidas de contingenciamento a serem adotadas, a falta de dados epidemiológicos locais dificultou a validação do modelo. Assim, durante o curso da pandemia, e enquanto dados locais eram gerados e atualizados, o modelo também foi atualizado, com parâmetros locais mais bem definidos e com desempenho melhorado.

Em paralelo, também nos propusemos a desenvolver estudos de análises exploratórias dos novos dados epidemiológicos, capazes de auxiliar no entendimento do espalhamento do SARS-CoV-2 no Brasil e no Estado de São Paulo. Modelos de dispersão foram estudados, através de estatística espacial e visualização cartográfica, técnicas imprescindíveis nesse contexto. A partir de dados secundários de banco de dados em acesso público, como número de casos graves de Covid-19 notificados em plataforma oficial (SIVEP-Gripe) ao longo do tempo, pudemos entender e descrever como se deu a dispersão inicial da Covid-19 da metrópole, sentido interior.

No Estado de São Paulo, foi possível relacionar essa dispersão aos diversos equipamentos sociais e econômicos existentes, incluindo principais rodovias, hidrovias e pontes aéreas, bem como o fluxo de pessoas e comércio de bens de consumo ou produção. E quando consideramos o país como um todo, foi possível

observar o avanço da Covid-19 no mesmo sentido metrópoles-interior e relacionar com o aumento da capacidade laboratorial brasileira ao longo do tempo para o diagnóstico molecular do SARS-CoV-2.

A partir do momento em que as duas pedras fundamentais desse estudo estavam mais bem definidas, ou seja, já era possível visualizar através de um modelo matemático como possivelmente seria a dinâmica de transmissão do SARS-CoV-2 no interior do país e entendida como havia se dado a introdução e dispersão da Covid-19 para o interior, foi possível acompanhar a transmissão comunitária do SARS-CoV-2 no Estado de São Paulo. Foi proposta uma metodologia de acompanhamento dos casos graves confirmados de Covid-19 e também de Síndrome Respiratória Aguda Grave (SRAG) que pudessem ser relacionadas com as medidas de contingenciamento propostas pelo governo do Estado de São Paulo, o Plano São Paulo.

Foi calculado o número de reprodução do SARS-CoV-2 ao longo do tempo para cada Departamento Regional de Saúde (DRS) do Estado de São Paulo. A divisão do território estadual em DRS se refere a uma divisão administrativa, através da Secretaria de Saúde do estado de São Paulo, para a coordenação das atividades de saúde em cada região. Durante a pandemia, essa divisão, que já é bem estabelecida desde 2006 pelo decreto DOE nº 51.433, foi utilizada para implementação de medidas de controle não-farmacológicas regionalizadas, a partir de um protocolo estabelecido pelo comitê de contingenciamento da Covid-19 no Estado. Assim, avaliamos, ao longo dos duzentos primeiros dias de pandemia, como se comportou a transmissão da Covid-19 em um cenário de transmissão comunitária e intervenções não-farmacológicas, na ausência de vacinas disponíveis.

Todos os estudos apresentados foram produtos de trabalhos em grupos de pesquisa, envolvendo infectologistas e epidemiologistas, geógrafos, e modeladores, foram revisados por pares e publicados em revistas científicas internacionais. Em conjunto, eles buscam demonstrar como é possível prever a dinâmica de transmissão de uma doença infecciosa, as possíveis medidas de controle que podem ou devem ser implementadas, além de uma análise inicial de dispersão de primeiros casos através do território e, enfim, como acompanhar a transmissão através do tempo e relacionar com as medidas não farmacológicas adotadas.

Objetivos

Objetivos Gerais

- Realizar um estudo de modelagem matemática capaz de projetar a dinâmica de transmissão do SARS-CoV-2 no Brasil e avaliar possíveis medidas de intervenções não-farmacológicas.

- Realizar um estudo de análise exploratória e modelagem espacial capaz de descrever a dispersão inicial do SARS-CoV-2 no Estado de São Paulo.

- Realizar um estudo de análise exploratória e modelagem espacial capaz de descrever a dispersão inicial do SARS-CoV-2 e relacionar com a capacidade de diagnóstico laboratorial no Brasil.

- Realizar um estudo de acompanhamento da evolução temporal da Covid-19 no estado de São Paulo e relacionar com as medidas de controle de transmissão ao longo do tempo.

Objetivos Específicos

- Desenvolvimento de um modelo matemático dinâmico clássico SEIRQ de transmissão de doença infecciosa para cidades pequenas e médias do interior do Brasil

- Mapear os primeiros casos de Covid-19 no interior do estado de São Paulo e realizar análise geoespacial da introdução e dispersão do SARS-CoV-2 para o interior do estado de São Paulo

- Mapear os primeiros casos de Covid-19 no Brasil e realizar análise geoespacial do credenciamento de novos laboratórios públicos para diagnóstico molecular do SARS-CoV-2.

- Calcular o número de reprodução do SARS-CoV-2 através do tempo nos diferentes Departamentos Regionais de Saúde (DRS) do estado de São Paulo e relacionar com as medidas de intervenção do governo estadual (Plano São Paulo).

Ética em pesquisa

Este projeto foi dispensado de avaliação pelo Comitê de Ética em Pesquisa (CEP) da Faculdade de Medicina de Botucatu-UNESP, por se tratar de um estudo ecológico, populacional, e com uso de dados secundários disponíveis em bancos públicos abertos.



OPEN

Addressing the COVID-19 transmission in inner Brazil by a mathematical model

G. B. Almeida^{1✉}, T. N. Vilches^{2,4}, C. P. Ferreira^{3,4} & C. M. C. B. Fortaleza^{1,4}

In 2020, the world experienced its very first pandemic of the globalized era. A novel coronavirus, SARS-CoV-2, is the causative agent of severe pneumonia and has rapidly spread through many nations, crashing health systems and leading a large number of people to death. In Brazil, the emergence of local epidemics in major metropolitan areas has always been a concern. In a vast and heterogeneous country, with regional disparities and climate diversity, several factors can modulate the dynamics of COVID-19. What should be the scenario for inner Brazil, and what can we do to control infection transmission in each of these locations? Here, a mathematical model is proposed to simulate disease transmission among individuals in several scenarios, differing by abiotic factors, social-economic factors, and effectiveness of mitigation strategies. The disease control relies on keeping all individuals' social distancing and detecting, followed by isolating, infected ones. The model reinforces social distancing as the most efficient method to control disease transmission. Moreover, it also shows that improving the detection and isolation of infected individuals can loosen this mitigation strategy. Finally, the effectiveness of control may be different across the country, and understanding it can help set up public health strategies.

It has been a year since the first confirmed case of a novel coronavirus pneumonia in Wuhan, China. Now, the world experiences its very first pandemic of the globalized era. SARS-CoV-2 has rapidly spread through the currently connected continents, and the World Health Organization has declared a health emergency on international concern, which made many countries taking serious mitigation and suppression strategies¹.

These strategies take importance when we look at the epidemic dynamics. The first studies estimated that the basic reproductive number of COVID-19 was 2.68 (95% CrI 2,47-2,86)², which means one infected person can spread the virus to almost three people in a totally susceptible community. As there is no treatment or vaccine wide available, the best way to control the virus is to diminish social contact. China has shown to the world that when people stay at home, the virus circulation can be controlled, and we have more time for preparing health systems, producing individual protection equipment, developing research, and minimizing the consequences of the epidemic³. However, in Brazil, this kind of mitigation strategy (social distancing) does not work for self-employed people and low-income families since their maintenance depends on their own work. Besides, the number of people living in the same house can vary from 1.7 to 7.7 in the country. Only 52.5% of Brazilian households have basic sanitation and less than two residents per bedroom. Moreover, 6% of the Brazilian population lives in slums where access to safe water, basic sanitation, waste management, and hygienic conditions is not guaranteed⁴⁻⁷.

In Brazil, the introduction of COVID-19 happened later than in many other locations, and that gave us time to analyze all the new scientific evidence and the control measures taken overseas⁸. However, a country with continental dimensions cannot work with a single plan response. In the higher urban hierarchy cities, like São Paulo, for example, the disease spread initially from the medium and high levels of social classes to the lower ones⁹. However, what should be the scenario for inner Brazil, and what can we do on infection control in each of these locations?

Mathematical modeling has taken significant importance when applied to epidemics^{10,11}. Since the earliest population studies on plague or measles, the methods have been refined. Today, with the parameters well established and more sensitive parameterization, such as contact patterns matrices, we may estimate how an epidemic will behave in a specific population and what should be our immediate response to the problem¹².

¹Medical School of Botucatu, São Paulo State University, Botucatu 18618-687, Brazil. ²Institute of Mathematics, Statistics, and Scientific Computing, University of Campinas, Campinas 13083-859, Brazil. ³Institute of Biosciences, São Paulo State University, Botucatu 18618-689, Brazil. ⁴These authors contributed equally: T. N. Vilches, C. P. Ferreira, and C. M. C. B. Fortaleza. ✉email: gb.almeida@unesp.br

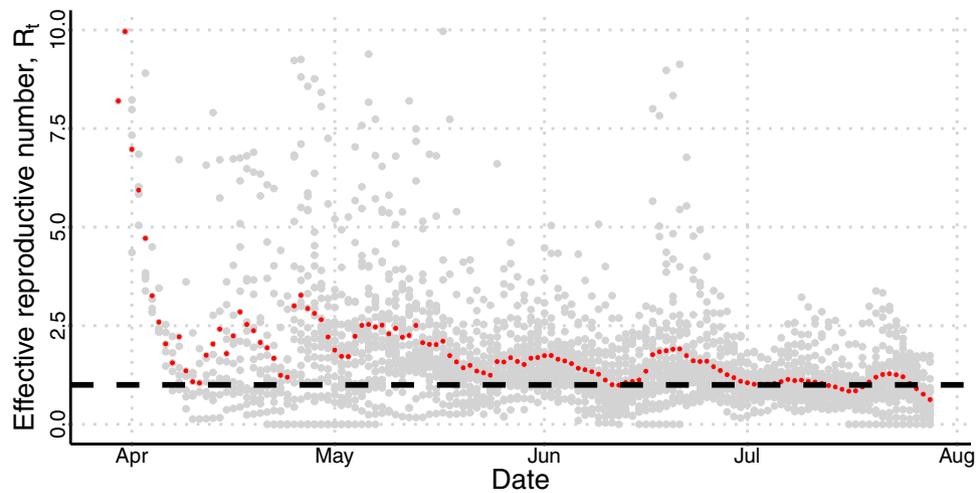


Figure 1. Temporal course of R_t in each municipality involved in the study. In red, the average value, and in grey, the individual values. The dashed line shows the threshold of $R_t = 1$. Above it, the transmission of the disease increases; below it, the disease's transmission decreases.

Mathematical models may draw best and worst-case scenarios for a COVID-19 epidemic situation in small and medium cities in inner Brazil¹³. Our main objective is to study how the disease might behave in specific cities of the country and see what happens when combining two strategies: diminishing social contact plus testing and isolating positive cases. We point out that we are not aiming to characterize the temporal dynamics of the COVID-19 transmission in a city, state, or country, but highlight the difference in the disease transmission across the country and emphasize that control must be done differently in each one of these regions. Besides, we clustered the cities based on a set of characteristics in order to see if these were able to give us any clue about the disease dynamics. The consequences of relaxing restrictions are a theme of debate now, and in Brazil, it is happening before the epidemic's peak has occurred, while the number of cases is still growing over the country. Mask wearing, mass testing, early detection of imported cases, and monitoring effective reproduction number are strategies that have been discussed and adopted around the world¹⁴.

Results

The temporal evolution of the effective reproduction number R_t is shown in Fig. 1. It was calculated for each municipality using data of daily incidence of cases and the knowledge about the generation interval of COVID-19¹⁵. In red, we plotted the average value, and in gray, the individual values. Mean R_t 's higher values were primarily observed at the beginning—absolute values from five to ten—which quickly changed to values between 1.8 and 2.7, after 10 days, since control measures were rapidly adopted across the country. It is crucial to notice that the large variability of R_t , and the oscillation, observed at each day when plotting cities together, can be associated with the fact that the cities started epidemic at different moments and the delay into report cases on weekends. Over time, R_t slowly decreases until June. Although R_t achieved values even lower than one during the epidemic's course, this was not able to control the outbreak in any part of the country.

Figure 2 shows the municipalities ranked by the cumulative number of cases per 10,000 inhabitants, from the least to the most infected one until the 60th day of the epidemic, and also ranked by the proportion of fatal cases, i.e., the number of deaths divided by the number of cases. In Fig. 2a, we can see the cumulative cases for all infected individuals (the sum takes into account the number of infected individuals in all age class) obtained from the mathematical model (in green line), and from the reported data (in blue line). Following the dashed grey lines that connect both data, we can compare the simulations with laboratory-confirmed cases of COVID-19. Figure 2b shows the proportion of fatal cases obtained from reported data and from the mathematical model simulations. Pink lines focus measures obtained considering only individuals in age classes older than 50 years. Unfortunately, there is no available data regarding this population to be compared. The average distance between the cities' observed ranks and their simulated ranks is 5.59 ± 6.87 (median = 4) when comparing cases, and 7.93 ± 6.64 (median = 6) when comparing deaths, both in all age classes.

Figure 3 shows the results when control measures are brought to the model, i.e., the percentage of reduction in the number of cases versus the reduction in contact rate, $(1 - \xi)$. The two panels were done for different values of ψ , where ψ is the fraction of infected population tested, supposing late detection ($\tau = 0.5$), $\nu = 0.55$, and the other parameters are given in Table 2. In (a) we have $\psi = 0.1$ and in (b) $\psi = 0.3$. We highlighted four different cities based on their estimated basic reproductive number denoted, R_0 (Table 1); the other cities are displayed in grey lines. We can observe a large variation among control efficacy in the group of municipalities under study, reflecting the country's heterogeneity, especially in inner Brazil. Considering only the four cities highlighted, we can see that reducing the contact rate by 20% cause a variation from 1 to 55% on reducing the number of cases. Moreover, this variation increases when the fraction of the population tested increases.

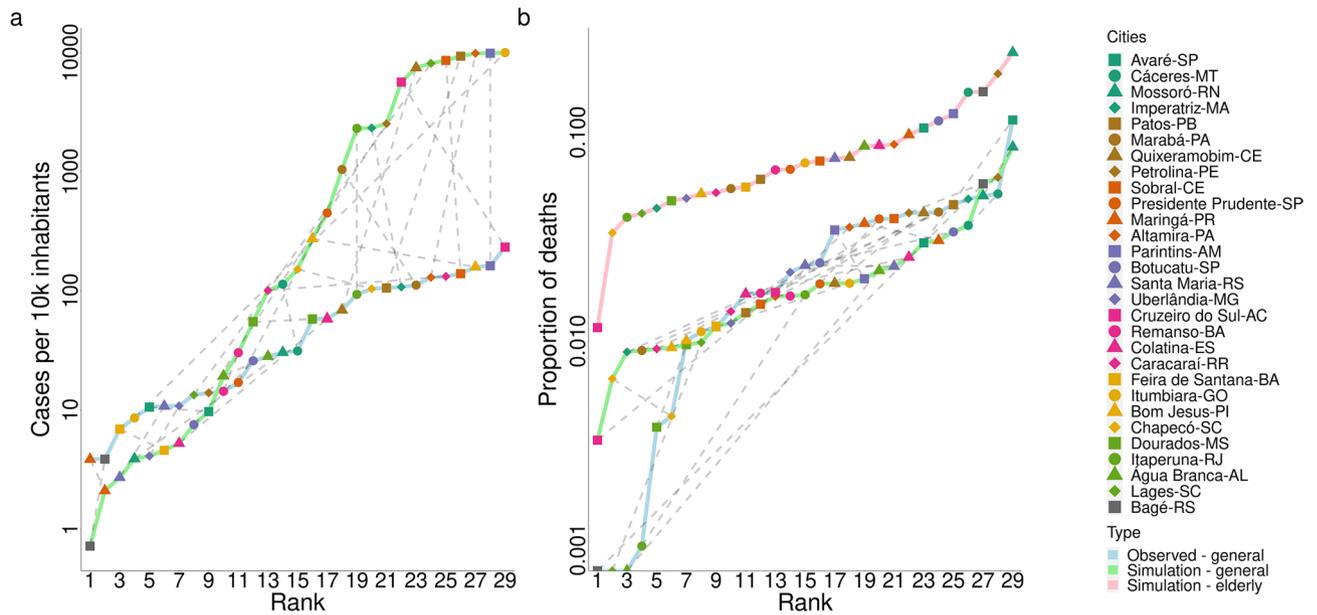


Figure 2. Simulation results and reported data for each municipality on the 60th day of the epidemic. In (a), we have the cumulative number of cases per 10,000 inhabitants versus city’s rank from the least infected to the most infected; in (b), the proportion of fatal cases versus city’s rank. The sum was done from day 1 to 60 of the epidemic course in each city. The first day was chosen to be the one at which the number of infected cases was higher than 10. The dotted grey lines connect the same city in the observed data and in the simulated data to highlight similarity on both results.

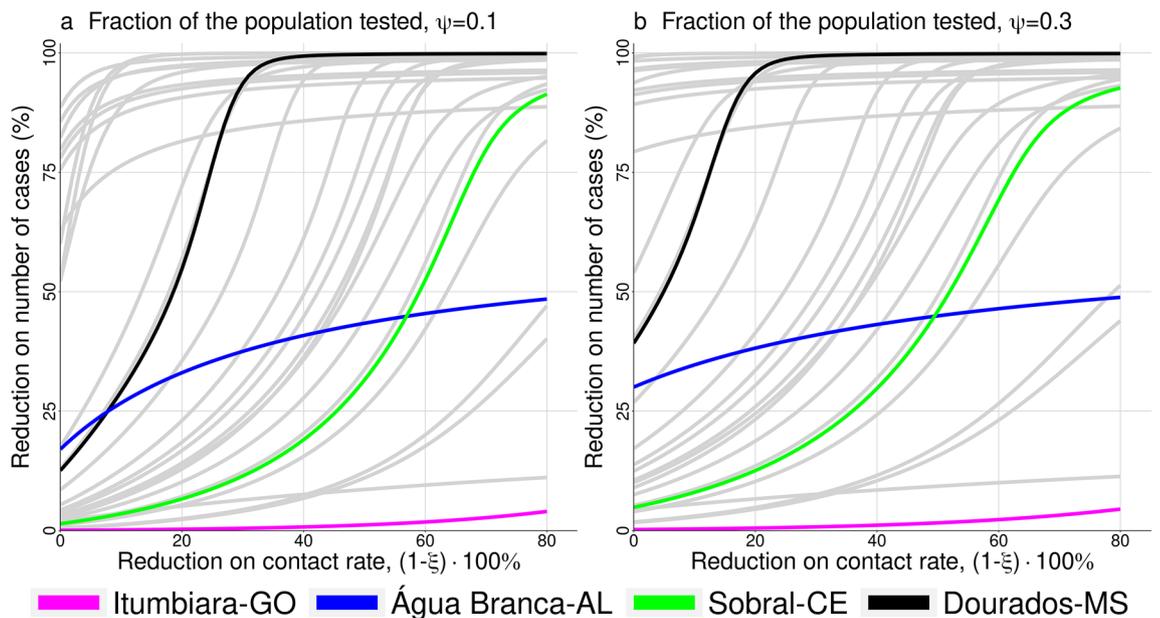


Figure 3. Reduction on the number of cases versus reduction on the contact rate, $1 - \xi$, both in percentage. In (a), $\psi = 0.1$ and in (b), $\psi = 0.3$; where ψ is the fraction of the population tested. Among the 29 municipalities involved in the study, we highlight four of them: Itumbiara, Água Branca, Sobral, and Dourados; the other ones are shown in grey lines.

Figure 4 shows the reduction on the number of cases versus the time of starting control, t_s , in two scenarios that differs by the fraction of tested population, $\psi = 0.1$ and $\psi = 0.3$. The parameters are the same as in Fig. 3 with $(1 - \xi) = 0.6$, and the same cities are highlighted. The efficacy of control diminished as the time of control start is delayed. For some municipalities the reduction in the number of cases is less than 50%. In general, increasing the fraction of the population tested, control efficacy is increased.

The partial rank correlation coefficient (PRCC) obtained from a global sensitivity analysis¹⁶ is shown in Fig. 5. We run 3000 simulations that correspond to different input parameter sets, all of them related to control

Municipality	Temperature	Humidity (%)	Density (inhab/km ²)	Population size inhabitants	R ₀ (estimated)	HDI
Água Branca-AL	23.9	83.3	42.6	19,377	0.80	0.549
Altamira-PA	26.3	85.6	0.6	99,075	4.64	0.665
Avaré-SP	21.3	75.6	68.4	82,934	1.05	0.862
Bagé-RS	18.1	73.4	28.5	116,794	0.46	0.740
Bom Jesus-PI	26.2	61.9	4.1	22,629	1.50	0.668
Botucatu-SP	19.3	67.0	85.9	127,328	1.14	0.800
Cáceres-MT	19.3	75.0	85.9	87,942	1.69	0.708
Caracará-RR	27.2	80.1	0.4	18,398	1.22	0.624
Chapecó-SC	19.3	76.0	293.1	183,530	1.89	0.790
Colatina-ES	25.0	77.5	78.9	111,788	0.99	0.746
Cruzeiro do Sul-AC	25.7	84.5	8.9	78,507	2.71	0.510
Dourados-MS	22.3	77.6	48.0	196,035	1.67	0.747
Feira de Santana-BA	25.2	82.2	416.0	556,642	1.31	0.712
Imperatriz-MA	26.6	80.2	180.8	247,505	2.69	0.731
Itaperuna-RJ	23.3	76.6	86.7	95,841	2.53	0.730
Itumbiara-GO	24.3	72.7	37.7	92,883	6.65	0.752
Lages-SC	16.6	81.1	56.6	156,727	3.44	0.770
Marabá-PA	27.0	83.3	15.4	233,669	2.44	0.668
Maringá-PR	22.9	70.4	733.1	357,077	1.06	0.808
Mossoró-RN	27.7	81.4	123.8	259,815	1.16	0.720
Parintins-AM	27.0	86.3	123.8	102,033	4.67	0.658
Patos-PB	27.2	70.1	212.8	100,674	3.76	0.701
Petrolina-PE	25.4	60.1	64.4	293,962	2.95	0.702
Presidente Prudente-SP	24.0	66.3	368.9	207,610	2.25	0.806
Quixeramobim-CE	26.4	73.3	22.0	71,887	3.03	0.642
Remanso-BA	26.7	68.6	8.3	38,957	1.12	0.579
Santa Maria-RS	19.4	81.3	146.0	261,031	1.03	0.784
Sobral-CE	26.0	85.9	88.7	188,233	3.54	0.714
Uberlândia-MG	22.8	73.9	146.8	604,013	1.33	0.789

Table 1. Municipalities and key factors that may modulate COVID-19 transmission. Each line brings the variables value of the city pointed in the first column. In the case of temperature and humidity the values are the average one observed in April month in each locality³¹. The other factors like density, population size, and Human Development Index (HDI) come from government's website¹³; R₀ are estimated from data³².

measures; the output is control efficacy. The parameters are $(1 - \xi)$, $(1 - \nu)$, ϵ and Ψ , respectively, the reduction of daily contacts among individuals, the reduction in the transmission rate of isolated individuals, the rate at which infected individuals are detected, and the fraction of infected individuals that are identified. As expected, the increase of any control measures promotes the increase of control efficacy, displaying positive values of correlation for any input parameter and the output. Nevertheless, the parameters contribute differently to it as can be seen by its absolute values, which rank them from less to more important (from lower to larger absolute value). We performed the analysis for all four highlighted cities in previous figures, and the results were the same. Here, we display the result for Sobral-CE.

Two different dendrograms were obtained from clustering the municipalities by their similarity, and they are shown in Fig. 6. The first one was built using the model's input data, like the proportion of fatal cases per age group, and the age pyramid. In the second one, we included population density, human development index (HDI), as well as the value of temperature and humidity in April month. One city might belong to different clusters in each dendrogram, and the differences between the dendrograms are highlighted by gray lines connecting both. The distance among groups increased when we re-clustered them; in the new dendrogram, the groups are more dissimilar among them.

Discussion

Our model shows epidemic dynamics for COVID-19 in various cities in inner Brazil. The epidemic's dynamic features on each municipality were modeled by using local and regional epidemiological data, as the value of R₀, the proportion of fatal cases per age group, and sociodemographic data (age pyramid and social contact matrices). Mitigation strategies, such as social distancing of all individuals and detection followed by isolation of infected ones, were tested and compared. The cities were clustered, taking into account several variables that could influence disease transmission among individuals.

At the beginning of the epidemic, a substantial amount of the reported cases are imported cases. The data set does not distinguish between imported and local transmission cases, but the method used to evaluate R_t takes

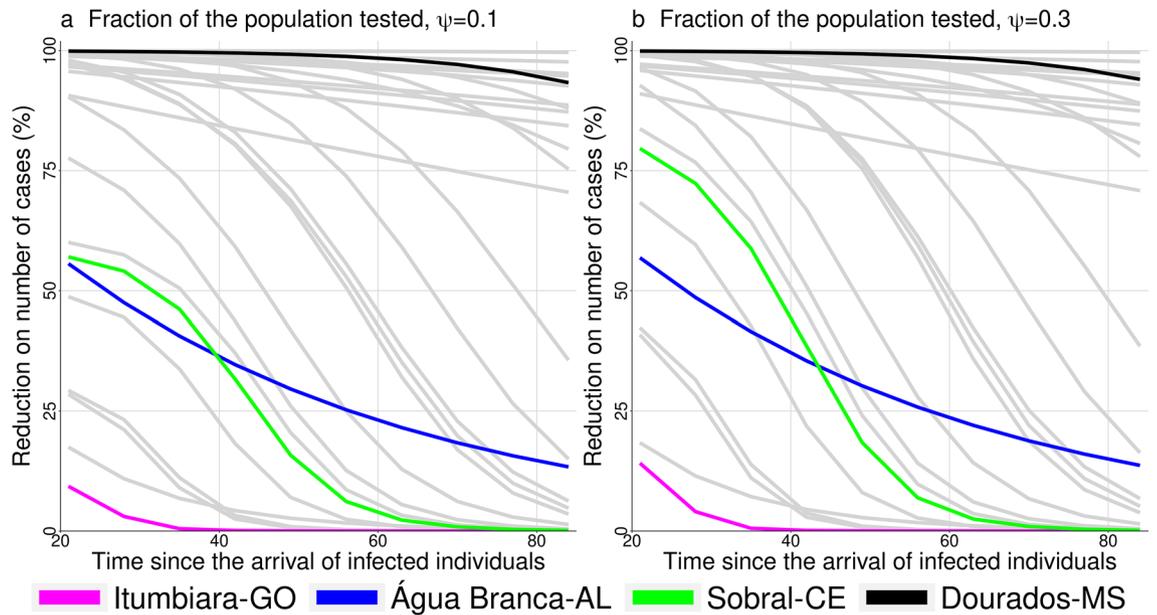


Figure 4. Reduction on the number of cases versus time of starting control. In (a), $\psi = 0.1$ and in (b), $\psi = 0.3$; where ψ is the fraction of the population tested. Among the 29 municipalities involved in the study we highlight four of them: Itumbiara, Água Branca, Sobral, and Dourados; the other ones are shown in grey lines.

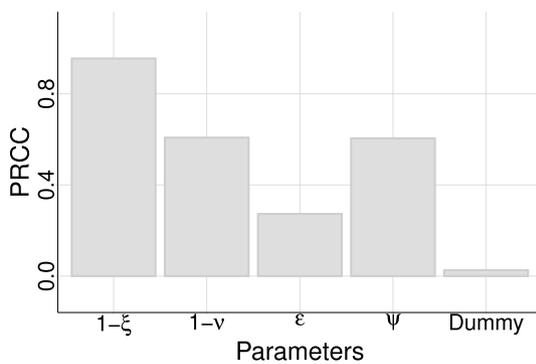


Figure 5. Sensitivity analysis using control efficacy as the output. A negative-control (dummy-parameter) was used to assign a zero value for a sensitivity index. Parameters values below the dummy are considered as not contributing to the model output. The result corresponds to the city of Sobral-CE, but the rank is obtained for the other cities.

into account the rising of infections coming from a local transmission already in course on the population. This can explain the high values seeing at the epidemic’s early stages (Fig. 1 and Table 1). Moreover, the available data displays the date of case report. Ideally, the calculation of the R_t should be performed using the date of symptoms onset. Therefore, the limitation regarding the delay between symptoms onset and reporting must be considered. If we consider that the delay is somehow homogeneous across the country, the R_t calculation is shifted in time. Besides mitigation strategies to halt or diminish disease transmission, deceleration in the initial epidemic’s growth rate can be driven by many factors like heterogeneity in population structure, behavior change of individuals, and increased herd immunity¹⁷.

Ranking the cities by the number of cumulative cases (Fig. 2a), we can see that, in general, the model provide a good prediction for disease behavior (as can be seen in Maringá, Altamira, and Cáceres), being the average distance between the rank of reported data and rank of simulated data of 5.59 ± 6.87 . It is important to highlight that those simulations were made considering no control measures. However, several cities in Brazil had enough time to implement social distancing and preventive measures after the arrival of the first case in São Paulo city, by February 2020. Adherence to social distancing, mask use, and self-isolation has been different across the country, but measure it is yet a challenge. Recently, several works have been trying to connect the transmission rate with mobility index, but a good model that link both measures is still missing¹⁸. This could explain the differences between the model’s prediction and the data collected. Among the cities that are in the top, six have a medium Human Development Index (HDI) and more than 85% of the population on the age classes until 50 years old. At least, regarding to simulations results, the rank follows with a good accuracy the one seen for the R_0 value.

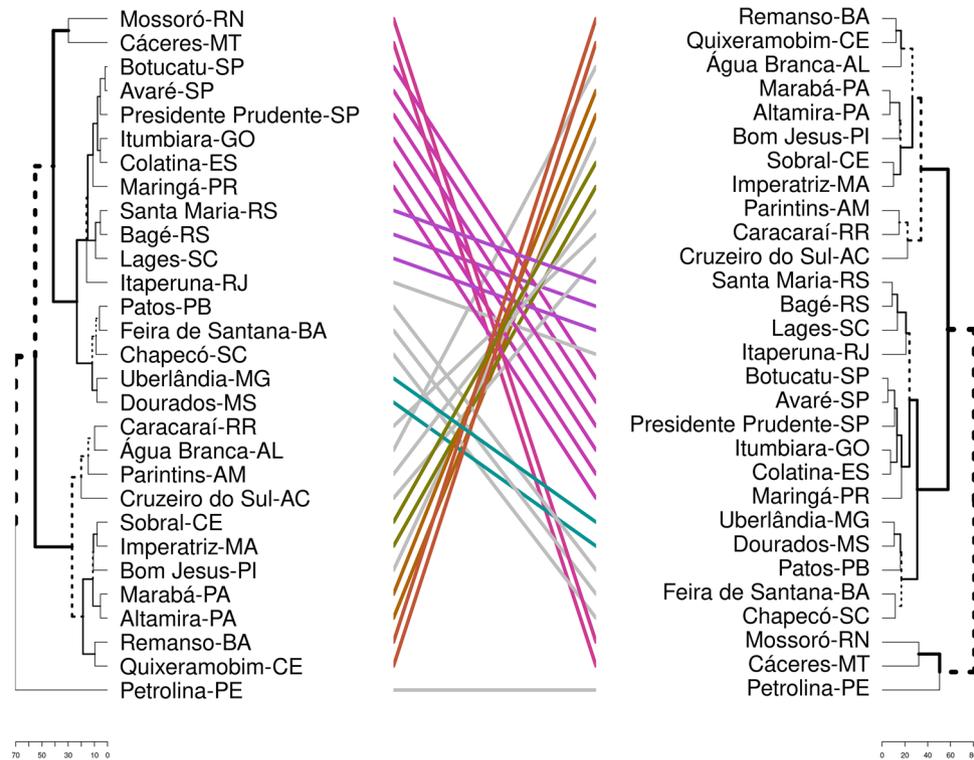


Figure 6. The municipalities are clustered in two ways, from left to right: (1) the proportion of fatal cases per age group and the age pyramid; (2) the same variables plus Human Development Index, population density, temperature, and humidity. The municipalities that changed group because of re-clustering are connected by gray line, while the ones that were kept together are connected through RGB color system.

Further, as it was told, the available data refers to the date of case report. Once we are ranking the cities, if the delay in confirmation is homogeneous across the country, the results related to the cities' rank must not change.

Regarding the proportion of fatal cases (Fig. 2b), since there is no available age-specific data for each city, our simulations use the reported lethality of COVID-19 at state level, which showed to be a good approach for most cities, such as Maringá and Chapecó. However, cities such as Bagé, Lages and Água Branca would have their mortality overestimated since they did not report any death until their 60th day of the epidemic. Here, the average distance between the rank of reported data and simulated data rank is 7.93 ± 6.64 . Interesting to note that among the cities that are in the top, five of them have very high HDI, and more than 16% of the population on the age classes older than 50 years old. In general, both results (Fig. 2) are in agreement with what is expected, medium HDI and youngest population explain the higher number of cases, while low HDI and older population are associated with higher fatality cases.

Overall, the fact that the model performs better for some localities compared to the other also reflects degrees of heterogeneity of COVID-19 test across the country¹⁹. Moreover, several local social and economic features can modulate the chance of death, not to forget access to health services and hospitals might be an important issue in each region of Brazil. Comparing the average distance between the rank of reported data (all age classes) with the rank of simulated data (only age classes older than 50 years old), we get 6.69 ± 6.29 for fatal cases, which is a better result when compared to the previous one. We hypothesize that as the disease impacts the older age classes strongly, they are the responsible for most of the death and, therefore, the mortality rate might be consistent within the state.

It is expected that any kind of control on disease transmission will affect the epidemic's course by delaying and reducing its peak. The gain on smaller numbers of infected individuals during the course of the epidemic is obtained by increasing its duration. Since there is no broad available vaccine, mitigation strategies rely on social distancing, isolation of infected individuals, self-isolation when you are a suspected case, mandatory quarantine applied to all populations, and travel restrictions^{20,21}. So, we drew scenarios with different strategies and interventions. We can clearly see that we have an optimal control measure for each city, depending on the target. Hypothetically, let us consider that a reduction of 60% on the number of cases is needed to avoid the epidemic's critical outcomes, such as health-system collapse. As we can see in Fig. 3, Itumbiara would not reach the needed reduction, probably because it has a high R_0 ($= 6.65$). In that case, it would be necessary to increase even more the control efforts. For Sobral and Dourados, both cities would reach the reduction on the number of infections, but with different control intensity, around 22% for Sobral and 63% for Dourados. This happens because Dourados has a lower R_0 than Sobral, 1.67 and 3.54 respectively. Água Branca is one particular case in which $R_0 < 1$, and this explains why control measures seem to be less useful. The variability of control efficacy

is associated to country's heterogeneity that may be quantified by its mean temperature (from 12° to 27° along the country), its population density (from 2.66 to 67.77 inhab/km²), its human development index (from 0.450 to 1), and many other factors²². Eilersen and Sneppen discussed the cost-benefit of limited isolation and testing in COVID-19 mitigation²³. Using an agent-based epidemiological model, they could compare several scenarios related to mitigation strategies such as testing and quarantining and concluded that this is much cheaper in terms of lost workdays than an extended lockdown. Also, the effect of quarantine on disease dynamics increases when testing is more widespread.

The effect of delaying the start of control measures was modeled as well (see Fig. 4). Again, the result shows a specific pattern at each municipality, but they all have in common one fact: the earlier cities start control, the greater is the reduction in the number of cases. Testing more people in the first 30 days is undoubtedly the best choice, and testing more people may also allow delaying social distance. Since the introduction of SARS-CoV-2 in Brazil, public laboratory certification for the molecular diagnosis of COVID-19 ranged from four laboratories to twenty-six in eight weeks. One can notice laboratories capacity is also increasing on time. This decrease the time of virus detection over the country, but in a heterogeneous way since there is a geographic concentration of laboratories in São Paulo state¹⁹. Amaku *et al.*²⁴ implemented a modified version of the classical SEIR compartmental model to compare two different test-trace-and-quarantine strategies to control the COVID-19 outbreak in the State of São Paulo, Brazil: indiscriminately testing the entire population of the State, and testing only symptomatic cases and their immediate contacts. They concluded that the second one is the most cost-effective strategy, and it can be applied especially in situations where social distancing is challenging to implement. Moreover, if the State of São Paulo had decided to adopt this strategy early, on April the 1st, it would have been possible to reduce the total number of cases by 90%.

The sensitivity analysis ranks the importance of parameters on control efficacy, which is (decreasing order): the reduction in the contact rate of the entire population due to control measures ($1 - \xi$), the reduction in the contact rate of isolated individuals ($1 - \nu$), the fraction of infected individuals that are identified (ψ) and the rate of testing (ϵ), highlighting the importance of mandatory isolation and testing individuals for COVID-19 (see Fig. 5). Combining isolation of detected COVID-19 positive cases with social distancing can provide an efficient way of halting or diminishing disease incidence on population, but the control effectiveness will depend on each municipality's characteristic. In Brazil (and other low-to-medium income countries), the expected peak of the disease was never observed; instead, it achieved a plateau sustained by a pattern of dispersion from major metropolitan areas to the interior²². Each state decides how to deal with a non controlled disease and an economy that may not support non-pharmacological control measures anymore. In São Paulo's case, the territory is divided into seventeen health departments (DRS, in Portuguese) with respect to epidemiological control. Since the beginning of June, the state decided to adopt a reopening plan - that brings back people mobility and non-essential services - which can be more restrictive or more flexible, considering the growth rate of COVID-19 cases and deaths, and bed occupancy rates in each DRS. The same restrictive measures rule all cities belonging to a DRS, that can be adapted in response to the temporal-spatial behavior of the epidemic²⁵.

We sustain the hypothesis that each city must be individually studied. However, it is possible to cluster cities (as it has been done in São Paulo state), considering similar characteristics, which ends up showing patterns of epidemic dynamics. In vast and heterogeneous countries like Brazil, we expect that many factors, such as population density, temperature, and mobility, modulate disease transmission. Quantifying and identifying such contributions can help governments to make decisions about mitigation strategies. The knowledge about other respiratory infection diseases that assault the population in different parts of Brazil, such as Influenza, can provide a pool of important information useful to forecasting COVID-19 in many municipalities.

Following this idea, in Fig. 6, the municipalities are clustered in two different ways. In the dendrogram on the left, we clustered cities by similar characteristics included in the model: the proportion of fatal cases by age group and age pyramid. Following the dendrogram we can identify three big groups: (I) Mossoró, Cáceres, Botucatu, Avaré, Presidente Prudente, Itumbiara, Colatina, Maringá, Santa Maria, Bagé, Lages, Itaperuna, Patos, Feira de Santana, Chapecó, Uberlândia, Dourados; (II) Caracará, Água Branca, Parintins, Cruzeiro do Sul, Sobral, Imperatriz, Bom Jesus, Marabá, Altamira, Remanso, Quixeramobim; (III) Petrolina. In each group, the average age and the average mortality rate are, respectively, 32.7 ± 1.5 and 0.019 ± 0.012 , 27.7 ± 1.8 and 0.033 ± 0.015 , 28 and 0.030 . Observe that groups II and III are very similar (when we compared them by the average values of age and mortality rate). The dendrogram on the right was generated including the cited characteristics plus new ones: temperature, humidity, population density, and HDI. This was done as an exercise to illustrate that we can add or remove characteristics from the clusters in order to find patterns, but it is essential to know which one of these characteristics is important on disease dynamic. For instance, comparing both dendrogram and Fig. 2, the re-clustering added Petrolina to the group of Cáceres and Mossoró, that display a similar number of cases and proportion of deaths; and Água Branca is set together with Quixeramobim and Remanso, being its number of cases between the number of cases of these two cities. Since the first clustering, Itumbiara, Água Branca, Sobral, and Dourados belong to different sub-groups and, therefore, have quite a different epidemic behavior. But, in the second clustering, the distance between Itumbiara and Dourados increases, while the distance between Água Branca and Sobral did not change too much, in accordance with Fig. 2. Moreover, Botucatu and Avaré belong to the same cluster and follow a similar epidemic evolution pattern. In summary, the two main groups that are identified can be distinguish by the HDI of their cities and average age of citizens. This emphasizes the statement that models that include, in some way, temperature, humidity, HDI and population density may better reflect the reality. This can spotlight groups of cities where it is expected that the control efficacy and the disease growth are similar. The results are sustained by Costa *et al.*²⁶ that used a stochastic metapopulation model, inter-municipality mobility, and hypothetical mitigation scenarios, and showed that the diversity of outcomes related to the disease transmission in Brazil is observed in several geographical scales.

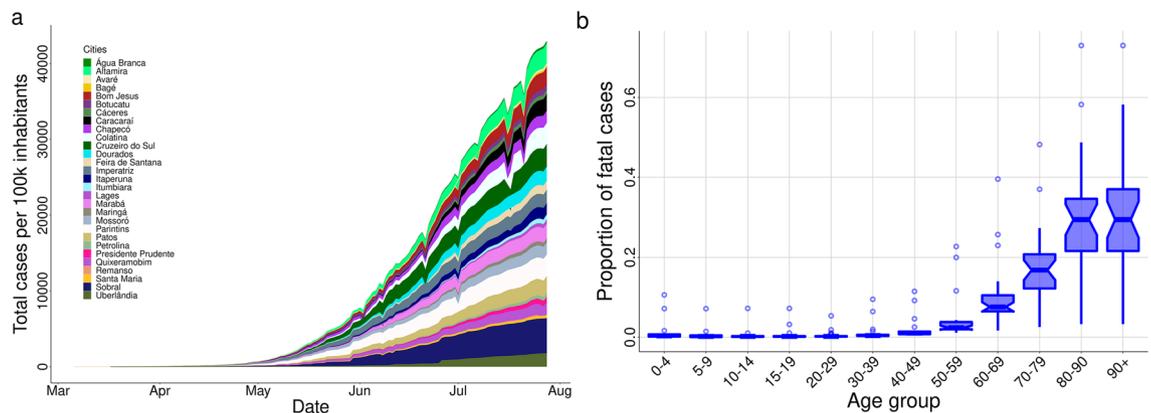


Figure 7. In (a), temporal evolution of the cumulative number of reported cases in each municipality; in (b), the boxplot of the proportion of reported fatal cases for different age groups in twenty Brazilian states enrolled in the study through their municipalities.

Like any other model, the approach developed here has its limitations. The fact that control measures may change the disease dynamic by decreasing or increasing its velocity of spreading may jeopardize model prediction. Also, spatial heterogeneity and social inequalities were not considered in the model, but it is known that in cities belonging to the higher urban hierarchy, COVID-19 spread first among the medium and high level of social classes, and afterward, it achieves the low social classes. Mitigation strategies, such as social distance and shelter-in, do not work for self-employed and low-income families, and to consider it would bring more complexity to the model. Moreover, the available data has two relevant limitations: testing is limited to symptomatic cases who seek health services, and the only date available in the data set is the date of the case report. No other relevant dates - i.e., exposure, onset, or laboratory confirmation - are available. These data limitations impact results diminished if we consider that the bias is homogeneous across the country. The bias generated by the under-reported data impacts our model's parameter estimation, since the diagnosis capacity, compared to the number of cases in the population, changes over time. Recent works have demonstrated and have argued that the delay into case reports and the mitigation strategies may directly impact the R_t estimation^{27,28}. With more information about the available data and complex models, fitting the model parameters to the epidemic curve would be an interesting approach worthy of study.

However, here we were able to show that different control measures should be taken for different cities and, most importantly, each city may have an optimal combination of social distance with testing and isolating positive cases that control the epidemic's curve and permit the health systems to be prepared for the peak of the number of cases. Cities in inner Brazil, such as Cruzeiro do Sul-AC, Imperatriz-MA, Altamira-PA, Bom Jesus-PI, and Parintins-AM that are clustered together, are susceptible to a delay in the arrival of the infections, and epidemic, which may decrease people's risk perception and enhance the disease spreading²⁹. As a consequence, those cities display a larger number of cases per number of inhabitants. We suggest the authorities to give special attention to those cities and perform an extensive educational campaign in order to control the infection. Our results also showed that testing and isolating people could perform a massive difference in controlling the epidemic. Due to a limited number of tests in Brazil, they have been mostly performed to confirm symptomatic cases, without a strategy of contact tracing. This plan should be revised, in accordance with other works³⁰.

By a mathematical model and clustering cities, we suggest patterns of the evolution of the number of cases and control strategies for COVID-19 epidemic. As testing is a major issue for many nations at this moment of the pandemic, social distance in different degrees should be established.

Methods

Municipalities. We aimed at a study capable of representing most small and medium cities of Brazil. Therefore, we decided to choose representative municipalities, with regional importance, from different states and regions, with varied population density, temperature, humidity, human development index (HDI), as well as age structure. From the North region we have: Altamira-PA, Marabá-PA, Cruzeiro do Sul-AC, Parintins-AM, Caracará-RR; from the Northeast region: Água Branca-AL, Sobral-CE, Quixeramobim-CE, Bom Jesus-PI, Imperatriz-MA, Mossoró-RN, Patos-PB, Petrolina-PE, Feira de Santana-BA, Remanso-BA; from the Central-West region: Dourados-MT, Cáceres-MT, Itumbiara-GO; from the South region: Santa Maria-RS, Bagé-RS, Lages-SC, Chapecó-SC, Maringá-PR; and from the Southeast region: Uberlândia-MG, Avaré-SP, Botucatu-SP, Colatina-ES, Itaperuna-RJ, Presidente Prudente-SP. Figure 7 shows each one's geographic location on a Brazil map, with a heatmap showing the interpolation result (distance weighted interpolation) of the total number of cases per 100,000 inhabitants in those cities recorded on 28th July 2020. Table 1 summarizes some information about the cities listed in the present study. In particular, temperature and humidity correspond to the average values in April month³¹.

Effective reproduction number. We calculated the effective reproduction number (R_t) for all chosen cities using the method proposed by Wallinga *et al.*¹⁵ and data of daily incidence of cases ($b(t)$), obtained from

epidemic reports from each municipality³². We considered a simple susceptible-exposed-infected-recovered model with an average latent period (η^{-1}) of 3.0 days and an infectious period (τ^{-1}) of 6.4 days, as well as the COVID-19's mortality rate reported for each state (σ). The rates of leaving the exposed and infectious classes are denoted by $s_1 = \eta + \mu$ and $s_2 = \tau + \mu + \sigma$, where μ^{-1} denote the life expectancy for Brazil. Therefore, the generation interval distribution ($g(t)$) is the combination of two exponential distributions $s_1 e^{-s_1 t}$ and $s_2 e^{-s_2 t}$ given by³³

$$g(t) = \sum_{i=1}^2 \frac{s_1 s_2 e^{s_i t}}{\prod_{j=1, j \neq i}^2 (s_j - s_i)} \quad \text{with } t \geq 0.$$

The duration of a generation interval is thereby implicitly specified as an exponential distribution with mean $T_c = 1/s_1 + 1/s_2$. The expression above is valid when the infection force, Λ , satisfies the following inequality $\Lambda > \min(-s_1, -s_2)$. Also, as we are dealing with a distribution, we need to normalize $g(t)$. Using this equation we can evaluate R_t as

$$R_t = \frac{b(t)}{\int_0^\infty b(t-a)g(a)da} \quad \text{with } \int_0^\infty g(t)dt = 1.$$

The R_0 of each city was considered to be the average value of R_t in the second week of the epidemic in the city. The first day was considered to be the one in which the cumulative incidence of infections reached ten cases. This choice was taken to guarantee that a local transmission was established in the city. We performed a spline interpolation and 7-day moving average on the data before used it to estimate R_t . The average value of R_t at each calendar day from April to August can be seen in Fig. 1. It also displays the value obtained for each municipality (in grey points). Outliers were omitted from this plot.

Clustering. After listing the cities, we clustered them in order to search for patterns. By taking each city as a model, studying the main characteristics, and crossing into a cluster study, we believe it is possible to extrapolate this study's results to other cities that are not plotted here. We first grouped cities by their proportion of fatal cases per age group and age pyramid. Afterward, we added population density, temperature, humidity, HDI index and clustered them again. We used a hierarchical agglomerative clustering method, combining cluster threw the complete linkage criterion and Manhattan distance as a metric to measure dissimilarity between the observation sets³⁴. The result is shown in Fig. 6.

Data availability. Time series of the number of cases for each municipality in Brazil is not reported on any official government's website. The Federal government does not provide it for open use. Therefore, we used daily cases reported on open sources in Brazil provided by a task force of volunteers (researchers and reporters) that compile the daily epidemiological reports of each state³². We used confirmed COVID-19 cases in the analysis, whose data refers to the date of case report and only mild and severe cases appear in this database (hospitalized cases and people seeking for medical assistance and health services).

Moreover, other issues may influence as well, such as the turnaround time of the performed tests and the fact that the data set does not distinguish between imported and autochthonous cases. To avoid the delay in reporting, we removed the last two weeks of data at the moment of the analysis. However, sub-notification is an issue that is difficult to be handle. Supposing that those issues occur in a homogeneous way throughout the country, we expected that the results would be impacted only by a scale factor, but keeping the conclusions regarding the temporal pattern of COVID-19 cases in each city. The age-dependent mortality is available separately in the epidemic's reports from each state, but not for each city. In this work, we used data from 20 different states from Brazil to simulate 29 different cities. For each city, the reported time series of cases per 100 thousand inhabitants are shown in Fig. 8a, while the proportion of fatal cases in each age group is shown in Fig. 8b.

Mathematical model. The proposed model is an age-structured one that divides the human population into fifteen age groups: 0 to 4 years, five years interval from 5 to 70 years, and greater than 70 years³⁵. The variables of the model are $t, S_i := S_i(t), E_i := E_i(t), I_i := I_i(t), Q_i := Q_i(t), R_i := R_i(t)$; respectively, time, susceptible, exposed, infected, detected and isolated infected individuals, and recovered one. The index i is the age class. The natural mortality rate μ appears in all age classes, and from 1 to 15, the parameter α_i takes into account the transition among them. Individuals are born susceptible, and they become exposed, when contacting infected or isolated individuals at rate β_1 and $\beta_2 = \nu\beta_1$ ($\nu \in [0, 1]$), respectively. The parameter $c_{i,j}$ represents the fraction of daily contacts that age group i has with age group j ³⁶. Target control can be done by varying $\xi_i \in [0, 1]$, being $\xi_i = 0$ complete protection of class i and $\xi_i = 1$ no protection of class i against the infection. After a period of time η^{-1} exposed individuals becomes infectious. At rate ϵ , a fraction $\psi \in [0, 1]$ of infected individuals are identified and isolated. Additional mortality related to the disease is considered in the compartments of infected and isolated individuals, σ_i . Finally, these individuals become recovered at rates γ and τ . The ODE model is given by

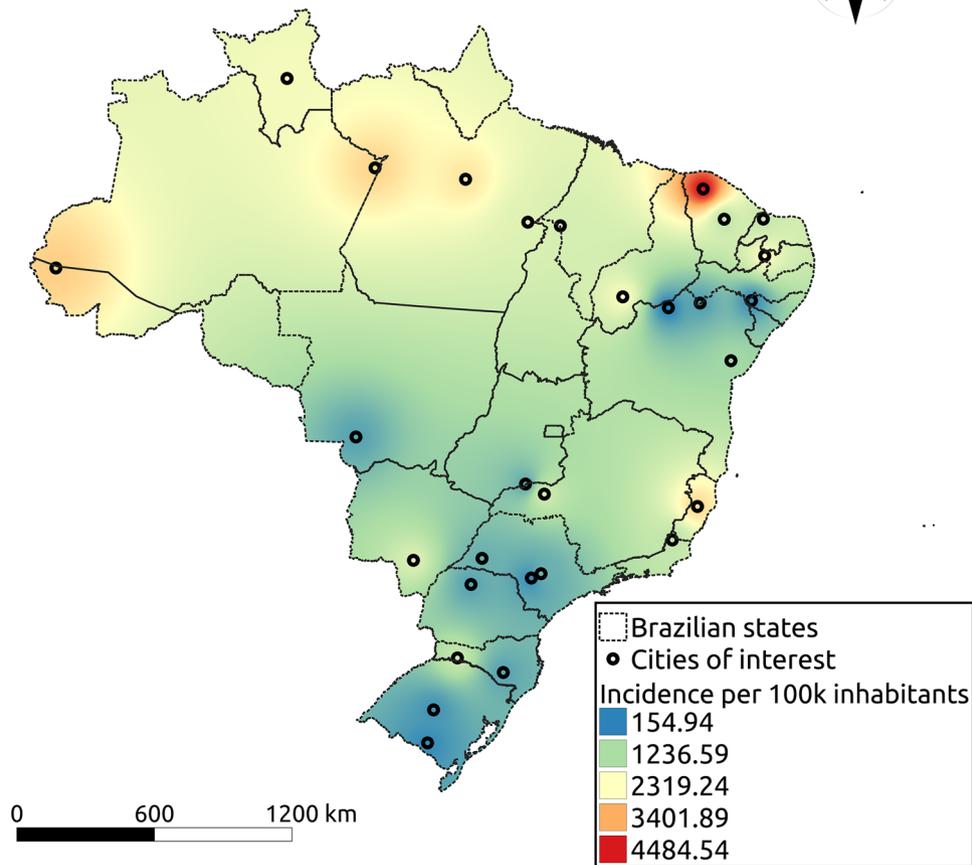


Figure 8. The geographic location of the municipalities enlisted in the study. The heatmap shows the interpolation result of the total number of cases per 100 thousand inhabitants in those cities recorded on July 28th. Cool colors mean less infected individuals while warm colors more infected individuals, and the scale goes from 153.6 (blue) to 4617.7 (red) cases per 100 thousand inhabitants. The cities were geocoded using the software Qgis (v3.10), and the interpolation was performed using the software’s tool for Inverse Distance Weighted Interpolation (<https://www.qgis.org/en/site/>).

$$\begin{aligned}
 \frac{dS_{i+1}}{dt} &= \mu N \delta_{i+1,1} + \alpha_i S_i - \left(\beta_1 \sum_{j=1}^{15} c_{i+1,j} \frac{I_j}{n_j} + \beta_2 \sum_{j=1}^{15} c_{i+1,j} \frac{Q_j}{n_j} \right) \xi_{i+1} S_{i+1} - (\mu + \alpha_{i+1}) S_{i+1} \\
 \frac{dE_{i+1}}{dt} &= \alpha_i E_i + \left(\beta_1 \sum_{j=1}^{15} c_{i+1,j} \frac{I_j}{n_j} + \beta_2 \sum_{j=1}^{15} c_{i+1,j} \frac{Q_j}{n_j} \right) \xi_{i+1} S_{i+1} - (\mu + \alpha_{i+1} + \eta) E_{i+1} \\
 \frac{dI_{i+1}}{dt} &= \alpha_i I_i + \eta E_{i+1} - (\sigma_i + \mu + \alpha_{i+1} + \gamma + \varepsilon \psi) I_{i+1} \\
 \frac{dQ_{i+1}}{dt} &= \alpha_i Q_i + \varepsilon \psi I_{i+1} - (\sigma_i + \mu + \alpha_{i+1} + \tau) Q_{i+1} \\
 \frac{dR_{i+1}}{dt} &= \alpha_i R_i + \tau Q_{i+1} + \gamma I_{i+1} - (\mu + \alpha_{i+1}) R_{i+1}
 \end{aligned}
 \tag{1}$$

with $i = 0, \dots, 14$, $\alpha_0 = \alpha_{15} = 0$, $\alpha_1 = \dots = \alpha_{14} = \alpha$, $\delta_{1,1} = 1$, and $\delta_{j,1} = 0$ with $j \neq 1$. Besides, $n_j = S_j + E_j + I_j + Q_j + R_j$, and $N = \sum_{i=1}^{15} n_j$ at $t = 0$. Table 2 summarizes model parameters, their description, range of values and units^{37,38}. Figure 9 shows the diagram of the compartmental model.

Defining $\bar{S} = S/N$ we can rewrite (1) as

Parameter	Description	Value
μ	Mortality rate	1/75 years ⁻¹
σ	Additional mortality rate	[0,0, 0,20]
α	Transition rate among age classes	1/5 years ⁻¹
η^{-1}	Latent period	3 days
γ^{-1}	Infectious period	6.4 days
τ^{-1}	Isolation period	{1, 2, 5, 6} days
ϵ	Detection and isolation rate	1/3 days ⁻¹
ψ	Fraction of infected that are detected	[0, 1]
ξ, ν	Reduction on the infection transmission	[0, 1]
β_1	Transmission rate	[0.4397, 0.4782] days ⁻¹
β_2	Transmission rate	[0.241835, 0.26301] days ⁻¹

Table 2. Parameters of the model, their values (or range of values) and units^{37,38}.

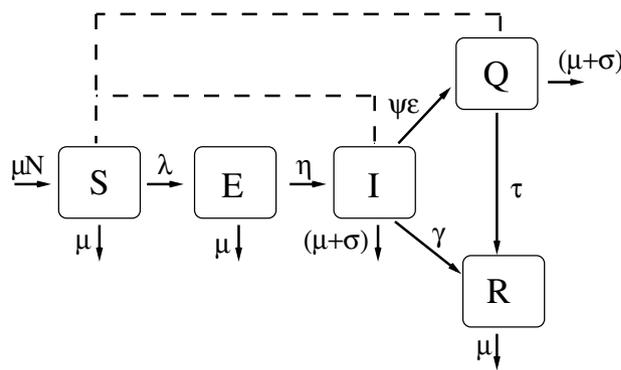


Figure 9. The variables of the model are susceptible (S), exposed (E), infected(I), isolated (Q) and recovered individuals (R). The continuous line indicates transitions between compartments and the dashed line indicates interactions between compartments that contributes to the infection force, λ . The model’s parameters are described at Table 2.

$$\begin{aligned}
 \frac{d\bar{S}_{i+1}}{dt} &= \mu \delta_{i+1,1} + \alpha_i \bar{S}_i - \left(\beta_1 \sum_{j=1}^{15} c_{i+1,j} \frac{\bar{I}_j}{\bar{n}_j} + \beta_2 \sum_{j=1}^{15} c_{i+1,j} \frac{\bar{Q}_j}{\bar{n}_j} \right) \xi_{i+1} \bar{S}_{i+1} - (\mu + \alpha_{i+1}) \bar{S}_{i+1} \\
 \frac{d\bar{E}_{i+1}}{dt} &= \alpha_i \bar{E}_i + \left(\beta_1 \sum_{j=1}^{15} c_{i+1,j} \frac{\bar{I}_j}{\bar{n}_j} + \beta_2 \sum_{j=1}^{15} c_{i+1,j} \frac{\bar{Q}_j}{\bar{n}_j} \right) \xi_{i+1} \bar{S}_{i+1} - (\mu + \alpha_{i+1} + \eta) \bar{E}_{i+1} \\
 \frac{d\bar{I}_{i+1}}{dt} &= \alpha_i \bar{I}_i + \eta \bar{E}_{i+1} - (\sigma_i + \mu + \alpha_{i+1} + \gamma + \epsilon \psi) \bar{I}_{i+1} \\
 \frac{d\bar{Q}_{i+1}}{dt} &= \alpha_i \bar{Q}_i + \epsilon \psi \bar{I}_{i+1} - (\sigma_i + \mu + \alpha_{i+1} + \tau) \bar{Q}_{i+1} \\
 \frac{d\bar{R}_{i+1}}{dt} &= \alpha_i \bar{R}_i + \tau \bar{Q}_{i+1} + \gamma \bar{I}_{i+1} - (\mu + \alpha_{i+1}) \bar{R}_{i+1}
 \end{aligned}
 \tag{2}$$

with $\bar{n}_j = \bar{S}_j + \bar{E}_j + \bar{I}_j + \bar{Q}_j + \bar{R}_j$. The disease free equilibrium is given by

$$P_0 = (S_1^*, 0, 0, 0, 0, \dots, S_{15}^*, 0, 0, 0, 0)$$

where

$$S_1^* = \frac{\mu}{\mu + \alpha_1}, S_i^* = \frac{\alpha_i}{\mu + \alpha_i} \quad \text{with } i \in 2, \dots, 15.$$

In order to obtain the next generation matrix^{39,40}, we used the reduced system given, in its vectorial form, by

$$\begin{aligned}
 \frac{d\bar{E}}{dt} &= \mathcal{A}\bar{E} + \text{diag}(\beta_1\xi) \text{diag}(\bar{S}) \mathcal{C} \text{diag}^{-1}(\bar{n})\bar{I} + \text{diag}(\beta_2\xi) \text{diag}(\bar{S}) \mathcal{C} \text{diag}^{-1}(\bar{n})\bar{Q} - \text{diag}(\mu + \alpha + \eta)\bar{E} \\
 \frac{d\bar{I}}{dt} &= \mathcal{A}\bar{I} + \text{diag}(\eta)\bar{E} - \text{diag}(\sigma + \mu + \gamma + \epsilon\psi + \alpha)\bar{I} \\
 \frac{d\bar{Q}}{dt} &= \mathcal{A}\bar{Q} + \text{diag}(\epsilon\psi)\bar{I} - \text{diag}(\sigma + \mu + \alpha + \tau)\bar{Q}.
 \end{aligned}
 \tag{3}$$

Bold symbols represent vectors as $\mathbf{x} = [x_1, \dots, x_{15}]^T$ and $\text{diag}(\mathbf{x})$ represent diagonal matrices, $M = [m_{ij}]$, in which $m_{ii} = x_i$. \mathcal{C} is the contact-distribution matrix among the age groups⁴¹, and

$$\mathcal{A} = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ \alpha_1 & 0 & 0 & \dots & 0 & 0 \\ 0 & \alpha_2 & 0 & \dots & 0 & 0 \\ 0 & 0 & \alpha_3 & \ddots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \alpha_{14} & 0 \end{pmatrix}.$$

The matrix of infection terms, \mathcal{F} , and the matrix of transition terms, \mathcal{V} , are given, respectively, by

$$\mathcal{F} = \begin{pmatrix} 0_{15 \times 15} & \text{diag}(\beta_1\xi) \text{diag}(\bar{S}^*) \mathcal{C} \text{diag}^{-1}(\bar{n}) & \text{diag}(\beta_2\xi) \text{diag}(\bar{S}^*) \mathcal{C} \text{diag}^{-1}(\bar{n}) \\ 0_{15 \times 15} & 0_{15 \times 15} & 0_{15 \times 15} \\ 0_{15 \times 15} & 0_{15 \times 15} & 0_{15 \times 15} \end{pmatrix}$$

and

$$\mathcal{V} = \begin{pmatrix} \text{diag}(\mu + \alpha + \eta) - \mathcal{A} & 0_{15 \times 15} & 0_{15 \times 15} \\ -\text{diag}(\eta) & \text{diag}(\sigma + \mu + \gamma + \alpha + \epsilon\psi) - \mathcal{A} & 0_{15 \times 15} \\ 0_{15 \times 15} & -\text{diag}(\epsilon\psi) & \text{diag}(\sigma + \mu + \alpha + \tau) - \mathcal{A} \end{pmatrix},$$

in which $\bar{S}^* = \mu [\text{diag}(\mu + \alpha) - \mathcal{A}]^{-1} \delta$, with $\delta = [1, 0, \dots, 0]^T$, is the disease-free equilibrium of (2). The basic reproductive number denoted by R_0 is given by the spectral radius of the next generator operator matrix given by $\mathcal{F}\mathcal{V}^{-1}$ (i.e. its dominant eigenvalue). The disease-free equilibrium \bar{S}^* is locally asymptotically stable if $R_0 < 1$, and unstable if $R_0 > 1$. R_0 is the mean number of secondary cases that a primary case generates in a whole susceptible population, which implies before control measures. A simple and direct way to calculate the effort to be done to control an epidemic is given by $P_c = 1 - 1/R_0$, where P_c is the fraction of population that likely to be infected without mitigation. This represents the worst scenario since the deterministic approach has several assumptions like large population, well-mixed individuals, and no spatial structure.

Simulations. In all simulations, the parameter β_1 was calibrated, for a given R_0 (Table 1), using the next-generation matrix, under no control measure. The addition mortality rates (days⁻¹) are calculated through the expression

$$\sigma_i = -\gamma \ln(1 - p_i), \quad i = \{1, \dots, 15\}$$

where p_i is the probability that an individual at age group i dies during their infectious period. For each city, we used data reported from their states to estimate p_i (see Fig. 8b).

The simulations start with ten infected individuals (in the age class of 25 to 50 years) introduced in a wholly susceptible population. Control started later, after one month since the introduction of infected individuals. Control was explored by reducing contact rate among age classes (using the parameter ξ), decreasing the time of detection of infected individuals (ϵ^{-1}), increasing the fraction of individuals that are detected and isolated (ψ), and decreasing the contribution of detected and isolated individuals to the disease transmission (ν).

Two different scenarios were analyzed. The first one deals with a situation where the detection and isolation of infected individuals occur quickly. Therefore, we set up ϵ^{-1} to 1 and 2 days and $\tau^{-1} = 6$ ($\approx \gamma^{-1}$) days. The second one suppose that detection takes longer time, then ϵ^{-1} ($\approx \gamma^{-1}$) was set up to 5 and 6 days and $\tau^{-1} = 2$ days. The other parameters are $\beta_2 = 0.55\beta_1$ days⁻¹, $\eta^{-1} = 3$ days, $\gamma^{-1} = 6.4$ days, and $\mu = 3.65 \times 10^{-5}$ days. In general, figures were done with the set of parameters that represent the late detection.

Since the time of starting control impacts the evolution of disease transmission, the efficacy of control was measured varying this parameter in the simulation. For this, we measure the reduction (in percentage) on the number of infected individuals with and without control. Target and no target control over higher age classes was explored by ranking and comparing the municipalities by the cumulative number of infected individuals, and by the proportion of lethal cases. Finally, a sensitivity analysis based on partial rank correlation coefficient (PRCC) was done to discuss the contribution of each model control parameter to the control efficacy, measured as the percentage of infected cases that are avoided. The PRCC measures the monotonic relationship between an input parameter and the output variable when the linear effects of other independent variables are discounted¹⁶. The input parameters were ϵ , ξ , ν , and ψ ; and they were sampled using the Latin Hypercube Sampling method. The first one took from a uniform distribution from 0.166 to 0.2 (late detection) and from 0.5 to 1 (early detection), and the others one from an uniform distribution in the range of 0 to 1. A PRCC close to one means that the input parameter and the output are strong and positively related, while negative values stand for negative correlation.

Received: 24 April 2020; Accepted: 6 May 2021

Published online: 24 May 2021

References

- 2019 novel coronavirus (2019-ncov): Strategic preparedness and response plan. <https://www.who.int/docs/default-source/coronavirus/srp-04022020.pdf> (2020).
- Wu, J., Leung, K. & Leung, G. Nowcasting and forecasting the potential domestic and international spread of the 2019-ncov outbreak originating in wuhan, china: a modelling study. *The Lancet* **395**, 689–697 (2020).
- Lau, H. *et al.* The positive impact of lockdown in Wuhan on containing the Covid-19 outbreak in China. *J. Travel Med.* <https://doi.org/10.1093/jtm/taaa037> (2020).
- de Souza Marotta Alfaia, R. G., Costa, A. M. & Campos, J. C. Municipal solid waste in brazil: A review. *Waste Management Res.* **35**, 1195–1209. <https://doi.org/10.1177/0734242X17735375> (2017).
- Gomes, F., Santo, M., Gryscek, R., Bertolozzi, M. & França, F. Access to drinking water and sewage treatment in Brazil: a challenge for the control of waterborne infectious diseases. *Rev. Inst. Med. Trop. SP* <https://doi.org/10.1590/S1678-9946202062071> (2020).
- In urban concentrations of brazil, 61.9% of the residents lived in areas with good or average living conditions. <https://agenciadenoticias.ibge.gov.br/en/agencia-press-room/2185-news-agency/releases-en/18961-in-urban-concentrations-of-brazil-61-9-of-the-residents-lived-in-areas-with-good-or-average-living-conditions> (2017). Accessed 2021-01-06.
- Pereira, R., Nascimento, G., Gratão, L. & Pimenta, R. The risk of covid-19 transmission in favelas and slums in Brazil. *Public Health* **183**, 33–42. <https://doi.org/10.1016/j.puhe.2020.04.042> (2020).
- Boletim epidemiológico - coe covid-19 - ministério da saúde brasil. <https://portalarquivos.saude.gov.br/images/pdf/2020/Abril/21/BE13---Boletim-do-COE.pdf> (2020).
- Candido, D. D. S. *et al.* Routes for COVID-19 importation in Brazil. *J. Travel Med.* <https://doi.org/10.1093/jtm/taaa042> (2020).
- Bootsma, M. & Ferguson, N. The effect of public health measures on the 1918 influenza pandemic in U.S. cities. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 7588–7593. <https://doi.org/10.1073/pnas.0611071104> (2007).
- Grassly, N. & Christophe Fraser, C. Mathematical models of infectious disease transmission. *Nat. Rev. Microbiol.* **6**, 477–487. <https://doi.org/10.1038/nrmicro1845> (2008).
- Enserink, M. & Kupferschmidt, K. With Covid-19, modeling takes on life and death importance. *Science* **367**, 1414–1415. <https://doi.org/10.1126/science.367.6485.1414-b> (2020).
- Instituto brasileiro de geografia e estatística. <https://cidades.ibge.gov.br/> (2020). Accessed 2020-03-03.
- Leung, K., Wu, J., Liu, D. & Leung, G. First-wave covid-19 transmissibility and severity in China outside Hubei after control measures, and second-wave scenario planning: a modelling impact assessment. *The Lancet* [https://doi.org/10.1016/S0140-6736\(20\)30746-7](https://doi.org/10.1016/S0140-6736(20)30746-7) (2020).
- Wallinga, J. & Lipsitch, M. How generation intervals shape the relationship between growth rates and reproductive numbers. *Proc. R. Soc. B* **274**, 599–604. <https://doi.org/10.1098/rspb.2006.3754> (2007).
- Marino, S., Hogue, I., Ray, C. & Kirschner, D. A methodology for performing global uncertainty and sensitivity analysis in systems biology. *J. Theor. Biol.* **254**, 178–196. <https://doi.org/10.1016/j.jtbi.2008.04.011> (2008).
- Park, S., Champredon, D., Weitz, J. & Dushoff, J. A practical generation-interval-based approach to inferring the strength of epidemics from their speed. *Epidemics* **27**, 12–18. <https://doi.org/10.1016/j.epidem.2018.12.002> (2019).
- Peixoto, P., Marcondes, P. C. & Oliva, S. Modeling future spread of infections via mobile geolocation data and population dynamics. An application to covid-19 in Brazil. *PLoS One* **15**, e0235732. <https://doi.org/10.1371/journal.pone.0235732> (2020).
- Grotto, R. *et al.* Increasing molecular diagnostic capacity and Covid-19 incidence in Brazil. *Epidemiol. Infect.* <https://doi.org/10.1017/S0950268820001818> (2020).
- Ebrahim, S., Ahmed, Q., Gozzer, E., Schlagenhauf, P. & Memish, Z. Covid-19 and community mitigation strategies in a pandemic. *BMJ* <https://doi.org/10.1136/bmj.m1066> (2020).
- Anderson, R., Heesterbeek, H., Klinkenberg, D. & Hollingsworth, T. D. How will country-based mitigation measures influence the course of the Covid-19 epidemic?. *The Lancet* **395**, 931–934. [https://doi.org/10.1016/S0140-6736\(20\)30567-5](https://doi.org/10.1016/S0140-6736(20)30567-5) (2020).
- Fortaleza, C., Guimarães, R., de Almeida, G., Pronunciate, M. & Ferreira, C. Taking the inner route: spatial and demographic factors affecting vulnerability to covid-19 among 604 cities from inner são paulo state, brazil. *Epidemiol. Infect.* **148**, 1–5. <https://doi.org/10.1017/S095026882000134X> (2020).
- Eilersen, A. & Sneppen, K. Cost-benefit of limited isolation and testing in Covid-19 mitigation. *Sci. Rep.* <https://doi.org/10.1038/s41598-020-75640-2> (2020).
- Amaku, M. *et al.* Modelling the test, trace and quarantine strategy to control the Covid-19 epidemic in the state of São Paulo, Brazil. *Infect. Dis. Model.* **6**, 46–55. <https://doi.org/10.1016/j.idm.2020.11.004> (2021).
- de Almeida, G. *et al.* Two hundred days of Covid-19 in São Paulo state, Brazil. *Epidemiol. Infect.* <https://doi.org/10.1017/S0950268820002927> (2020).
- Costa, G. S., Cota, W. & Ferreira, S. C. Outbreak diversity in epidemic waves propagating through distinct geographical scales. *Phys. Rev. Res.* **2**, 043306. <https://doi.org/10.1103/PhysRevResearch.2.043306> (2020).
- Rao, A. S. S. *et al.* How relevant is the basic reproductive number computed during Covid-19, especially during lockdowns?. *Infect. Control Hosp. Epidemiol.* <https://doi.org/10.1017/ice.2020.1376> (2020).
- Park, S. W. & Bolker, B. M. A note on observation processes in epidemic models. *Bull. Math. Biol.* **82**, 1–8 (2020).
- Prusaczyk, B. Strategies for disseminating and implementing Covid-19 public health prevention practices in rural areas. *The J. Rural. Heal.* **37**, 142–144 (2021).
- Moghadas, S. M. *et al.* The implications of silent transmission for the control of Covid-19 outbreaks. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 17513–17515 (2020).
- Normais climatológicas do brasil, 1981–2010, instituto nacional de meteorologia (inmet). <http://www.inmet.gov.br/portal/index.php?r=clima/normaisClimatologicas> (2020).
- Open data in Brazil. <https://brasil.io/home/> (2020).
- Akkouchi, M. On the convolution of exponential distributions. *J. Chungcheong Math. Soc.* **21**, 501–510 (2008).
- Johnson, R. & Wichern, D. *Applied Multivariate Statistical Analysis* (Pearson, 2007).
- Keeling, M. & Rohani, P. *Modeling Infectious Diseases in Human and Animals* (Princeton University Press, 2008).
- Arregui, S., Aleta, A., Sanz, J. & Moreno, Y. Projecting social contact matrices to different demographic structures. *PLoS Comput. Biol.* **14**, e1006638. <https://doi.org/10.1371/journal.pcbi.1006638> (2018).
- Byrne, A. W. *et al.* Inferred duration of infectious period of SARS-COV-2: rapid scoping review and analysis of available evidence for asymptomatic and symptomatic Covid-19 cases. *BMJ Open* <https://doi.org/10.1136/bmjopen-2020-039856> (2020).
- Prete, C. A. *et al.* Serial Interval Distribution of SARS-CoV-2 Infection in Brazil. *J Travel Med.* **28**, taaa115 (2020).
- Diekmann, O., Heesterbeek, J. & Roberts, M. The construction of next-generation matrices for compartmental epidemic models. *J. R. Soc. Interface* **7**, 873–885 (2009).
- Van den Driessche, P. & Watmough, J. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Math. Biosci.* **180**, 29–48. [https://doi.org/10.1016/S0025-5564\(02\)00108-6](https://doi.org/10.1016/S0025-5564(02)00108-6) (2002).

41. Mossong, J. *et al.* Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med.* <https://doi.org/10.1371/journal.pmed.0050074> (2008).

Acknowledgements

T.N.V. and C.P.F. thanks support from São Paulo Research Foundation (FAPESP) Grants 18/24811-1 and 18/24058-1.

Author contributions

C.M.C.B.F. and G.B.A. conceived the problem, T.N.V and C.P.F conduct the simulations. All authors analysed the results and reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to G.B.A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021

RESEARCH ARTICLE

The use of health geography modeling to understand early dispersion of COVID-19 in São Paulo, Brazil

Carlos Magno Castelo Branco Fortaleza¹✉*, Raul Borges Guimarães²✉, Rafael de Castro Catão³‡, Cláudia Pio Ferreira⁴‡, Gabriel Berg de Almeida¹‡, Thomas Nogueira Vilches⁵‡, Edmur Pugliesi¹‡

1 Department of Infectious Diseases, Botucatu Medical School, São Paulo State University (UNESP), Botucatu, São Paulo State, Brazil, **2** Department of Geography, Faculty of Science and Technology, São Paulo State University (UNESP), Presidente Prudente, São Paulo State, Brazil, **3** Department of Geography, Federal University of Espírito Santo, Vitória, Espírito Santo State, Brazil, **4** Institute of Biosciences, São Paulo State University (UNESP), Botucatu, São Paulo State, Brazil, **5** Institute of Mathematics, Statistics and Scientific Computation, University of Campinas (UNICAMP), Campinas, São Paulo State, Brazil

✉ These authors contributed equally to this work.

‡ These authors also contributed equally to this work.

* carlos.fortaleza@unesp.br



OPEN ACCESS

Citation: Fortaleza CMCB, Guimarães RB, Catão RdC, Ferreira CP, Berg de Almeida G, Nogueira Vilches T, et al. (2021) The use of health geography modeling to understand early dispersion of COVID-19 in São Paulo, Brazil. PLoS ONE 16(1): e0245051. <https://doi.org/10.1371/journal.pone.0245051>

Editor: Javier Sanchez, Atlantic Veterinary College, CANADA

Received: April 26, 2020

Accepted: December 21, 2020

Published: January 7, 2021

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0245051>

Copyright: © 2021 Fortaleza et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its [Supporting information](#) files.

Abstract

Public health policies to contain the spread of COVID-19 rely mainly on non-pharmacological measures. Those measures, especially social distancing, are a challenge for developing countries, such as Brazil. In São Paulo, the most populous state in Brazil (45 million inhabitants), most COVID-19 cases up to April 18th were reported in the Capital and metropolitan area. However, the inner municipalities, where 20 million people live, are also at risk. As governmental authorities discuss the loosening of measures for restricting population mobility, it is urgent to analyze the routes of dispersion of COVID-19 in São Paulo territory. We hypothesize that urban hierarchy is the main responsible for the disease spreading, and we identify the hotspots and the main routes of virus movement from the metropolis to the inner state. In this ecological study, we use geographic models of population mobility to check for patterns for the spread of SARS-CoV-2 infection. We identify two patterns based on surveillance data: one by contiguous diffusion from the capital metropolitan area, and the other hierarchical with long-distance spread through major highways that connects São Paulo city with cities of regional relevance. This knowledge can provide real-time responses to support public health strategies, optimizing the use of resources in order to minimize disease impact on population and economy.

Introduction

The International Health Regulations (IHR), administered by World Health Organization (WHO), was last revised in 2005, under the influence of the global response to the SARS emergency and the risk of the H5N1 influenza pandemic [1]. Since then, it has guided coordinated

Funding: We thank Centro de Vigilância Epidemiológica of São Paulo State (CVE), Health Department, to provide the data. CPF and TNV thank the support from São Paulo Research Foundation (FAPESP) grant 18/24058-1, 18/24811-1, respectively.

Competing interests: The authors have declared that no competing interests exist.

international cooperation during public health emergencies such as the Zika virus and Ebola epidemics [2]. However, the current COVID-19 pandemic is the greatest challenge faced by IHR thus far [3]. Although the WHO has issued several guidelines related to the current epidemic, the adherence level varies among nations and, inside nations, provinces, and states [4].

Up to the present day, non-pharmacological interventions, like social distancing, radical lockdown, and extensive testing for SARS-CoV-2 infection, have been applied by different countries, with widely varying degrees of success [5, 6]. In some countries, such as Brazil, scientific research on the effectiveness of those strategies has been severely hampered by political bias, which interferes with public health decisions [7].

São Paulo, the most populous state in Brazil (45 million inhabitants), is also the most severely affected by COVID-19. The state government has challenged the Brazilian President's denial of the pandemic and declared the closure of commerce, schools and other non-essential services. However, despite the ferocious spread of the virus on the state capital and metropolitan area, the slowly evolving of the epidemic in the state's inner cities (until April 18th), where 20 million people live, has led to protests against governmental measures. In this context, there is a sense of urgency about predicting routes of epidemic spreading in the inner state and the population's risks.

Here, we discussed a detailed analysis of the spatial dispersion of COVID-19 in São Paulo State, Brazil, intending to provide real-time responses to support public health strategies. Using data since the first confirmed cases of COVID-19 in São Paulo State, we assess the importance of geographic space on the spread of the epidemic. We hypothesize that urban hierarchy is the main responsible for the disease spreading, and we identify the hotspots and the main routes of virus movement from the metropolis to the inner state. This premise is also supported by [8] where multivariate analyses showed that demographic density and high classification of regional relevance were associated with early introduction and high COVID-19 incidence and mortality rates. We cross validate the confirmed cases with urban mobility, urban hierarchy, and land use at each spatial localization, in work developed here. The results highlight the importance of the main routes that cross São Paulo State and the regional airports on introducing the disease in the territory, just as the main municipalities act as critical centers of disease spreading to the inner state. Knowing in advance the path of COVID-19 dispersion can support decision-makers to optimize health service, and plan strategies of quarantine measures. This approach can be made in other states of Brazil and other developing countries, observing local and regional mobility and urban network [9].

Methods

Geographical data modelling

Spatial analysis of surveillance data includes exploratory data analysis, spatial modeling, and cartographic visualization [10]. The first one uses spatial statistical methods to measure centrality and dispersion of data sets to detect spatial patterns and check for relationships between variables of the complex phenomenon under investigation. The second one examines the elementary forms of spatial organization that explains the phenomenon under study, such as railways, land cover, demographic, economic, and social factors [11]. Lastly, cartographic visualization provides a synthesis of the previous procedures, aiming the elaboration of a thematic map that can be presented to managers for decision making in emergencies in public health [9, 10].

Focusing on São Paulo State, its center and periphery structure, main roads, and network flux that gives population and trade mobility, the geographic spread of COVID-19 was studied. For this, several maps were made to summarize information about quantity and localization of

confirmed cases, urban hierarchy, area of influence and urban-rural typology of cities, modes of transport, and population vulnerability. The maps were constructed based on graphic-semiology principles, the theory of colors and visual communication [10, 12, 13].

We used surveillance data (number of confirmed cases of coronavirus) updated on April 18th, 2020. The data was obtained from Brasil IO's compiled databank (<https://brasil.io/dataset/covid19/boletim/>) kept by volunteers' task force (researchers and reporters). This group daily catches, from the epidemiological reports of each state, the number of confirmed cases and death by SARS-CoV-2 and make it publicly available. Because of the recognition of SARS-CoV-2 as a pandemic by the WHO, laboratory certification in Brazil ranged from few laboratories to 26, in eight weeks since the beginning of the epidemic; the majority is located at São Paulo State [14]. Data reporting of Severe Acute Respiratory Illness (SARI) is mandatory in Brazil. A specific form (national database SIVEP-Gripe) collects information that allows us to estimate reported delay, disease fatality at which age class, and identify confirmed cases of the disease. This permits surveillance of all respiratory diseases in Brazil. Only cases that were hospitalized belong to this data set; therefore, underreporting is expected. On the other hand, this is probably homogeneous along with the municipalities and will not impact on the observed pattern of disease spreading. Out of this national surveillance system (SARI), test capacity can vary among cities because many of them made agreements with factories to test the population that lives on the site where the factory is settled. The same procedure has been done in universities, schools, and firms that returned their activities. However, this data has not been taken into account here, since SARI only reports severe cases. Lastly, test capacity has grown fastly in São Paulo State; currently, 13 out of 35 laboratories of Brazil are settled in this state. A broader serological survey is on course in Brazil, to detect underreporting and follow the population susceptibility along with the course of the epidemic. This may help in defining target groups for vaccination.

Data about each municipality, such as territorial management, trade and services, financial services, health care services, educational institutions, media and communication markets, culture and sport, mode of transport, and land use, was used to identify the fundamental entities of spatial structure that trigger coronavirus dispersion in São Paulo territory [15]. This information was compiled from census data done by the Federal Government and other thematic studies. The metropolis of São Paulo appears as the largest urban complex in the country, with almost 22 million inhabitants and a high level of integration with other municipalities that comprise the national territory. It is listed as an alpha global city by the Globalization and World Cities Research Network (GaWC). In the second level of the hierarchy, we have Rio de Janeiro and Brasília (the capital of Brazil). Focusing on São Paulo State, it has two main axes of urban and trade mobility (roadways, railways, and airways), the first one connecting São Paulo with Rio de Janeiro, and the second one connecting São Paulo with Brasília and Central-West Region of Brazil. Besides these main transportation axes, we have a secondary flux network connecting the metropolis of São Paulo to country municipalities and the South of Brazil. This secondary flux aggregates roadways, railways, airways, and waterways (Tietê-Paraná). Over this intense flux of people and trade, a complex structure of cities emerges, reinforcing this network composed of high hierarchy cities (as nodes) linked by the best transportation system of the country (as edges). With almost 48 million inhabitants, São Paulo State concentrates 23.6% of the country's population and 33% of its income. Besides, São Paulo State has the highest number of primary (Metropolis) 2/15, and secondary (Regional Capital) 26/97 cities on the urban hierarchy level in Brazil, and a high number of other cities classified as Subregional Center 77/352, and Zone Center 51/398, respectively, at tertiary and quaternary levels.

Out of 645 municipalities in São Paulo State, 145 have laboratory-confirmed cases on April 18th and were used in this study. The first confirmed case was at São Paulo metropolis on

February 25th. In all maps that we will present, the studied feature (number of coronavirus cases or time-lapse since the first case) is located at each municipality's city hall.

In the first map, we plotted the number of laboratory-confirmed cases reported from February 25th to April 18th. For this, the proportional symbol maps scale was used to draw circles proportionally to the number of cases in each municipality. Proportional symbol maps are often constructed by beginning with the largest symbol size (the largest radius of the circle corresponds to the largest data value) to minimize symbol overlap. To measure the spatial trend on data, we use a weighted standard deviation ellipse. In this case, data of each municipality i was (until the date at which each ellipse was drawn), centered at the city hall (position coordinate (x_i, y_i)) and weighted by the number of cases in the municipality [16].

Three ellipses were drawn to show, at different times, the main direction of disease spreading. Although the SARS-CoV-2 was introduced in São Paulo on February 25th, it took time to move towards the inner municipalities because of the strong mitigations strategy adopted by São Paulo State to halting the disease's spread. The average time spent by the disease, since its introduction on the metropolis, to achieve the regional centers, the municipalities under major and minor influence, and the rural municipalities were respectively 22, 31, 34 and 55 days [8] (the classification of the municipalities follows the criteria established by the Brazilian Institute for Geography and Statistics (2017) [17]). Therefore, three calendar date were chosen to cover the period of study: March 29, April 8, and April 18; which are 10 days apart from each other. Over this information, we highlight the main roads that cross São Paulo State, and its 645 municipalities' urban-rural typology.

The standard deviations for the x- and y-axis are given by

$$\sigma_{1,2} = \sqrt{\frac{\sum_i \tilde{x}_i^2 + \sum_i \tilde{y}_i^2 \pm \sqrt{\sum_i \tilde{x}_i^2 - \sum_i \tilde{y}_i^2 + 4 \sum_i \tilde{x}_i \tilde{y}_i}}{2n}}^2$$

where

$$\bar{x} = \sum_i x_i, \quad \bar{y} = \sum_i y_i, \quad \tilde{x}_i = x_i - \bar{x}, \quad \tilde{y}_i = y_i - \bar{y},$$

and the summation symbol i takes into account the number of municipalities with registered cases of COVID-19. Observe that (\bar{x}, \bar{y}) represents the mean center of the feature. A Standard deviation ellipse summarizes both the dispersion and orientation of the observed set of samples. If the data is normally distributed, one standard deviation represents approximately 68% of all occurrences.

The second map shows the movement of airplanes during March and April of 2020, connecting municipalities of São Paulo State among them and with other states and countries. The data were obtained from [18]. Line thickness is proportional to the number of passengers moving from one place to another. For the purpose of the study, the flux inside São Paulo territory is highlighted.

The third map shows the urban hierarchy centrality level of municipalities. The regional importance of each city can also be seen in this figure from the tree diagram. The data were obtained from [15]. According to the literature, five urban hierarchy levels are defined: metropolis, regional capital, subregional center, zone centers, and local center. Many variables are used in this classification, such as services establishments, inter-urban relations, banking establishments, social information, cultural and sports offering, and territorial management [15, 19].

The fourth map was constructed by interpolating over the total number of days at which coronavirus transmission was reported in each municipality. We used Inverse Distance Weighting (IDW) as an interpolator and a circle as the neighborhood shape for the interpolation procedure. The Root Squared Error (RMS) permitted to set up the radius of the circle (25 km) and the minimum and maximum numbers of neighbors, respectively 2 and 12, to optimize global accuracy of the interpolated curve. Inside of this radius, the nearest neighbors (with reported cases) of each point s_0 were used in the interpolator and the contribution of each one was weighted by the inverse of its distance. This gave us an RMS of 6.56. Assuming that the measured values closest to the prediction location have more influence on the predicted value than those far away, the following equation was used

$$\hat{z}(s_0) = \sum_i w_i z(s_i),$$

where $\hat{z}(s_0)$, w_i , and $z(s_i)$ are the estimated value at position s_0 , the weight attributed to each pair of coordinates ($1/|s_i - s_0|$) and the numerical value observed at position s_i . In the summation symbol, i takes into account the number of neighbors.

The interpolator created a surface on which the values from points (municipalities) are combined and recorded in a data matrix, simplifying information, and creating regional patterns. As it has spatiotemporal data, it must be read with the darkest data in the red palette as the oldest that passes through orange, yellow going to the blue palette, which are the municipalities that were later infected. Although we are using the time-lapse since the first case reported in each municipality to create the interpolated surface that permits us to predict the epidemic course over the state, some cities entered community transmission only after 10 to 15 days from the first confirmed case. This reflects not only the stochastic nature of the introduction of a new pathogen in a community, but also the fact that data of COVID-19 in Brazil do not distinguish between imported and autochthonous cases. This is an exploratory analysis that permits us to follow the virus's dispersion pattern, glimpsing the next cities that will probably be affected by the disease. The palette of colours of the map comprises the period from February 25th to April 18th, with white color indicating disease absence.

The last map is a schematic cartogram of the elementary spatial structures that drive and modulate disease spreading in São Paulo territorial. It shows the main modes of transportation, together with the key municipalities that acted as agents of the initial spreading of COVID-19. It also highlights the geographic position of the metropolis and the vulnerable population.

Finally, we want to emphasize that we are not looking for epidemiological links that explain disease transmission among municipalities, but we seek geographical links that conditionate the regional pattern of disease spreading along São Paulo territorial.

Results and discussion

[Fig 1](#) shows on grayscale the 645 municipalities painted according to their classification of urban-rural typology that takes into account population density, accessibility to goods and services, and land use; São Paulo State has 13% of its area classified as rural [17, 20]. In pink, we highlighted the main roads that cross the state, dividing them into primary and secondary axes according to the flux of people and goods. The municipalities with reported COVID-19 cases are shown in red circles, which size is proportional to the number of cases recorded until April 18th. Three weighted standard deviation ellipses are shown on March 29th, April 08th, and April 18th. The angles are 128, 135, and 137 degrees, respectively, and the semi-major axis measures 34, 89, and 110 km. As time passes and the epidemic evolves, we can notice a change

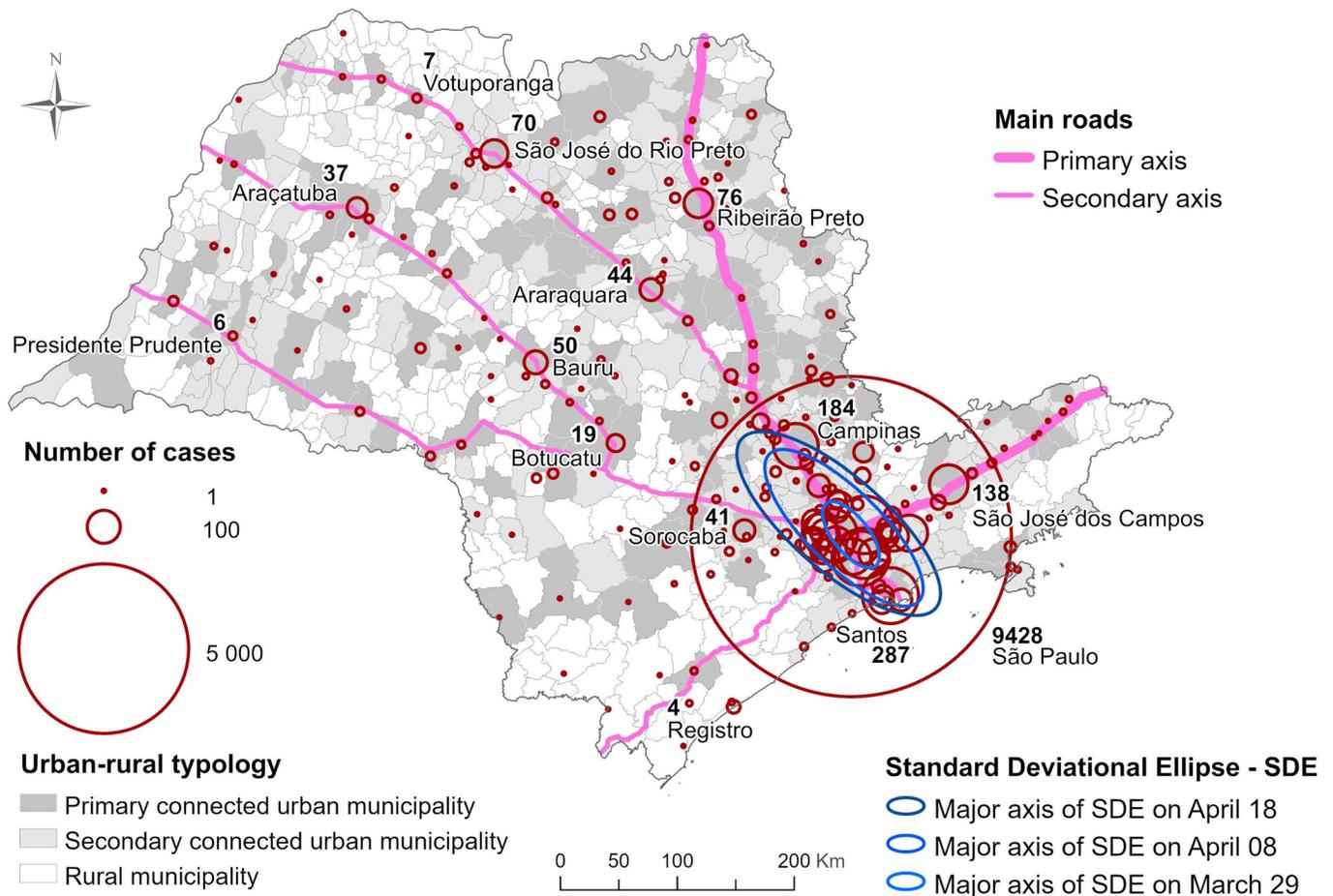


Fig 1. Distribution of confirmed COVID-19 cases in São Paulo State as of April 18th 2020, Brazil. The size of the circles is proportional to the number of cases reported in each municipality. The map also shows the main roads that cross the state, the typology of each municipality, and the direction of disease spread on three different moments of epidemic course. The map was made using the software ArcGIS (version 10.8).

<https://doi.org/10.1371/journal.pone.0245051.g001>

of direction and velocity of disease spreading. A simple calculus gives us 5.5 km/day and 2.1 km/day ($\Delta S/\Delta t$ where ΔS is the difference between the semi-major axis measure and Δt is the time elapsed between each ellipse). Interesting to note that on March 24th, mobility restriction was imposed at São Paulo State. Planned to be finished on April 07th, this restriction mobility was extended several times until May 27th when São Paulo quarantine plan started [21].

To emphasize the mobility restrictions imposed by São Paulo government and how connected are its municipalities, Fig 2 shows the airplanes moving in and out of São Paulo during March and April months of 2020. It was registered a movement of 2 to 11350 individuals per connection (107 different connections among different cities) using this option as transportation in March, and 2 to 4 individuals per connection (16 connections among different cities), considering that every flight has at least two individuals. In the figure, line thickness is proportional to the number of individuals moving among municipalities. The reduced number in April reflects the travel restrictions imposed by the government of São Paulo to reducing coronavirus spread in the state. We can see that the inner state is well connected not only by roadways (as shown in Fig 1) but also through airways. A highlight to cities of Campinas and São Paulo that have international airports. In order to compare, in March and April of 2019, the number of passengers that left São Paulo State was three times greater than in 2020; and

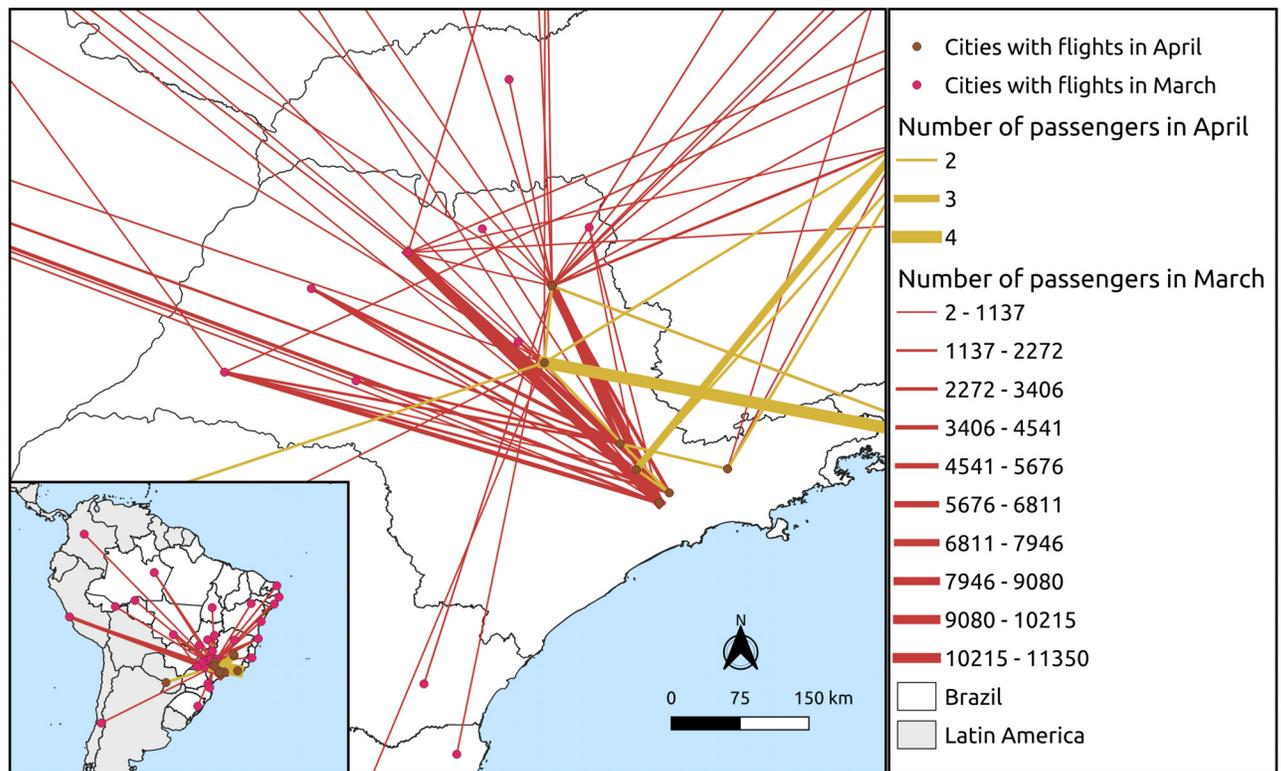


Fig 2. Airway connections in March and April of 2020 at São Paulo State, Brazil. Red (March month) and orange (April month) colors show airport connections among different cities of São Paulo, as well as among São Paulo State and other states in Brazil or other countries. Line thickness is proportional to the number of passengers moving from one city to another. The map was made using the software QGIS 3.10.

<https://doi.org/10.1371/journal.pone.0245051.g002>

the number of passengers arriving at São Paulo State's airport was 15 times in 2019 than in 2020.

People's movement is facilitated, and encouraged, due to transportation availability and commercial and social activities. Likely SARS-CoV in 2003, the SARS-CoV-2 fastly spread among cities and countries due to airline network and ground transportation [22–24]. In the case of São Paulo State, the delay in closing the airports located at the inner municipalities probably contributed to the hierarchical dispersion of the disease on its territorial.

Fig 3 displays the cities with a high level of urban hierarchy centrality that we can find at São Paulo State: the metropolis of São Paulo, the regional capital of Campinas, the subregional centers of São José do Rio Preto and Ribeirão Preto, the zone centers of Presidente Prudente, Marília, Bauru, Araçatuba, Sorocaba, São José dos Campos, Santos, Araraquara, and Piracicaba, and local center of Barretos, Franca, São João da Boa Vista, São Carlos, Rio Claro, Limeira, Ourinhos, Botucatu, Jaú, and Catanduva. The black border delimits the regions subordinate to the cities level one and two in the hierarchy, and the yellow one the regions subordinate to cities of level three. The diagram on tree summarizes the regions of influence of each city displayed on the map. We hypothesize that city hierarchy plays an important role in the disease spreading over the territory.

The exploratory analysis of data on confirmed cases in São Paulo State generated a dispersion map in which the color spectrum indicates the areas ranging from earlier to the more recent introduction of SARS-CoV-2 (Fig 4). The colors have to be reading such as a predictor of an earlier or later arrival of the disease in each city of the map because they comprise only

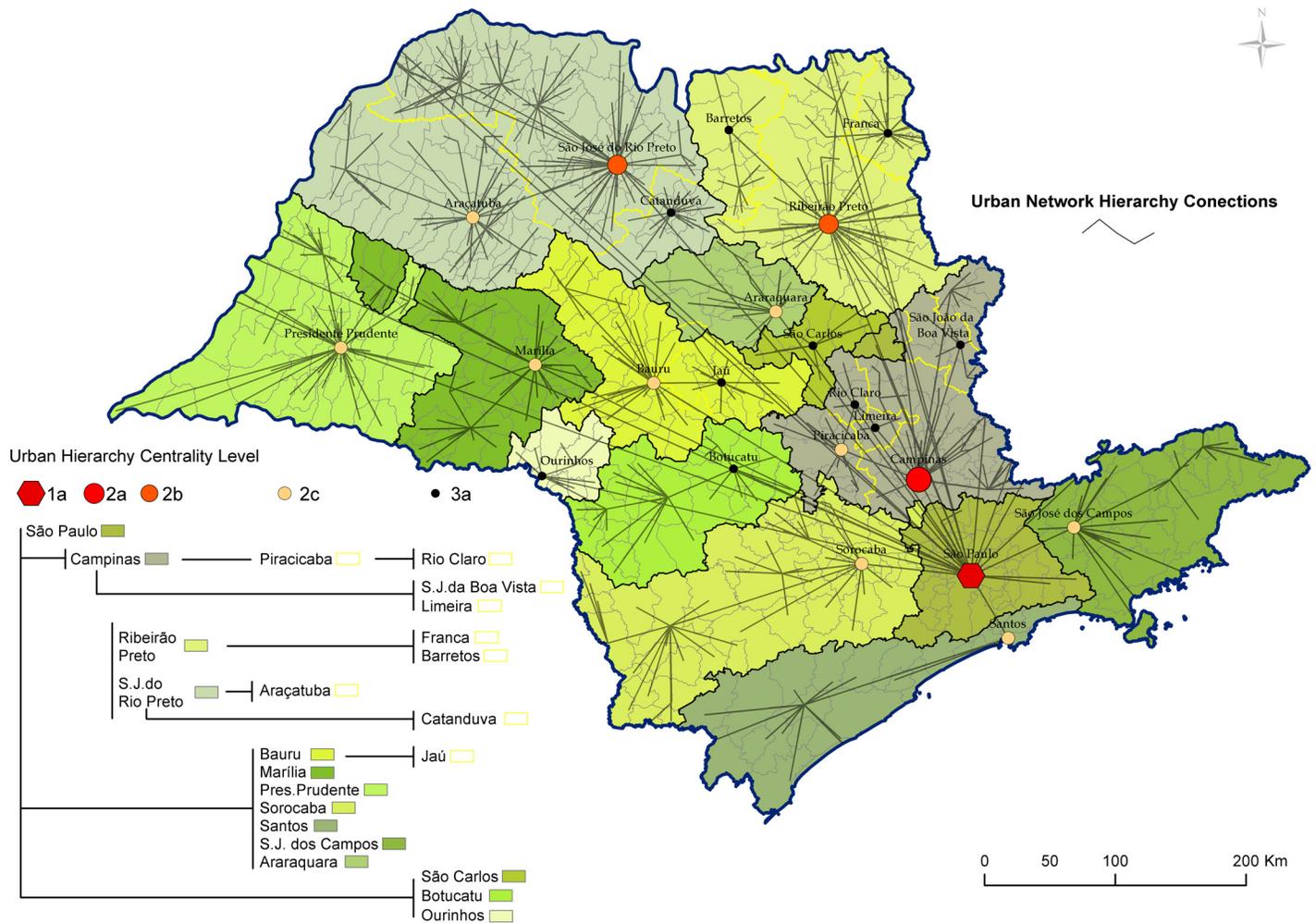


Fig 3. Level of urban hierarchy find at São Paulo State: Metropolis, regional capital, subregional center, zone centers, and local center. The tree diagram schematizes the hierarchy of the main cities, and the lines that divided the territory (the yellow and black borders) show the regions of influence of cities classified as level 1, 2, and 3 of urban hierarchy. The map was made using the software ArcGIS (version 10.8).

<https://doi.org/10.1371/journal.pone.0245051.g003>

the study period; therefore, it is a first insight into disease dispersion. The white areas on the map show the municipalities without COVID-19 cases, and at the same time, far away from the ones where the disease was already reported. Overlapping this map with the one shown in Fig 1, we can see that most municipalities without case reported are classified as rural ones. The main roads that cross the state are also highlighted, and we hypothesize is that they also play an important role in the disease spreading over the territory.

To understand the regional pattern of SARS-CoV-2 spreading Fig 5, presents the elementary spatial structures identified as the main ones responsible for the disease spread inside the state. They comprise the main roadways as well as the airports that give people and trade mobility, and the hotspots of the disease introduction and spread. The airports and the cities are displayed by circles proportional to their role on COVID-19 spread. In the case of cities, we classified them as principal (São Paulo, Campinas, São José dos Campos, Ribeirão Preto, and São José do Rio Preto) and secondary (Santos, Araçatuba, Presidente Prudente, Bauru, Marília, São Carlos, Sorocaba, Rio Claro, and Piracicaba) urban centers in the level of relevance for the disease spread. Cities belonging to the metropolitan area, such as Santos, São

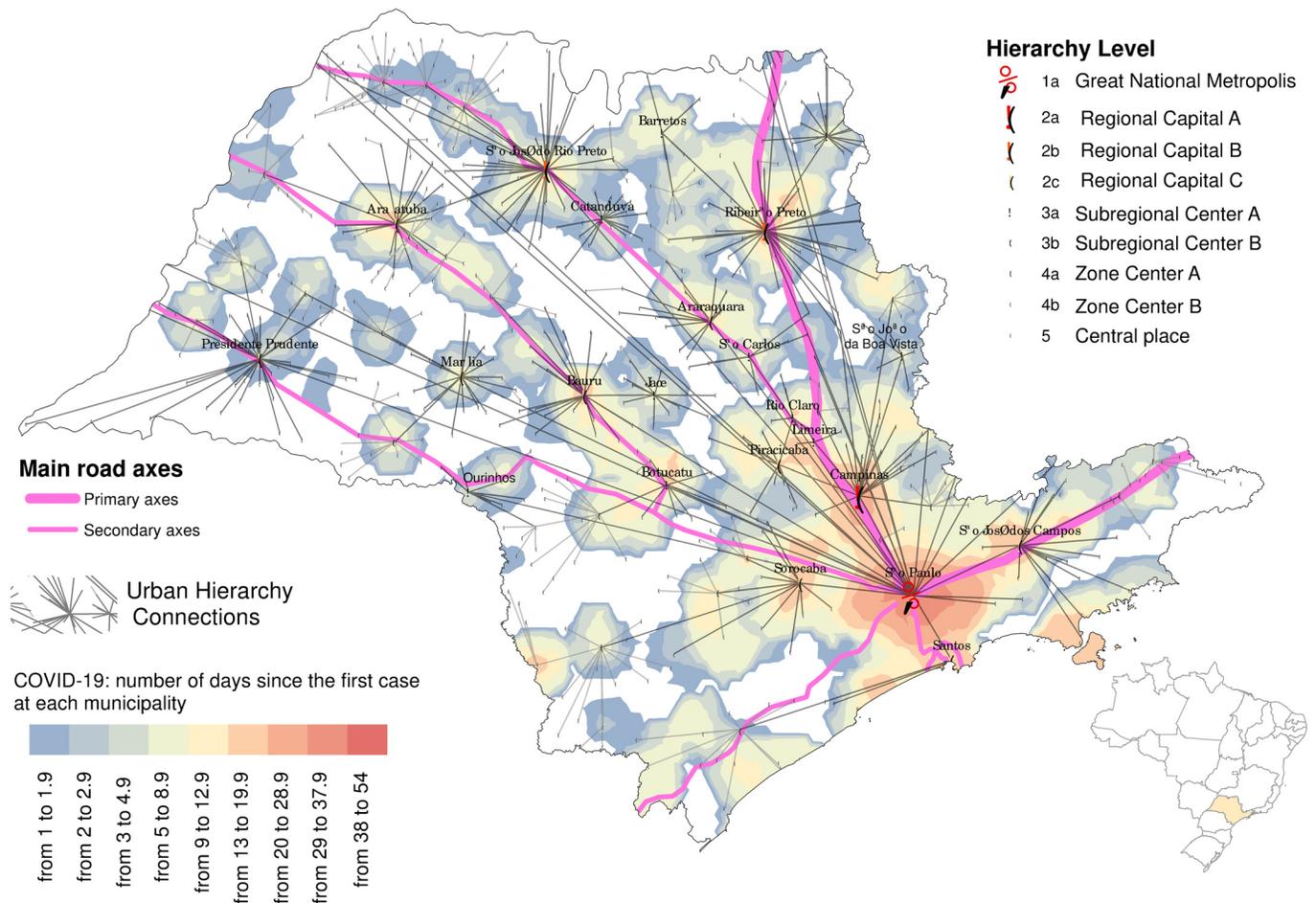


Fig 4. Dispersion map for COVID-19 in São Paulo State Brazil from March 25 to April 18, 2020. The color spectrum indicates early introduction areas (in red) to those of the more recent COVID-19 introduction (in blue). The main roadways that cross São Paulo State and the urban hierarchy level of each municipality is displayed. The map was made using the software ArcGIS (version 10.8).

<https://doi.org/10.1371/journal.pone.0245051.g004>

José dos Campos, and Campinas, are classified as “contiguity”; the other ones are connected to São Paulo City through a primary or secondary axis. The metropolitan area of São Paulo and the state’s region where there is a massive concentration of elderly population (older than 60 years of age) are highlighted. The latter is called vulnerable because disease lethality among them is high. For these listed cities, demographic characteristics, number of reported cases and, disease lethality (up to April 18, 2020) are presented in [Table 1](#). Santos, which has a considerable mortality per 100,000 inhabitants, is the one in the list with the more significant number of the older population (≥ 50 years).

Based on the results of the exploratory analysis (Figs 1 and 4) and population mobility studies (Figs 2 and 3), two dispersion patterns were postulated. In the first one, virus dispersion occurs by contiguity, from a region of initial introduction, that is the Metropolitan Region of the Capital, the City of São Paulo (contagious diffusion) to its nearest neighborhoods. In the second one, there is a long-distance dispersion following structural axes (roadways and airways) that connect São Paulo city to peripheral municipalities of regional importance (hierarchical diffusion). From these, diffusion by contiguity occurs again to smaller municipalities.

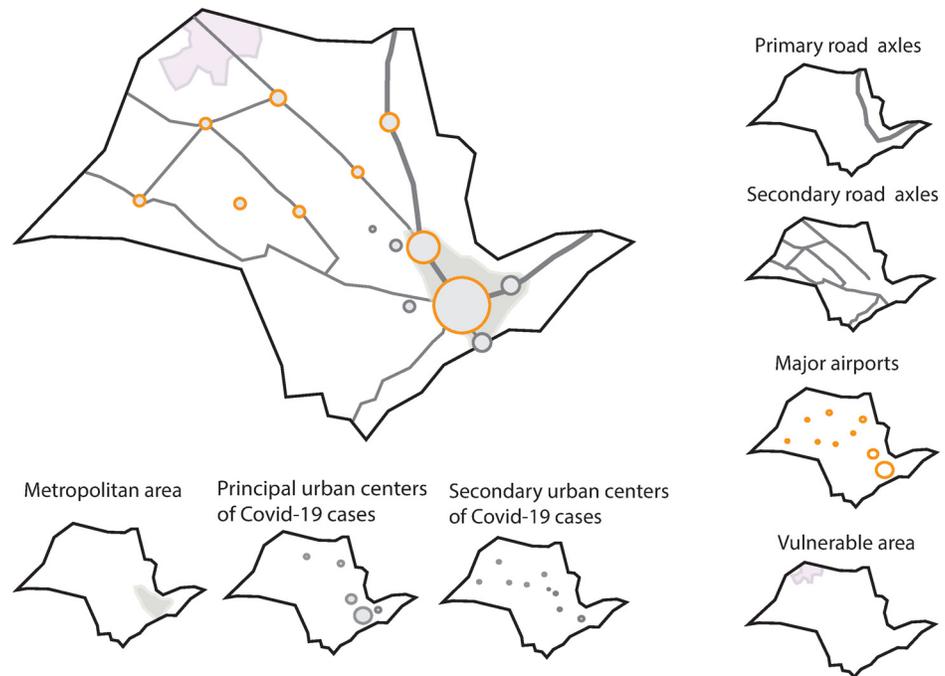


Fig 5. Elementary spatial structures associated with COVID-19 spread in São Paulo State, Brazil. Around the main map, we display the structures that comprise it, such as road axes, regional airports, the metropolitan area of São Paulo city, municipalities keys as centers of disease dispersion from the metropolitan area to inner state, and municipalities where the oldest population of São Paulo lives.

<https://doi.org/10.1371/journal.pone.0245051.g005>

Table 1. Epidemiologic COVID-19 data for São Paulo State capital and hotspots cities for disease introduction and spread on April 18th (see Fig 5).

Municipality	Population (inhabitant)	Dist. ¹ (Km)	Connection with the capital ²	Cumul. cases	Incid. ³	Cumul. deaths	Mort. ³	Date of arrival ⁴
São Paulo (capital)	12252023	-	-	9428	76.95	686	5.6	2020-03-25
Campinas	1204073	95	Contiguity	184	15.28	7	0.58	2020-03-18
Ribeirão Preto	703293	314	Primary axis	76	10.81	5	0.71	2020-03-26
São José do Rio Preto	408558	440	Secondary axis	70	17.13	4	0.98	2020-03-18
São J. dos Campos	721944	91	Contiguity	138	19.12	3	0.42	2020-03-18
Santos	433311	55	Contiguity	287	66.23	19	4.38	2020-03-30
Sorocaba	679378	100	Secondary axis	41	6.03	8	1.18	2020-03-27
Piracicaba	404142	162	Secondary axis	19	4.70	2	0.49	2020-03-30
Bauru	376818	343	Secondary axis	50	13.27	3	0.80	2020-04-03
Presidente Prudente	228743	550	Secondary axis	6	2.62	2	0.87	2020-04-08
Araçatuba	197016	530	Secondary axis	37	18.78	0	0.00	2020-03-31
Marília	216745	438	Secondary axis	8	3.69	1	0.46	2020-04-03
São Carlos	251983	231	Primary axis	7	2.78	2	0.79	2020-04-06
Rio Claro	186273	176	Secondary axis	14	7.52	3	1.61	2020-04-03

¹. distance from the capital;

². classification according to Fig 5;

³. incidence or mortality per 100,000 inhabitants.

⁴. Date of the disease arrival at each municipality.

<https://doi.org/10.1371/journal.pone.0245051.t001>

A relationship between disease spreading and territorial geography was also established in other epidemics [22, 25]. Differently from São Paulo State [26], showed that the first wave of SARS-CoV-2 pandemic in Germany followed a dispersion pattern called relocation diffusion process since the arrival of infections in Germany coincided with a traditional carnival festivity. Therefore, a single infected individual transmitted the infection to several others. After the festivities, people went back to their homeland, creating long-range connections, and new spots of infection spread, which were randomly distributed across the country. In São Paulo State, since all non-essential activity was limited, the spread followed the main routes of commercial relationships and supply distribution, in a hierarchical diffusion, firstly reaching the most important cities in São Paulo State, and locally spreading within their regions.

Currently, SARS-CoV-2 is reported in all São Paulo territorial; the last city to be achieved by this coronavirus (September 1th) was Santa Mercedes, a rural municipality with 2,945 inhabitants and 580 km far away São Paulo City (also out of the main roads of coronavirus dispersion). On September 27th, São Paulo State reports 972,237 confirmed cases and 35,108 deaths. The isolation index is 48%, the Intensive Care Unit (ICU) occupation is 45,6%, and disease lethality is 3,6% [27]. Schools and universities still closed, and the state has its own plan of quarantine measures (“Plano São Paulo”) that, based on the growth rate of COVID-19 cases and deaths and bed occupancy rates in each Regional Health Departments (DRS), can be more or less flexible. Cities belonging to the same DRS (we have seventeen) are ruled by the same quarantine measures. Phase 1 is considered a contamination phase, and only essential services are permitted. Phase 2 is considered an attention phase with the possibility of some services such as commerce opening. Food courts are still banned in this phase. Phase 3 is considered a controlled phase with some flexibilization. Phase 4 has less restriction than phase 3, and at phase 5, all services are allowed to open, maintaining all specific protocols. Now, in December 2020, all São Paulo State is at phase 3 [28].

Our predictions of routes and risks of COVID-19 in inner São Paulo State (Fig 5) have been thus far validated by surveillance data. Given the extensive mobility between smaller municipalities and those cities with regional economic relevance [17], it is reasonable to infer that the regional spread of SARS-CoV-2 infections depends on the success of non-pharmacological strategies applied in the latter. We also state that similar methodological approaches can direct public health strategies in other developing countries, especially those that either have great territorial extension and/or have diverse patterns of urbanization and mobility.

Limitations of the analysis include: (i) the no-identification of asymptomatic individuals and, potentially, mild or moderate infectious, since only symptomatic cases that seek for medical assistance have been tested; (ii) data dependence, *i.e.* data set does not distinguish between imported and autochthonous cases; (iii) the assumption that all individuals have the same degree of susceptibility and transmissibility of the disease, regardless the environment they live; (iv) the transmission is homogeneous within the cities; (v) mitigation strategies are the same everywhere. All those characteristics may vary according to the city because the number of tests that is distributed and performed among cities is not homogeneous; the number of contacts among people changes according to the city characteristics, such as the use of public transportation [29]; and people’s adherence to social distancing really differed across the state, which may be related to the epidemics delay into reach the small inner cities, affecting people’s risk perception [30].

Moreover, the data source in Brazil has been updated with some delay, regarding the occurrence of the infections [31]. Nevertheless, since the data used in this study is related to the arrival of infections in each city, which happened in early 2020, we expect the numbers to be trustful at the point of the analysis. Despite there is no data at a granular level, such as

information about the address of the infection occurrence, the data is enough to perform the analysis and reach our goal, which was to study the spread of SARS-CoV-2 among cities.

Conclusion

Spatial analysis of coronavirus spread is an important tool for public health management, as it can highlighting the main routes of disease dispersion and the fragility of municipalities related to its socio-demographic characteristics. In the case of São Paulo State, this analysis evidenced the hotspots and main routes of disease dispersion from capital to inner state. Currently, non-pharmacological controls are the only tools to halt or diminish the disease spreading among both individuals and municipalities. The existence of two different ways of disease dispersal, by standard diffusion and hierarchical one, can provide alternative strategies to control disease spread in the São Paulo territory.

This work shows that it was possible to understand and even predict the route of COVID-19 spread in São Paulo State looking to the cities' hierarchy, which means that the spread of the epidemic does not follow a diffusion process but reaches the cities based on their regional importance and activities. After that, the epidemic spread to contiguous cities following a diffusion standard process. We state that those cities are responsible for the arrival of the epidemics in the inner São Paulo State and demand attention.

Supporting information

S1 Data set.
(XLSX)

Author Contributions

Conceptualization: Carlos Magno Castelo Branco Fortaleza, Raul Borges Guimarães, Cláudia Pio Ferreira.

Data curation: Carlos Magno Castelo Branco Fortaleza, Raul Borges Guimarães, Rafael de Castro Catão, Cláudia Pio Ferreira, Gabriel Berg de Almeida, Thomas Nogueira Vilches, Edmur Pugliesi.

Formal analysis: Carlos Magno Castelo Branco Fortaleza, Raul Borges Guimarães, Rafael de Castro Catão, Cláudia Pio Ferreira, Gabriel Berg de Almeida, Thomas Nogueira Vilches, Edmur Pugliesi.

Funding acquisition: Cláudia Pio Ferreira.

Investigation: Carlos Magno Castelo Branco Fortaleza, Raul Borges Guimarães, Rafael de Castro Catão, Cláudia Pio Ferreira, Edmur Pugliesi.

Methodology: Carlos Magno Castelo Branco Fortaleza, Raul Borges Guimarães, Rafael de Castro Catão, Cláudia Pio Ferreira, Gabriel Berg de Almeida, Edmur Pugliesi.

Project administration: Carlos Magno Castelo Branco Fortaleza, Raul Borges Guimarães.

Software: Edmur Pugliesi.

Supervision: Carlos Magno Castelo Branco Fortaleza, Raul Borges Guimarães, Cláudia Pio Ferreira.

Validation: Carlos Magno Castelo Branco Fortaleza, Raul Borges Guimarães, Rafael de Castro Catão, Cláudia Pio Ferreira, Gabriel Berg de Almeida, Edmur Pugliesi.

Visualization: Carlos Magno Castelo Branco Fortaleza, Raul Borges Guimarães, Rafael de Castro Catão, Cláudia Pio Ferreira, Gabriel Berg de Almeida, Edmur Pugliesi.

Writing – original draft: Carlos Magno Castelo Branco Fortaleza, Raul Borges Guimarães, Cláudia Pio Ferreira, Gabriel Berg de Almeida.

Writing – review & editing: Carlos Magno Castelo Branco Fortaleza, Raul Borges Guimarães, Rafael de Castro Catão, Cláudia Pio Ferreira, Gabriel Berg de Almeida, Thomas Nogueira Vilches, Edmur Pugliesi.

References

1. World Health Organization(WHO). International Health Regulations; 2005. Available from: <https://www.who.int/ihr/publications/9789241580496/en/>.
2. Bennett B, Carney T. Public health emergencies of international concern: global, regional, and local responses to risk. *Medical Law Review*. 2017; 25(2):223–239. <https://doi.org/10.1093/medlaw/fwx004>
3. Jee Y. WHO International Health Regulations Emergency Committee for the COVID-19 outbreak. *Epidemiology and health*. 2020; 42. <https://doi.org/10.4178/epih.e2020013> PMID: 32192278
4. Kandel N, Chungong S, Omaar A, Xing J. Health security capacities in the context of COVID-19 outbreak: an analysis of International Health Regulations annual report data from 182 countries. *The Lancet*. 2020. [https://doi.org/10.1016/S0140-6736\(20\)30553-5](https://doi.org/10.1016/S0140-6736(20)30553-5) PMID: 32199075
5. Kraemer MUG, Yang CH, Gutierrez B, Wu CH, Klein B, Pigott DM, et al. The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science*. 2020; 368(6490):493–497. <https://doi.org/10.1126/science.abb4218> PMID: 32213647
6. Liu W, Yue XG, Tchounwou PB. Response to the COVID-19 Epidemic: The Chinese Experience and Implications for Other Countries; 2020.
7. COVID-19: learning from experience. *Lancet (London, England)*. 2020; 395(10229):1011. [https://doi.org/10.1016/S0140-6736\(20\)30686-3](https://doi.org/10.1016/S0140-6736(20)30686-3)
8. Fortaleza CMCB, Guimarães RB, de Almeida GB, Pronunciante M, Ferreira CP. Taking the inner route: spatial and demographic factors affecting vulnerability to COVID-19 among 604 cities from inner São Paulo State, Brazil. *Epidemiology & Infection*. 2020; 148. <https://doi.org/10.1017/S095026882000134X> PMID: 32594926
9. Huang R, Liu M, Ding Y. Spatial-temporal distribution of COVID-19 in China and its prediction: A data-driven modeling analysis. *The Journal of Infection in Developing Countries*. 2020; 14(03):246–253. <https://doi.org/10.3855/jidc.12585> PMID: 32235084
10. MacEachren AM. *How maps work: representation, visualization, and design*. Guilford Press; 2004.
11. Brunet R. Building models for spatial analysis. *L'Espace géographique*. 1993; 1(1):109–123.
12. Slocum TA, McMaster RB, Kessler FC, Howard HH. *Thematic cartography and geographic visualization*. 2003.
13. Messina JP, Brady OJ, Golding N, Kraemer MUG, Wint GRW, Ray SE, et al. The current and future global distribution and population at risk of dengue. *Nature Microbiology*. 2019; 4(9):1508–1515. <https://doi.org/10.1038/s41564-019-0476-8> PMID: 31182801
14. Grotto RMT, Lima RS, de Almeida GB, Ferreira CP, Guimarães RB, Pronunciante M, et al. Increasing molecular diagnostic capacity and COVID-19 incidence in Brazil. *Epidemiology & Infection*. 2020; 148. <https://doi.org/10.1017/S0950268820001818> PMID: 32807244
15. Regions of City's Influence;. Available from: <https://biblioteca.ibge.gov.br/visualizacao/livros/liv101728.pdf>.
16. Wang B, Shi W, Miao Z. Confidence analysis of standard deviational ellipse and its extension into higher dimensional Euclidean space. *PloS one*. 2015; 10(3):e0118537. <https://doi.org/10.1371/journal.pone.0118537> PMID: 25769048
17. Instituto Brasileiro de Geografia e Estatística. *Classificação e caracterização dos espaços rurais e urbanos do Brasil: uma primeira aproximação*; 2017. Available from: <https://biblioteca.ibge.gov.br/visualizacao/livros/liv100643.pdf>.
18. National Agency of Civil Aviation;. Available from: <https://www.anac.gov.br/assuntos/dados-e-estatisticas/dados-%0Aestatisticos/dados-estatisticos>.
19. Hutton TA. Service industries, globalization, and urban restructuring within the Asia-Pacific: new development trajectories and planning responses. *Progress in planning*. 2003; 61(1):1–74. [https://doi.org/10.1016/S0305-9006\(03\)00013-8](https://doi.org/10.1016/S0305-9006(03)00013-8) PMID: 32287934

20. Guide to applying the rural urban classification to data;. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/539241/Guide_to_applying_the_rural_urban_classification_to_data.pdf.
21. Peixoto PS, Marcondes D, Peixoto C, Oliva SM. Modeling future spread of infections via mobile geolocation data and population dynamics. An application to COVID-19 in Brazil. *PloS One*. 2020; 15(7): e0235732. <https://doi.org/10.1371/journal.pone.0235732> PMID: 32673323
22. Litaker JR, Chou JY, Novak S, Wilson JP. Implications of SARS: Medical geography and surveillance in disease detection. *Annals of Pharmacotherapy*. 2003; 37(12):1841–1848. <https://doi.org/10.1345/aph.1D244> PMID: 14632586
23. Bowen JT Jr, Laroe C. Airline networks and the international diffusion of severe acute respiratory syndrome (SARS). *Geographical Journal*. 2006; 172(2):130–144. <https://doi.org/10.1111/j.1475-4959.2006.00196.x>
24. Ali SH, Keil R. *Networked disease: Emerging infections in the global city*. vol. 44. John Wiley & Sons; 2011.
25. Kuo CL, Fukui H. Geographical structures and the cholera epidemic in modern Japan: Fukushima prefecture in 1882 and 1895. *International Journal of Health Geographics*. 2007; 6(1):25. <https://doi.org/10.1186/1476-072X-6-25> PMID: 17603906
26. Kuebart A, Stabler M. Infectious diseases as socio-spatial processes: The Covid-19 outbreak in Germany. *Tijdschrift voor economische en sociale geografie*. 2020; 111(3):482–496. <https://doi.org/10.1111/tesg.12429> PMID: 32836489
27. Center of Epidemiological Surveillance—SP/BR;. Available from: <http://200.144.0.84/cve-centro-de-vigilancia-epidemiologica-prof.-alexandre-vranjac/areas-de-vigilancia/doencas-de-transmissao-respiratoria/coronavirus-covid-19/situacao-epidemiologica>.
28. de Almeida GB, Pronunciate M, Grotto RMT, Pugliesi EA, Guimarães RB, Vilches TN, et al. Two hundred days of COVID-19 in São Paulo State, Brazil. *Epidemiology & Infection*. 2020; p. 1–9. <https://doi.org/10.1017/S0950268820002927> PMID: 33261679
29. Chen Y, Wang Y, Wang H, Hu Z, Hua L. Controlling urban traffic-one of the useful methods to ensure safety in Wuhan based on COVID-19 outbreak. *Safety Science*. 2020; 131:104938. <https://doi.org/10.1016/j.ssci.2020.104938> PMID: 32834520
30. Van Bavel JJ, Baicker K, Boggio PS, Capraro V, Cichocka A, Cikara M, et al. Using social and behavioural science to support COVID-19 pandemic response. *Nature Human Behaviour*. 2020; p. 1–12.
31. Do Prado MF, de Paula Antunes BB, Bastos LdSL, Peres IT, da Silva AdAB, Dantas LF, et al. Analysis of COVID-19 under-reporting in Brazil. *Revista Brasileira de Terapia Intensiva*. 2020; 32(2):224. <https://doi.org/10.5935/0103-507X.20200030>

Increasing molecular diagnostic capacity and COVID-19 incidence in Brazil

cambridge.org/hyg

From the Field

Cite this article: Grotto RMT, Santos Lima R, de Almeida GB, Ferreira CP, Guimarães RB, Pronunciate M, Azevedo E, Catão RdeC, Fortaleza CMCB (2020). Increasing molecular diagnostic capacity and COVID-19 incidence in Brazil. *Epidemiology and Infection* **148**, e178, 1–3. <https://doi.org/10.1017/S0950268820001818>

Received: 27 June 2020

Revised: 5 August 2020

Accepted: 12 August 2020

Key words:

COVID-19; diagnosis; epidemiology; infectious disease; laboratory tests

Author for correspondence:

Rejane Maria Tommasini Grotto, E-mail: rejane.grotto@unesp.br

Rejane Maria Tommasini Grotto^{1,2,3} , Rodrigo Santos Lima^{2,3}, Gabriel Berg de Almeida³ , Claudia Pio Ferreira⁴, Raul Borges Guimarães⁵, Micheli Pronunciate², Edmur Azevedo⁵, Rafael de Castro Catão⁶ and Carlos Magno Castelo Branco Fortaleza^{2,3}

¹School of Agriculture, São Paulo State University (Unesp), Botucatu, Brazil; ²Botucatu Medical School, São Paulo State University (Unesp), Botucatu, Brazil; ³Clinical Hospital of Botucatu Medical School (HCFMB), Botucatu, Brazil; ⁴Institute of Biosciences, São Paulo State University (Unesp), Botucatu, Brazil; ⁵School of Technology and Sciences, São Paulo State University (Unesp), Presidente Prudente, Brazil and ⁶Federal University of Espírito Santo (UFES), Vitória, Brazil

Abstract

Different countries have adopted strategies for the early detection of SARS-CoV-2 since the declaration of community transmission by the World Health Organization (WHO) and timely diagnosis has been considered one of the major obstacles for surveillance and health-care. Here, we report the increase of the number of laboratories to COVID-19 diagnosis in Brazil. Our results demonstrate an increase and decentralisation of certified laboratories, which does not match the much higher increase in the number of COVID-19 cases. Also, it becomes clear that laboratories are irregularly distributed over the country, with a concentration in the most developed state, São Paulo.

Different countries have adopted strategies for the early detection of SARS-CoV-2 since the declaration of community transmission of the virus by the World Health Organization (WHO), allowing early clinical intervention and the management of these patients with non-pharmacological measures such as hospital isolation, suspension of regular activities and respiratory support in intensive care units (ICUs) promptly after diagnosis [1]. Timely diagnosis has been considered one of the major obstacles for surveillance and healthcare (especially hospital) preparedness in low-to-middle income countries [2]. With that in mind, we studied the increase in COVID-19 molecular diagnostic capacity of public health laboratories in different regions in Brazil. We were especially interested in analysing the association of newly certified laboratories with the increase of COVID-19.

Therefore, we searched the epidemiological bulletins provided by the Ministry of Health of Brazil (available at <https://covid.saude.gov.br/>) for the weekly incidence of laboratory-confirmed cases; and the Union Official Diary (a daily publication of the Federal Government official decrees, available at <https://www.jusbrasil.com.br/diarios/DOU/>) to identify new certifications of public laboratories for the molecular diagnosis of SARS-CoV-2. Our analysis was carried out until 4th June, when the introduction of SARS-CoV-2 in the country completed 100 days. Both the number of newly certified laboratories and the weekly incidence of laboratory-confirmed COVID-19 were submitted to Joinpoint Regression, using software Joinpoint 4.8 (National Cancer Institute, Calverton, MD) [3]. This analysis detects changes in rate trends, and was performed using a log link function to fit the data. We also performed univariate and single-step multivariable Poisson Regression model, with the number of laboratories and the epidemiological weeks as predictors for the outcome of interest (rate of COVID-19 confirmed cases), using STATA 14 (Statacorp, College Station, TX), and georeferenced the time of introduction of COVID-19 and certification of laboratories in different areas in Brazil, using ArcGIS 10 (ESRI, Redlands, CA). We then applied the inverse distance weighted (IDW) technique to interpolate discrete cases and transform it in a continuous surface in raster format, highlighting the date of case arrival and the geographic region. IDW is a local deterministic interpolator which does not exceed data intervals between neighbours. We used 20 neighbours and a 1.5 power factor. The diffusion layer was overlaid with the laboratories mapped by municipality and week of certification.

Our results are shown in [Figure 1](#). We can observe on the map (panel C) the COVID-19 spatial diffusion pattern, starting in the main metropolis throughout the country, mainly São Paulo, Rio de Janeiro, Fortaleza, Recife and Manaus. The diffusion follows the path of the urban network going from major metropolis to middle size cities and then to small cities. The velocity of the spread is different among the regions of the country. In North of Brazil, the diffusion was extremely fast despite lack of road transport in several portions. In

© The Author(s), 2020. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

CAMBRIDGE
UNIVERSITY PRESS

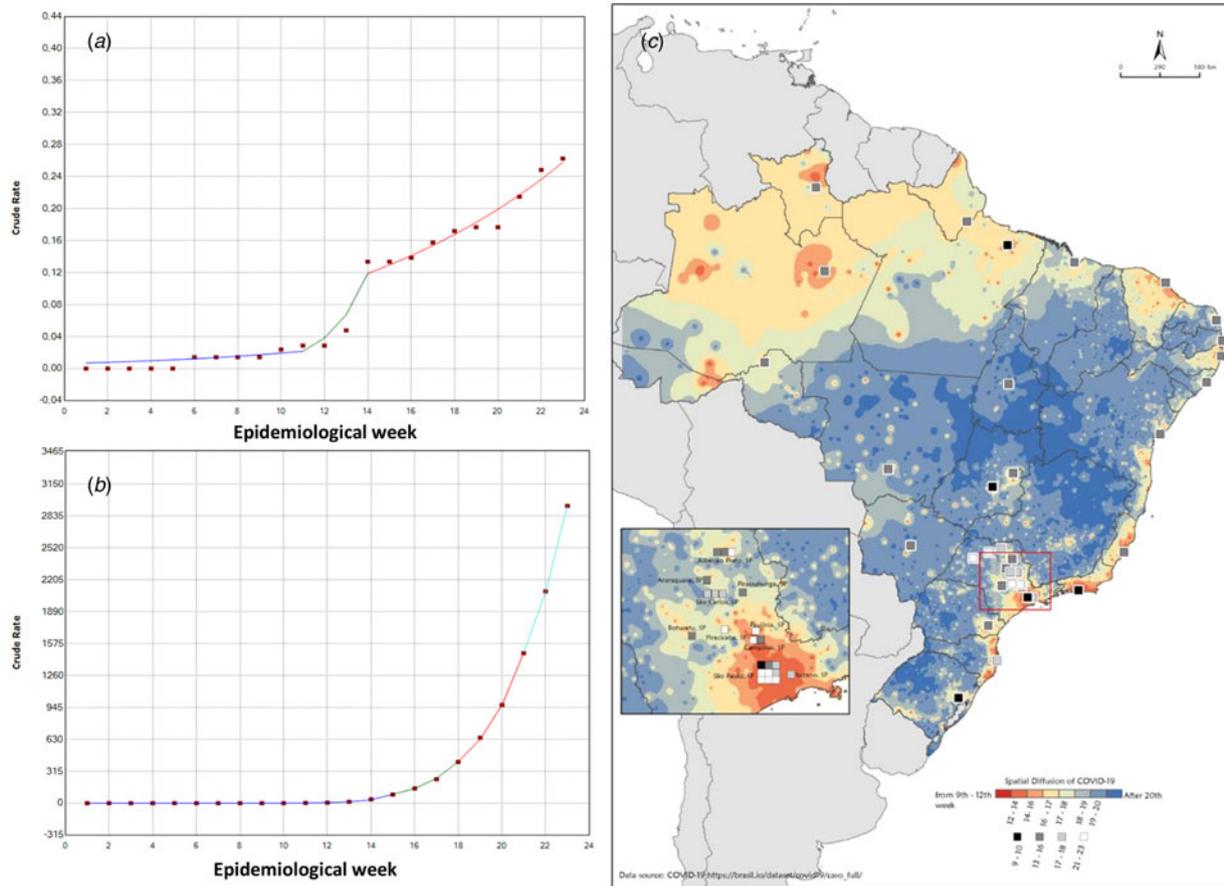


Fig. 1. Trends on SARS-CoV-2 infection and laboratory capacity in Brazil over one hundred days since first COVID-19 confirmed case. Section A shows the rate of certifications of public laboratories for the molecular diagnosis of SARS-CoV-2 by epidemiological week, in a Joinpoint Regression analyzes. Section B shows the rate of incidence of laboratory-confirmed cases by epidemiological week, also in a Joinpoint Regression analyzes. Section C shows temporal-spatial diffusion of COVID-19 in Brazil: warm colors designate early introduction (i.e. 12–14 epidemiological weeks), while cool colors designate recent introduction (i.e. 19–20 epidemiological week). Certified laboratories for SARS-CoV-2 infection molecular diagnosis are represented by squares in greyscale also in section C. The shades of grey range from the darkest for the three initial certified public health laboratories to the light grey representing those that were certified in later stages of the outbreak.

Northeast and Southeast regions, metropolitan areas and cities with higher populational density concentrated the early cases of COVID-19, spreading after towards the interior of each region. In South and Midwest regions, the diffusion was at the initial stages, centred in major urban areas and near major road axis.

Panel A shows the rate of certifications of public laboratories for the molecular diagnosis of SARS-CoV-2 by epidemiological week, whereas panel B shows the rate of incidence of laboratory-confirmed COVID-19 by epidemiological week, both in a Joinpoint Regression analysis (per 100 000 inhabitants). One can note an increase of certified laboratories, which does not match the much higher increase in the number of cases.

The laboratory's certification ranged from few laboratories (one in the North-Northeast region, one in the Middle East and South and two in Southeast) to 26 labs, in 8 weeks. Almost all regional centres of the country certificated at least one laboratory. In summary, at week 23 the ratio between the number of laboratories to population density (population per km² at the last census at 2010) at each Brazil's region was 6:4.14, 4:8.75, 28:86.92, 3:48.58 and 7:34.15, respectively, at North, Middle East, Southeast, South and Northeast. It is also clear from panel C that laboratories are irregularly distributed over the country, with a concentration in

the most developed state, São Paulo (13 of 35 laboratories). However, a decentralisation trend can be seen over the last epidemiological weeks, highlighted by the presence of newly certified laboratories specially in North and Northern regions of Brazil. It is important to reveal that in São Paulo State there are many cities performing mass testing, and this could explain the greater growth in the number of labs. Interestingly, the number of laboratories is positively associated with the number of COVID-19 cases in the univariate model (incidence rate ratio (IRR), 1.11; 95% confidence interval (CI), 1.11–1.11) but negatively associated after adjusting for epidemiological week (number of laboratories: IRR, 0.98; 95% CI 0.97–0.99; epidemiological week: IRR, 1.70; 95% CI 1.69–1.71). This finding can be interpreted in two directions. From an optimistic perspective, the slow increase in certifications of new laboratories for diagnosis does not necessarily correlates with the overall diagnostic capacity of the laboratory net, once already certified individual laboratories may increase their own capacity as well. This analysis could not be performed as no data were available regarding diagnostic capacity of each one of these laboratories. From a pessimistic perspective, the fast increase of COVID-19 incidence and the continuous spread into inner country, less developed areas of Brazil challenges diagnostic

capacity and therefore, accurate and timely health surveillance [4]. This implies that, given the fast increase of COVID-19 cases and the continuous spread into inner Brazil [5], the laboratories (which are continuously increasing their capacity) may still be insufficient to provide accurate data in a setting of exhaustion of hospital (especially ICUs) capacity [6]. The primacy of the growth of cases over laboratory capacity is reinforced by the increase in hospital admissions and deaths (as reported in official data (<https://covid.saude.gov.br/>) and recent studies [7, 8]).

Health surveillance has been a strong pillar of response to previous public health emergencies in Brazil, including pandemic H1N1 influenza and Zika virus [9]. Challenges for COVID-19 response are not restricted to health surveillance [10], but strengthening an accurate knowledge of its behaviour can direct preventive strategies (including infection control). Serial antibody prevalence surveys may be an option but are still hampered by inaccurate serological tests [11]. Given that quarantine measures started to be relaxed in national territory, this is a critical moment where diagnosis missing can jeopardise the epidemic control. Therefore, Brazilian states must increase their capacity of timely molecular diagnosis, not only to face this pandemic, but as a network for preparedness for future public health emergencies [12].

Author contributions.

Grotto and Fortaleza: concept and design the study and drafting of the manuscript; Lima, De Almeida and Pronunciate: acquisition of data; Ferreira, Guimarães, Catao and Pugliesi: analysis and interpretation of data; all authors: critical revision of the manuscript for important intellectual content.

Conflict of interest. None.

Data availability statement. The authors state that the database used in the analyses can be available as a supplementary file to the paper or provided to interested researchers upon reasonable request.

References

1. **Davies NG et al.** (2020) Effects of non-pharmaceutical interventions on COVID-19 cases, deaths, and demand for hospital services in the UK: a modelling study. *Lancet Public Health* **5**, e375–e385. doi: [10.1016/S2468-2667\(20\)30133-X](https://doi.org/10.1016/S2468-2667(20)30133-X).
2. **Khubone T et al.** (2020) Electronic health information systems to improve disease diagnosis and management at point-of-care in low and middle income countries: a narrative review. *Diagnostics (Basel)* **10**, E327.
3. **Kim HJ et al.** (2000) Permutation tests for Jointpoint regression with applications to cancer rates. *Statistics in Medicine* **19**, 335–351.
4. **Werneck GL et al.** (2020) The COVID-19 pandemic in Brazil: chronicle of a health crisis foretold. *Cadernos de Saude Publica* **36**, e00068820.
5. **Fortaleza CMCB et al.** (2020) Taking the inner route: spatial and demographic factors affecting vulnerability to COVID-19 among 604 cities from inner São Paulo State, Brazil. *Epidemiology and Infection* **148**, e118. Published 2020 Jun 19.
6. **Requia WJ et al.** (2020) Risk of the Brazilian health care system over 5572 municipalities to exceed health care capacity due to the 2019 novel coronavirus (COVID-19). *Science of the Total Environment* **730**, 139144.
7. **Lobo AP et al.** (2020) COVID-19 epidemic in Brazil: where are we at?. *International Journal of Infectious Diseases* **97**, 382–385.
8. **Moreira RS.** COVID-19: intensive care units, mechanical ventilators, and latent mortality profiles associated with case-fatality in Brazil. *Cadernos de Saúde Pública*, **36**, e00080020. Epub May 18, 2020.
9. **Croda J et al.** (2020) COVID-19 in Brazil: advantages of a socialized unified health system and preparation to contain cases. *Revista da Sociedade Brasileira de Medicina Tropical* **53**, e20200167.
10. **The Lancet** (2020) COVID-19 in Brazil: ‘So what?’. *Lancet (London, England)* **395**, 1461.
11. **Hallal P et al.** (2020) Remarkable variability in SARS-CoV-2 antibodies across Brazilian regions: nationwide serological household survey in 27 states. *medRxiv* 2020.05.30.20117531; doi: [10.1101/2020.05.30.20117531](https://doi.org/10.1101/2020.05.30.20117531).
12. **Pabbaraju K et al.** (2020) A public health laboratory response to the COVID-19 pandemic. *Journal of Clinical Microbiology*. (published online ahead of print, 2020 Jun 8) JCM.01110-20. doi:10.1128/JCM.01110-20.

Two hundred days of COVID-19 in São Paulo State, Brazil

cambridge.org/hyg

From the Field

Cite this article: de Almeida GB *et al* (2020). Two hundred days of COVID-19 in São Paulo State, Brazil. *Epidemiology and Infection* **148**, e295, 1–4. <https://doi.org/10.1017/S0950268820002927>

Received: 29 June 2020

Revised: 27 October 2020

Accepted: 27 November 2020

Key words:

COVID-19; emerging infections; epidemics; epidemiology; infectious disease epidemiology

Author for correspondence:

Gabriel Berg de Almeida,
E-mail: gb.almeida@unesp.br

Gabriel Berg de Almeida^{1,2} , Micheli Pronunciate², Rejane Maria Tommasini Grotto^{1,3} , Edmur Azevedo Pugliesi⁴, Raul Borges Guimarães⁴, Thomas Nogueira Vilches⁵, Renato Mendes Coutinho⁶, Rafael de Castro Catão⁷, Claudia Pio Ferreira⁸ and Carlos Magno Castelo Branco Fortaleza^{1,2}

¹Clinical Hospital of Botucatu Medical School (HCFMB), Botucatu, Brazil; ²Department of Infectious Diseases, São Paulo State University (UNESP), Botucatu Medical School, Botucatu, Brazil; ³São Paulo State University (UNESP), School of Agriculture, Botucatu, Brazil; ⁴São Paulo State University (UNESP), School of Technology and Sciences, Presidente Prudente, Brazil; ⁵University of Campinas (UNICAMP), Institute of Mathematics, Statistics and Scientific Computation, Campinas, Brazil; ⁶Federal University of ABC (UFABC), Center for Mathematics, Computation and Cognition, Santo André, Brazil; ⁷Department of Geography, Federal University of Espírito Santo (UFES), Vitória, Brazil and ⁸São Paulo State University (UNESP), Institute of Biosciences, Botucatu, Brazil

Abstract

Two hundred days after the first confirmed case of COVID-19 in Brazil, the epidemic has rapidly spread in metropolitan areas and advanced throughout the countryside. We followed the temporal epidemic pattern at São Paulo State, the most populous of the country, the first to have a confirmed case of COVID-19, and the one with the most significant number of cases until now. We analysed the number of new cases per day in each regional health department and calculated the effective reproduction number (R_t) over time. Social distance measures, along with improvement in testing and isolating positive cases, general population mask-wearing and standard health security protocols for essential and non-essential activities, were adopted and impacted on slowing down epidemic velocity but were insufficient to stop transmission.

Background and epidemiology

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) was first identified in Brazil on 25 February 2020, in São Paulo [1]. The most populous city of Latin America was the route for COVID-19 importation, mainly from the USA and Italy [2]. It took 17 days for Brazil to reach a hundred cases, mainly reported in the capitals highly connected by airports and with an intensive flux of people. Two hundred days after the first confirmed case, the epidemic rapidly spreads across the country, and the disease advances through the interior.

São Paulo implemented state-wide quarantine measures quite early in the epidemic course. On 24 March, the government adopted a social distance recommendation for all people associated with closing trade and non-essential services. The decree suggested that people's movement should be limited to the immediate needs for food and health care. Despite the enforces, adherence to safety protocols was more significant in the capital, and COVID-19 took the inner route in São Paulo State [3]. Two patterns of disease dispersion were described by Fortaleza *et al.* [4]: one by contiguity, in which the virus spreads through areas of conurbation, and another hierarchical (long-distance spread through elementary spatial structures, such as highways, to cities with several degrees of connectivity).

São Paulo State is divided into 17 health departments with respect to epidemiological control, each one represented by a major city: Araraquara, Araçatuba, Baixada Santista, Barretos, Bauru, Campinas, Franca, São Paulo Metropolitan Area, Marília, Piracicaba, Presidente Prudente, Registro, Ribeirão Preto, São João da Boa Vista, São José do Rio Preto, Sorocaba and Taubaté. COVID-19 was introduced in different moments and behaved in different ways in each one of these Regional Health Departments (DRS, in Portuguese). We aimed to study COVID-19 advance in all these regions by analysing new confirmed cases per day (after the first case of COVID-19 in Brazil) and calculating the effective reproduction number (R_t) of SARS-CoV-2 over time. Also, daily new cases of Severe Acute Respiratory Illness (SARI) and its R_t number were estimated as an alternative way to follow the temporal evolution of the disease in a country still struggling to increase testing capacity [5].

Since 27 May, São Paulo State adopted a plan of quarantine measures ('Plano São Paulo'), which can be more restrictive or more flexible, considering the growth rate of COVID-19 cases and deaths, and bed occupancy rates in each DRS. All cities belonging to a DRS are ruled by the same quarantine measures, which were called phases. Phase 1 (red phase) is considered a

© The Author(s), 2020. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

CAMBRIDGE
UNIVERSITY PRESS

contamination phase, in which the rates of spread of the disease are high and, the capacity of the health system is close to its limit, with permission only for essential services. Phase 2 (orange phase) is considered an attention phase, with the possibility of some services opening. For commerce, limit to maximum occupancy 20% of the location's capacity, with reduced hours: 4 h in a row on all days of the week or 6 h in a row on 4 days of the week, always adopting standard and sector-specific protocols. Food courts are still banned in this phase. Phase 3 (yellow) is considered a controlled phase, with some flexibilisation. For example, commerce may open with maximum occupancy limited to 40% of total capacity and reduced hours (8 h). Finally, phase 4 (green), a partial opening phase, in which all services are allowed to open, respecting the limit of 60% of capacity and maintaining all specific protocols. Bars, restaurants, beauty salons and barbershops will only be open from phases 3 and 4, yellow and green.

Methods

We monitored the number of SARI and confirmed cases of COVID-19, over time, for each Regional Health Department of São Paulo State, Brazil [6]. These permitted us to calculate the effective reproduction number (R_t) for COVID-19 in each of these regions and evaluate the evolution of the epidemic using the methodology proposed by Wallinga and Lipsitch [7]. Unfortunately, data from Brazil does not distinguish imported cases from local cases, making it inviable to use more recent methodologies to estimate R_t [8]. Since 20 March, community transmission of COVID-19 was declared for the whole country, impacting notifications.

Data were obtained from the national database SIVEP-Gripe, which registers all severe hospitalised cases of SARI and identifies confirmed COVID-19 cases, and covers the period from the date of the first confirmed case of COVID-19 (25 February) until 200 days after. A nowcasting procedure [9] was performed to correct delay in notifications that span 40 days before the last case. The last week was ignored in the analysis. After that, the data were smoothed using a moving average with a window of 7 days.

Results

Figure 1 shows the time evolution of new cases, and R_t for both COVID-19 confirmed cases and SARI. The first case of COVID-19 was reported in the DRS of São Paulo Metropolitan Area on 25 February. For SARI, we reported the cases beginning on 15 March, when a change of protocol was done to englobe COVID-19 on SARI notifications. The results obtained for COVID-19 are shown in full lines and for SARI in dashed lines. Vertical lines have been marked to signal the dates for changes in social distancing protocols and trade functioning. The 24 March (purple vertical line) represents the generalised quarantine for the state as a whole. From 27 May on, the São Paulo plan was implemented. Coloured lines indicate the phase in which each DRS belongs. Phases 1, 2 and 3 are represented by the colours red, orange and yellow. None of the DRS was in the reopening phase (phase 4, green) at the end of this study.

One can notice the difference of scales for COVID-19 confirmed cases and SARI notifications to follow the disease's spatial-temporal dynamics in each location. The confirmed cases of COVID-19 struggle with the country's capacity to acquire quick tests, additionally to the shortages of molecular testing supplies. SARI cases can vary from twice as high as COVID-19 cases

(e.g. Franca and Registro) to four times higher (e.g. Araquara and Barretos). Grotto *et al.* [5] showed that molecular diagnosis increased in São Paulo State over the epidemiological weeks, but it does not match the much higher increase in the number of cases, challenging diagnostic capacity and, therefore, accurate and timely health surveillance.

The trends in increasing or decreasing epidemic velocity are captured by both measures SARI and COVID-19 (see R_t curves). Oscillations can be clearly observed on the COVID-19 curve of new cases.

Discussion

Our results confirm the spatial-temporal dispersion of COVID-19 over São Paulo State described by Fortaleza *et al.* [3, 4]. Higher numbers of confirmed COVID-19 cases are seen earlier in São Paulo Metropolitan Area DRS, accompanied by conurbation areas, such as Campinas DRS and Baixada Santista DRS. On 27 May, they belonged to the red phase, indicating high rates of disease spreading and high hospital bed occupancy rates. This is evidence of the earlier introduction of SARS-CoV-2 in these regions and may explicit the contiguity model of disease dispersion.

It is interesting to highlight that the social distance measures and improvement in testing, mask-wearing and standard health security protocols, were adopted and impacted slowing down epidemic velocity in all DRS, but in different moments. In São Paulo Metropolitan Area, protocol adherence was greater in March and April, as we can see the R_t 's reduction over time. In the countryside, rules were not strictly followed in this very first moment because the number of cases and deaths were still not so alarming, giving the population a false security sensation. São Paulo State general quarantine started on 24 March and was extended until 27 May, when the São Paulo plan started. Over this period, all DRS showed R_t values close to 1, but only sometimes below 1. This was not sufficient to stop the epidemic, as the number of confirmed cases kept growing, but was able to slow down dissemination.

When taking special attention to more recent R_t values (June, for instance), we can see the interior of São Paulo State is at a critical phase of the epidemic. Inner regions such as Marília, Araraquara and Barretos DRS show R_t values sometimes much higher than 1. Bauru shows similar behaviour. An increase in the R_t in these locations is mostly due to more recent disease introduction, not following social distancing measures, and local issues on testing and isolating positive cases. This scenario is not compatible with a plan of reopening commerce and industry activities, which have been induced by some mayors who insist on questioning the São Paulo State quarantine plans.

Finally, we observed that we can follow COVID-19 epidemic behaviour by following SARI notifications. Obviously, there is a great difference on the scale of numbers. As commented before, this might be secondary to a lack of diagnostic power and a delay in the diagnosis of confirmed cases. Anyhow, as we can see in Figure 1, in all DRS, full lines and dashed lines run together and represent remarkably similar curves. A low-to-medium income country with serious structural issues, such as Brazil, looking at SARI incidence and how it increased in different areas might represent an alternative way to estimate the real epidemic's numbers.

Similar studies estimated R_t considering imported and local transmission cases because the first one seems to be particularly important for a newly introduced disease [10, 11]. Such methodology cannot be applied to Brazil's data since this information is not captured in the SARI mandatory notification form. However,

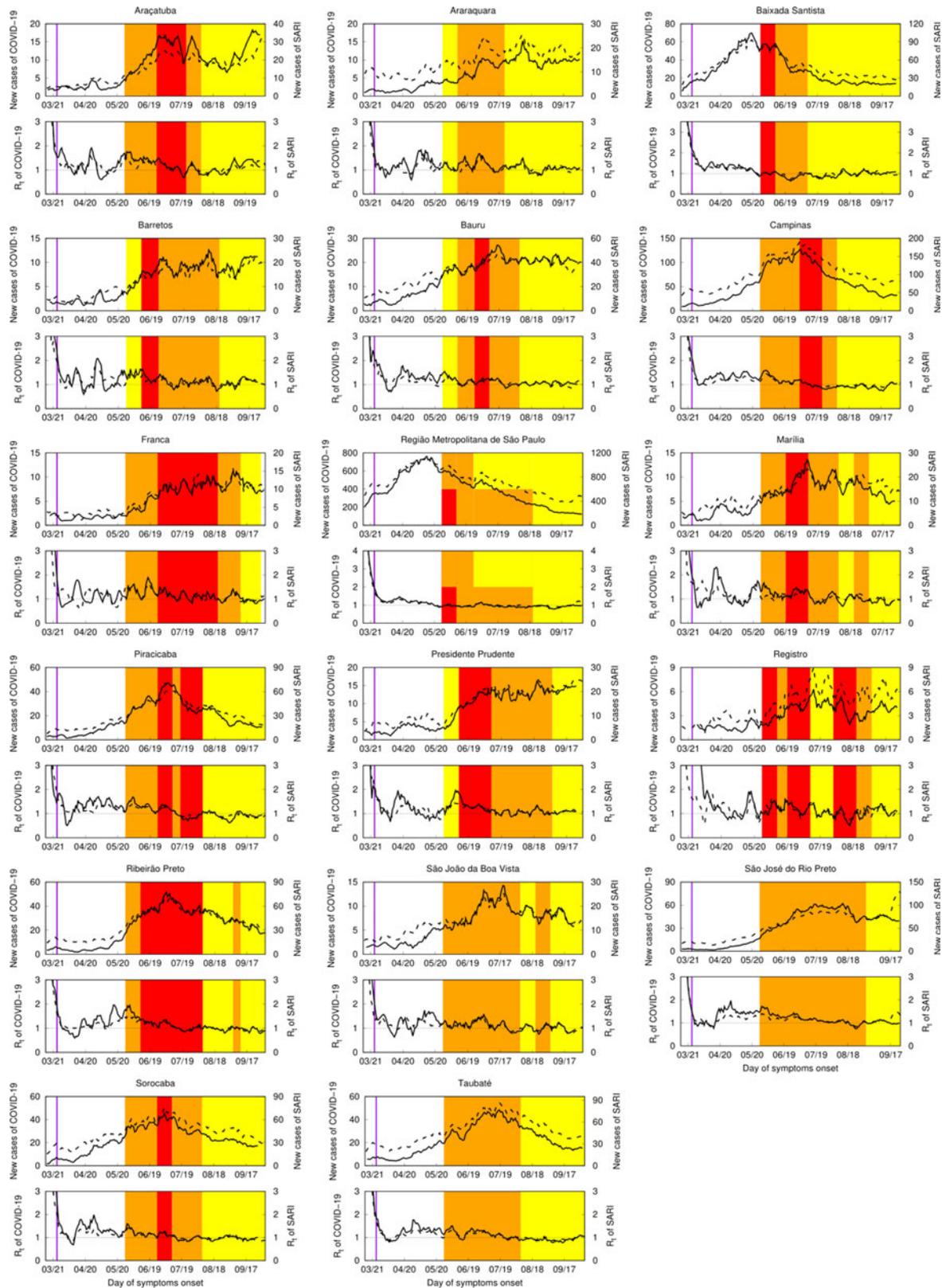


Fig. 1. Epidemic evolution of COVID-19 in the Regional Health Departments (DRS) of São Paulo State, Brazil. In each panel identified with the DRS name, top curves correspond to new confirmed cases of COVID-19 in full lines and new notifications for severe acute respiratory illness (SARI) in dashed lines. Bottom curves correspond to the temporal evolution of R_t for COVID-19 (full lines) and SARI (dashed lines). Vertical purple lines mark March 24th, the day of the first quarantine recommendations for São Paulo State. Red, orange, or yellow painted periods represent phases of the São Paulo plan. Phase one is addressed in red, phase 2 in orange, and phase 3 in yellow. None of DRS achieved phase 4 (green), a reopening phase, during the period of this study.

the obtained R_t values at the different phases of the epidemic are in the range observed in other studies [10, 11]. Also, as in other countries, slowing down on epidemic spreading was observed when the non-pharmacological measures were introduced, but it was not enough to control the epidemic, differently from many countries. Local issues on epidemic control are still a challenge for Brazilian studies. Delays in notifications, poorly computerised systems, difficulties in contact tracing, open-source data and political and economic fragilities are some. Finally, lock-down was never performed in any location of São Paulo state.

Studying R_t values and relating to the number of confirmed new cases and SARI cases permitted us to evaluate quarantine plans and their impact on disease spreading over time. Along with universal mask-wearing and testing and isolating positive cases, social distance measures were able to diminish epidemic velocity, impacting the reduction of R_t . Still, they were insufficient to stop transmission, as the R_t was mostly established over one, and the number of cases kept growing. Today, after 200 days since the first confirmed case of COVID-19 in Brazil, the São Paulo State situation is still alarming. Although many regions started showing a reduction in the number of new cases since August, R_t on most of these locations is greater than 1, making it clear new strategies on public health and epidemic control urges.

Acknowledgements. CPF acknowledges grants 2018/24058-1 and 2019/22157-5, São Paulo Research Foundation (FAPESP).

Data availability statement. The authors state that the database used in the analyses can be available as a supplementary file to the paper or provided to interested researchers upon reasonable request.

References

1. **Rodriguez-Morales AJ *et al.*** (2020) COVID-19 in Latin America: the implications of the first confirmed case in Brazil. *Travel Medicine and Infectious Diseases* **10**, 16–13.
2. **Candido DDS *et al.*** (2020) Routes for COVID-19 importation in Brazil. *Journal of Travel Medicine* **27**, 1–3. doi: 10.1093/jtm/taaa042.
3. **Fortaleza CMCB *et al.*** Taking the inner route: spatial and demographic factors affecting vulnerability to COVID-19 among 604 cities from inner São Paulo State, Brazil. *Epidemiology and Infection* **148**, 1–5.
4. **Fortaleza CMCB *et al.*** (2020) Elementary spatial structures and dispersion of COVID-19: health geography directing responses to public health emergency in São Paulo State, Brazil. *medRxiv* 2020.04.26.20080895. Published online. doi: 10.1101/2020.04.26.20080895.
5. **Grotto R *et al.*** (2020) Increasing molecular diagnostic capacity and COVID-19 incidence in Brazil. *Epidemiology and Infection* **2020**, 1–9.
6. **Coutinho R *et al.*** (2020) Time series of COVID-19 severe cases and deaths, corrected using nowcasting, by Regional Health Department of São Paulo State, Brazil, as of 2020-06-11. figshare. Dataset. <https://doi.org/10.6084/m9.figshare.12543836.v1>.
7. **Wallinga J and Lipsitch M** (2007) How generation intervals shape the relationship between growth rates and reproductive numbers. *Proceedings of the Royal Society Biological Sciences* **274**, 599–604.
8. **Thompson RN *et al.*** (2019) Improved inference of time-varying reproduction numbers during infectious disease outbreaks. *Epidemics* **29**, 100356.
9. **McGough S *et al.*** (2020) Nowcasting by Bayesian smoothing: a flexible, generalizable model for real-time epidemic tracking. *PLOS Computational Biology* **16**(4), e1007735. <https://doi.org/10.1371/journal.pcbi.1007735>.
10. **Shim E *et al.*** (2020) Transmission potential and severity of COVID-19 in South Korea. *International Journal of Infectious Diseases* **93**, 339–344.
11. **Tariq A *et al.*** (2020) Real-time monitoring the transmission potential of COVID-19 in Singapore, March 2020. *BMC Medicine* **18**, 166.

Conclusões

A análise crítica dos artigos publicados e replicados nessa tese nos permite realizar uma projeção com exploração de cenários hipotéticos da dinâmica de transmissão da Covid-19 em cidades pequenas e médias localizadas no interior do Brasil, além de descrever e interpretar a dispersão geográfica precoce da Covid-19 no Brasil e no Estado de São Paulo, das metrópoles para municípios menores. O acompanhamento da evolução do número de casos ao longo do tempo no Estado de São Paulo também nos permitiu avaliar a velocidade de transmissão do SARS-CoV-2 e relacioná-la às medidas restritivas adotadas pelo governo estadual através do Plano São Paulo.

Através de um modelo matemático dinâmico não autônomo de transmissão de doença infecciosa pudemos demonstrar que diferentes medidas de controle devem ser tomadas para diferentes cidades e, mais importante, cada cidade pode ter uma combinação ideal entre distanciamento social com testagem e isolamento dos casos positivos que controla a curva da epidemia e permite que os sistemas de saúde estejam preparados para o pico do número de casos. Ademais, a análise de *clustering* e aglomeração de municípios com padrões semelhantes de evolução das curvas epidêmicas pode auxiliar a propor estratégias unificadas de controle e contingenciamento. Demonstramos, através da metodologia científica e de equações diferenciais, que o distanciamento social em diferentes graus deve ser estabelecido nas diferentes localidades.

A análise espacial da disseminação do SARS-CoV-2, por sua vez, pôde destacar as principais vias de dispersão da doença e a fragilidade dos municípios com relação às suas características sociodemográficas. Foi possível evidenciar

os *hotspots* e as principais vias de dispersão da doença da capital para o interior do Estado de São Paulo. A existência de duas formas diferentes de dispersão, por contiguidade e hierárquica, pode aventar estratégias alternativas para controlar o espalhamento da doença pelo território paulista. Este trabalho mostrou que foi possível definir a rota de propagação da Covid-19 no Estado de São Paulo pela hierarquia das cidades, o que significa que a propagação da epidemia não segue um processo de difusão em todas as localidades, mas atinge os municípios a partir de sua relevância e conectividade. A partir de então, a epidemia se espalha para cidades contíguas seguindo um processo padrão de difusão. Afirmamos que essas cidades são responsáveis pela chegada da epidemia no interior do Estado de São Paulo e demandam atenção.

O acompanhamento da introdução dos primeiros casos confirmados em todo o Brasil, em conjunto com a análise geográfica dos novos laboratórios credenciados para o diagnóstico da Covid-19 também evidenciou uma dispersão da doença das metrópoles para o interior, ao mesmo tempo em que destacou a corrida brasileira para o aumento da sua capacidade de diagnóstico molecular do SARS-CoV-2. Apesar do aumento considerável do número de laboratórios públicos, esse esforço parece não ter sido suficiente. O avanço da doença para áreas no interior, mais remotas e menos desenvolvidas, ao mesmo tempo em que um número importante de laboratórios foi certificado em regiões centrais, urbanizadas e industriais, pode evidenciar uma fragilidade do sistema de vigilância em saúde comprometendo, portanto, a capacidade de resposta rápida em gestão de saúde pública.

Por fim, o estudo dos valores de R_t (através do número de novos casos confirmados de Covid-19 e de casos de SRAG) nos permitiram relacionar e

avaliar os planos de quarentena e seu impacto na propagação da doença ao longo do tempo no Estado de São Paulo. Em conjunto com o uso universal de máscaras pela população e a testagem e isolamento de casos positivos, medidas de distanciamento social foram capazes de diminuir a velocidade da epidemia, impactando a redução do valor de R_t . Ainda assim, todas essas medidas foram insuficientes para interromper o ciclo de transmissão do SARS-CoV-2, já que o valor de R_t ficou estabelecido na maior parte do tempo acima de um ($R_t > 1$), e o número de casos se manteve em ascendência.

Nossos estudos compreenderam uma primeira fase de transmissão e dispersão da Covid-19 no Brasil, quando ainda grande parte da população era susceptível, não havia imunização disponível e as medidas de distanciamento social e restrição das atividades nos municípios apresentavam-se como a maneira mais eficaz de contingenciamento da epidemia. A reabertura do comércio e dos serviços, a autorização para funcionamento de casas noturnas, *shows* e bares, e a volta às aulas das escolas e universidades, por exemplo, foi uma discussão que se estendeu por vários meses no Brasil, mesmo quando o número de novos casos ainda se mantinha elevado e a transmissão comunitária sustentada. Realizar estudos epidemiológicos, de modelagem matemática e estatística teve papel fundamental na demonstração, através de evidências científicas, de que determinadas intervenções em saúde pública poderiam ser necessárias para conter a evolução da epidemia e evitar a sobrecarga dos serviços de saúde.

Nesse contexto, é imprescindível ressaltar que todos os estudos foram realizados, em maior ou menor grau, a partir de dados secundários do sistema de vigilância epidemiológica. No Brasil, a rede de vigilância ainda possui diversas

fragilidades, o que pode interferir na qualidade dos dados e na capacidade de processamento dos mesmos. Os atrasos em notificações, o não preenchimento de informações essenciais e o acesso desigual a testes e serviços de saúde pela população, por exemplo, podem ser causa de subnotificação, e portanto, subestimar os dados reais, impactando na leitura dos resultados. Ressaltamos a importância de um sistema de vigilância eficaz, informatizado e com boa capilaridade, que possa trazer informações precisas, e, portanto, serem capazes de desencadear respostas rápidas e adequadas em saúde pública.

Hoje, após dois anos completos de pandemia, os desafios ainda são grandes e os estudos de modelagem continuam a contribuir. Apesar da disponibilidade de imunizantes, principalmente a partir de janeiro de 2021, o surgimento de novas variantes de preocupação (VOCs), com escape imunológico, maior potencial de transmissibilidade, aliados a ineficácia do Governo Federal em estruturar um plano nacional de vacinação coordenado, novas ondas epidêmicas ainda podem ocorrer, com a possibilidade de superlotação de leitos de enfermaria e UTI e acometimento de grupos populacionais mais vulneráveis, como as crianças, que ainda possuem uma baixa taxa de cobertura vacinal e os imunossuprimidos, que apresentam resposta vacinal muitas vezes insatisfatória.

Os estudos matemáticos, geográficos e epidemiológicos fazem parte de uma análise contínua no curso de uma epidemia. É evidente que todos os métodos possuem suas fragilidades, referentes principalmente a qualidade dos dados disponíveis para análise e à natureza intrinsecamente estocástica do comportamento de uma epidemia. Entretanto, a leitura da realidade através da

ciência pode contribuir sobremaneira para a tomada de decisão da gestão em saúde.

Referências

1. Novel Coronavirus (2019-nCoV) – Situation Report 1 (21 January 2020) - World Health Organization (WHO), disponível em https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200121-sitrep-1-2019-ncov.pdf?sfvrsn=20a99c10_4
2. Grassly NC, Fraser C. Mathematical models of infectious disease transmission. *Nat Rev Microbiol.* 2008 Jun;6(6):477-87. doi: 10.1038/nrmicro1845. PMID: 18533288; PMCID: PMC7097581.
3. Kermack, W.O., McKendrick, A.G., 1927. A contribution to the mathematical theory of epidemics. *Proc. Roy. Soc. Lond. A* 115, 700–721.
4. Novel Coronavirus (2019-nCoV) – Situation Report 10 (30 January 2020) - World Health Organization (WHO), disponível em https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200130-sitrep-10-ncov.pdf?sfvrsn=d0b2e480_2
5. Rodriguez-Morales AJ et al. (2020) COVID-19 in Latin America: the implications of the first confirmed case in Brazil. *Travel Medicine and Infectious Diseases* 10, 16–13.
6. Candido DDS et al. (2020) Routes for COVID-19 importation in Brazil. *Journal of Travel Medicine* 27, 1–3. doi: 10.1093/jtm/taaa042.
7. Sousa GJB, Garces TS, Cestari VRF, Moreira TMM, Florêncio RS, Pereira MLD. Estimation and prediction of COVID-19 cases in Brazilian metropolises. *Rev Lat Am Enfermagem.* 2020 Jun 26;28:e3345. doi: 10.1590/1518-8345.4501.3345. PMID: 32609282; PMCID: PMC7319758.
8. Yang HM, Lombardi Junior LP, Castro FFM, Yang AC. Mathematical model describing COVID-19 in São Paulo, Brazil - evaluating isolation as control mechanism and forecasting epidemiological scenarios of release. *Epidemiol Infect.* 2020 Jul 20;148:e155. doi: 10.1017/S0950268820001600. PMID: 32684175; PMCID: PMC7378372.
9. Canabarro A, Tenório E, Martins R, Martins L, Brito S, Chaves R. Data-driven study of the COVID-19 pandemic via age-structured modelling and prediction of the health system failure in Brazil amid diverse intervention strategies. *PLoS One.* 2020 Jul 30;15(7):e0236310. doi: 10.1371/journal.pone.0236310. PMID: 32730352; PMCID: PMC7392258.

10. Tang Y, Serdan TDA, Masi LN, Tang S, Gorjao R, Hirabara SM. Epidemiology of COVID-19 in Brazil: using a mathematical model to estimate the outbreak peak and temporal evolution. *Emerg Microbes Infect.* 2020 Dec;9(1):1453-1456. doi: 10.1080/22221751.2020.1785337. PMID: 32552473; PMCID: PMC7473191.
11. Vynnycky E, White RG - An introduction to infectious disease modelling – Oxford University Press, New York, 2010.