# SUBNATIONAL ANALYSIS OF THE INITIAL PHASE OF THE COVID-19 EPIDEMIC IN BRAZIL

**Anonymous authors**
Paper under double-blind review

## ABSTRACT

Since the beggining of pandemic, Brazil has reported the second highest number of COVID-19 deaths in the world. Here we characterise the early transmission that seeded the country-wide spread of the disease, and assess attempts to attenuate the spread through implementing non-pharmaceutical interventions (NPIs) at subnational level. The analysis presented uses a Bayesian hierarchical approach to model transmission based on mortality data. The statistical model encodes a causal inferential bias for generic infectious disease transmission — deaths are generated by infections which arise from earlier infections. As transmission is heterogeneous at subnational level, from differences such as the timing of seeding and hospital capacities, this is modelled by partially pooling parameters across geographic regions, using state-level mobility covariates for the reproduction number ($R_t$), and through inference of region-specific epidemiological parameters. We report extensive heterogeneity in the initial epidemic trajectory across Brazil underscoring the importance of sub-national analyses in understanding asynchronous state-level epidemics underlying the national spread and burden of COVID-19.

## 1 INTRODUCTION

Since the emergence of the SARS-CoV-2 in China in December 2019, global spread has been rapid, with over 660 million cases and over 6.5 million deaths reported globally as of the January 2023 (WHO, 2023). One of the countries particularly badly affected was Brazil – since report of its first case on 26th Feb 2020, its epidemic has grown quickly, with the country now reporting over 36 million cases and almost 700,000 deaths (WHO, 2023). In response to significant spread and community transmission of the virus within the country in the early months of 2020, Brazilian state and city officials implemented extensive public health measures to reduce the transmission of COVID-19, including declaring a state of emergency, mandating the closure of retail and service businesses, restricting transportation, and closing schools. Specific packages of interventions have been decided at the state level, with substantial variation between states in their comparative timing and the extent to which measures have been adopted (cep, 2020).

A better understanding of the epidemiological origin and the impact of those initial interventions deployed is required to guide future policy decisions aimed at preventing worsening of the public health emergencies. Motivated by this, we extend a previously published semi-mechanistic Bayesian hierarchical model of COVID-19 epidemiological dynamics from Flaxman et al. (2020) to assess the impact of interventions aimed at curbing transmission of COVID-19 across Brazil. In our framework we estimate the number of deaths, infections and transmission as a function of patterns in human mobility. We utilise this framework to explore the epidemiological situation in the early stages of the epidemic in detail at the state level, understand the highly heterogeneous spread of the virus across the country and understand the insufficiency of the initial response to the virus across states.

## 2 METHODS AND DATA

**Data:** Our model utilises daily consolidated deaths data at the state level to infer epidemiological characteristics of viral spread and transmission. The analyses presented here are based on daily death figures, disaggregated by state, from two sources that are published by the Brazilian Ministry of Health — *Painel Coronavírus*(cov, 2020) and *SIVEP-Gripe*(Ministério da Saúde, 2020). Analyses

are restricted to the period ending 9th May 2020, which is sufficiently long after NPIs in the first wave were implemented across Brazil for us to draw conclusions about their effectiveness. To account for under-reporting of deaths (i.e. deaths that either occur in the community or that occur in a hospital setting but fail to be recorded), additionally to using two datasets mentioned earlier, we also undertake a series of sensitivity analysis assuming different levels of death underreporting across states (see Appendix for further details).

**Model:** We adapt and extend a previously published Bayesian semi-mechanistic model of COVID-19 transmission and mortality from Flaxman et al. (2020), updating the model to explicitly incorporate population level metrics of mobility (Google Mobility data) that have previously been shown to reflect patterns of transmission across a variety of settings (Unwin et al., 2020) and a weekly autoregressive process intended to capture variation in patterns of transmission between states above and beyond that reflected in their comparative patterns of mobility (e.g. that could reflect variation in adoption of individual-level behaviours that would modify the impact of mobility reductions on transmission). Parameters governing the model are jointly estimated with partial pooling for 18 Brazilian states to evaluate the impact of control interventions on SARS-CoV-2 transmission and to explore variation in epidemic trajectories between states. Code for the model is available in an online repository . The key structural elements of the model are as follows.

The model is structured with a causal inferential bias, in which observed deaths $d$ are generated by earlier infections $i$:

$$d_{t,m} = \text{ifr} \sum_{\tau<t} i_{\tau,m}\pi_{t-\tau,m}\,, \tag{1}$$

Infections and deaths are indexed in time $t$, for earlier times $\tau$, and for state $m$. The probability of death following infection is given by the infection fatality ratio (ifr), and has a temporal delay distribution $\pi$. Infections in turn are generated from earlier infections

$$i_{t,m} = R_{t,m} \sum_{\tau<t} i_{\tau,m} g_{t-\tau,m}\,, \tag{2}$$

where $R_t$ denotes the reproductive number, which is sub-critical for values less than 1, otherwise supercritical in infections. The temporal distribution $g$ is the generation interval between infections.

The time-varying reproduction number $R_{t,m}$ is estimated at the state level, using a similar modelling approach to Unwin et al. (2020). Specifically the functional form for $R_{t,m}$ is specified in the following way:

$$R_{t,m} = R_{0,m} \cdot \sigma(-(\sum_{k=1}^{4}(\alpha_k + \beta_{m,k})X_{t,m,k}) - \epsilon_{m,w_m(t)}), \tag{3}$$

with $\sigma(x) = 2\exp(x)/(1+\exp(x))$ denoting the logistic function. The index $k$ labels the four different mobility covariates in $X_{t,m,k}$ used for each state ($m$) and time ($t$) from Google Mobility data (see Appendix Fig. 2). $\alpha_k$ is a coefficient linking each of these mobility metrics to transmission that is shared across states, and $\beta_{m,k}$ is an additional coefficient allowing a state-specific relationship between a particular mobility covariate and transmission. These mobility report data were used to estimate the effects of different interventions over time. The report provides the estimated percentage of change on movements of places such as retail, recreation, transit stations, workplaces or residential comparing to a baseline.

Prior distributions in Eq. 3 were specified as follows:

$$\alpha_k \sim (0, 0.5),\ \beta_{m,k} \sim (0, \gamma), \quad \text{with} \quad \gamma \sim (0, 0.5), \tag{4}$$

and the variable $\epsilon_{m,w_m(t)}$ is a weekly auto regressive process with two lags (AR(2)), centred around 0, which intend to capture extra variation between states that is not fully explained by mobility alone.

In addition to this elaboration to the formulation of $R_t$, we utilised Brazil-specific estimates of the key parameters governing the model. The distribution of times from infection to death was estimated using patient level data from SIVEP-Gripe dataset(Ministério da Saúde, 2020; Hawryluk et al., 2020). Based on these results, we model the infection-to-death distribution $\pi$ as a sum of the infection-to-symptom and symptom-to-death distributions,

$$\pi \sim \text{Gamma}(5.1, 0.86) + \text{Gamma}(13.4, 0.7)\,. \tag{5}$$

The infection fatality ratios used are based on those in Verity et al. (2020a), adapted to the demographic composition for each state based on the distributions of ages of individuals. Based on Walker et al. (2020) we adapted our IFR estimates for each state in an income-dependent manner. Specifically, we assumed that the state with the highest income (São Paulo) has a quality of care identical to that observed in the UK ((Verity et al., 2020a)), and that the state with the lowest income (Maranhão) had significantly worse healthcare outcomes - more similar to those that would be expected in a Lower Middle Income Country (Walker et al. (2020)). For the other states, we linearly interpolate the age-specific infection fatality probabilities based on state-level average income (ibg, 2019a). These age-specific infection fatality probabilities are then combined with predictions of the age-distribution of infections to produce an overall, state-specific IFR. Because of the inherent uncertainty associated with these modifications, we undertake a number of sensitivity analyses examining how IFR-related assumptions qualitatively impact our results.

In this work an extension of the hierarchical model from Flaxman et al. (2020) is adopted to reflect the uncertainty about underreported deaths. We address the effect of underreporting in the data set by setting a prior distribution to death underreporting $\psi \sim beta(\theta, \rho)$. The hyperparameters of the beta density are fixed in order to reflect in the mode the desired underreporting rate — see Figure 3 in the Appendix. Daily deaths $D_{t,m}$ are observed for days $t \in \{1, \ldots, n\}$ and Brazilian states $m \in \{1, \ldots, M\}$. These daily deaths are modelled using a positive real-valued function $d_{t,m} = \mathbb{E}[D_{t,m}\psi]$ that represents the expected number of deaths attributed to COVID-19, taking into account the designated underreported rate $\psi$. Daily deaths $D_{t,m}$ are assumed to follow a negative binomial distribution with mean $d_{t,m}$ and variance $d_{t,m} + \frac{d_{t,m}^2}{\phi}$, where $\phi$ follows a normal distribution, i.e.

$$D_{t,m} \sim \text{Negative Binomial}\left(d_{t,m}, d_{t,m} + \frac{d_{t,m}^2}{\phi}\right),\ \phi \sim N(0, 5) \tag{6}$$

$$\tag{7}$$

## 3 Results

Across the 18 Brazilian states with more than 50 deaths over the time period up to 9th May, we estimate that implemented NPIs had a substantial impact on transmission and spread of SARS-CoV-2, reducing $R_t$ from greater than three to 1.5 (95% CI 1.3-1.9) on average. Figure 1 shows the model estimated time-varying reproduction number ($R_t$) for 2 states – São Paulo and Maranhão (see Appendix for analyses of 18 states). São Paulo and Rio de Janeiro had the highest numbers of deaths, Amazonas had the highest predicted attack rates, while the epidemic in Maranhão is comparatively nascent in progression. Across all states, our results highlight that $R_t$ has dropped consistently following the implementation of public health interventions. For the 2 states considered in Figure 1, mobility indicators dropped by 33% on average by May 9th across Brazil compared to a baseline derived from data on the same date in the preceding year, and $R_t$ declined by greater than 50% on average, for example to 0.8 (95% CI: 0.6%-1.0%) in Maranhão. Despite these substantial reductions in $R_t$ however, our results also indicate it is unlikely that these measures have brought $R_t$ consistently below 1. This was observed across all states analysed and highlights the insufficiency of the initial NPIs implemented at a state level in controlling and preventing further growth of the COVID-19 epidemic.

Similar variation was observed when the predicted attack rates (the total proportion of the population infected over the time period considered) were assessed, with our analyses suggesting attack rates by the 9th May 2020 as high as 8% in Amazonas to as low as 0.1% in Minas Gerais (Table 1, more results in the Appendix).These results are driven in part by modelling assumptions regarding the extent of death underreporting and the assumed state-specific IFR and we therefore undertook a series of sensitivity analyses (Appendix 5) exploring different assumptions surrounding state-level IFR and the extent of death underreporting.

## 4 Discussion

Attempts to contain the spread of SARS-CoV-2 in the community in the initial, pre-vaccinations stages of the epidemic, have centred around the deployment of various NPIs that involve reducing

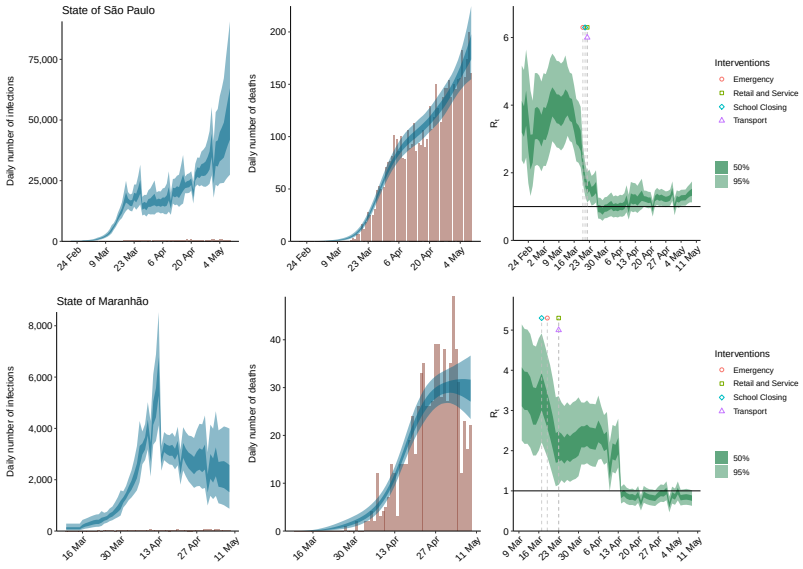


Figure 1: Estimates of infections, deaths and $R_t$ for 2 states. Left: daily number of infections, blue bands are predicted infections, dark blue 50% credible interval (CI), light blue 95% CI. Middle: daily number of deaths, brown bars are reported deaths, blue bands are predicted deaths, CI as in left plot. Right: time-varying reproduction number $R_t$. If the $R_t$ is above 1, the number of infections continues to grow. Icons are interventions shown at the time they occurred.

Table 1: Estimated infection fatality ratio (IFR), state population, reported deaths and deaths per million population, estimated number of infections in thousands, attack rate (AR), and time-varying reproduction number on 9th-May-2020 with 95% credible intervals, for selected states. More results available in the Appendix.

| State | IFR % | Population | Deaths | Deaths per million | Infections (thousands) | Attack rate % | $R_t$ |
|---|---|---|---|---|---|---|---|
| São Paulo | 0.7 | 46,289,333 | 5,142 | 111 | 1,370 [950, 1,880] | 3.0 [2.1, 4.1] | 1.4 [1.1, 1.7] |
| Rio de Janeiro | 0.8 | 17,366,189 | 4,564 | 263 | 1,050 [777, 1,380] | 6.1 [4.5, 8.0] | 1.1 [0.9, 1.3] |
| Amazonas | 0.8 | 4,207,714 | 1,446 | 344 | 337 [241, 454] | 8.0 [5.7, 10.8] | 1.2 [0.9, 1.4] |
| Maranhão | 1.0 | 7,114,598 | 745 | 105 | 121 [82, 172] | 1.7 [1.2, 2.4] | 0.8 [0.6, 1.0] |
| Paraíba | 1.2 | 4,039,277 | 161 | 40 | 59 [35, 92] | 1.5 [0.9, 2.3] | 2.1 [1.7, 2.5] |
| Minas Gerais | 1.0 | 21,292,666 | 153 | 7 | 31 [15, 54] | 0.1 [0.1, 0.3] | 1.5 [1.0, 2.0] |
| Amapa | 0.7 | 861,773 | 69 | 80 | 31 [15, 55] | 3.6 [1.7, 6.3] | 1.6 [1.2, 2.0] |

the number of contacts made between individuals (Ferguson et al.), in doing so attempting to disrupt chains of transmission, bring the reproduction number ($R_t$) below 1 and curb exponential growth of the epidemic. Using our framework, with $R_t$ parameterised as a function of Google mobility data (Aktay et al., 2020), we highlight the marked effect of NPIs on transmission in Brazil over the period considered here, which substantially slowed spread of the virus. Despite these reductions however, our results also highlight that the changes in population mobility covariates (which saw a 33% average reduction relative to baseline) were not stringent enough to reduce $R_t$ below 1 in many states.

Despite the high numbers of deaths reported for the country as a whole during the initial phase, there are also noticeable differences in the burden of COVID-19 experienced between states, with the distribution of deaths among states highly heterogeneous (Fig. 1 and Fig. 4). Our results support this geographical variation, highlighting specifically the variation in the likely timing of epidemic takeoff that has led to the asynchronous epidemics that have emerged across different states. This variation highlights that in the initial phase by the 9th of May, Brazilian states were at very different points in their respective epidemics. Such differences had material consequences for the likely evolution of the epidemic trajectories by state in the coming months. Moreover, they underscore the importance of granular, sub-national analyses in understanding the spread of SARS-CoV-2, revealing a level of heterogeneity that would be obscured by analyses conducted at the national level.

## REFERENCES

IBGE divulga o rendimento domiciliar per capita 2019, 2019a. URL https://agenciadenoticias.ibge.gov.br/agencia-sala-de-imprensa/2013-agencia-de-noticias/releases/26956-ibge-divulga-o-rendimento-domiciliar-per-c

IBGE Projeções da População, 2019b. URL https://www.ibge.gov.br/estatisticas/sociais/populacao/9109-projecao-da-populacao.html?=&t=resultados.

COVID-19 Observatory in Latin America and the Caribbean, 2020. URL https://www.cepal.org/en/topics/covid-19.

Painel Coronavírus, 2020. URL https://covid.saude.gov.br/.

Ahmet Aktay, Shailesh Bavadekar, Gwen Cossoul, John Davis, Damien Desfontaines, Alex Fabrikant, Evgeniy Gabrilovich, Krishna Gadepalli, Bryant Gipson, Miguel Guevara, Chaitanya Kamath, Mansi Kansal, Ali Lange, Chinmoy Mandayam, Andrew Oplinger, Christopher Pluntke, Thomas Harold Roessler, Arran Schlosberg, Tomer Shekel, Swapnil Vispute, Mia Vu, Gregory A Wellenius, Brian Williams, and Royce J Wilson. Google COVID-19 Community Mobility Reports: Anonymization Process Description (version 1.0). *ArXiv*, abs/2004.0, 2020.

Neil M Ferguson, Daniel Laydon, Gemma Nedjati-Gilani, N. Imai, K Ainslie, M. Baguelin, S. Bhatia, A. Boonyasiri, Z. M. Cucunuba, G. Cuomo-Dannenburg, Amy Dighe, Ilaria Dorigatti, Han Fu, K. Gaythorpe, Will Green, Arran Hamlet, W. Hinsley, L. C. Okell, S van Elsland, Hayley Thompson, Robert Verity, E. Volz, H Wang, Y Wang, P. G. T. Walker, C Walters, Peter Winskill, Charles Whittaker, C. A. Donnelly, Steven Riley, and A. C. Ghani. Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand (Report 9). Technical report. URL https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/news--wuhan-coronavirus/.

Seth Flaxman, Swapnil Mishra, Axel Gandy, H Juliette T Unwin, Helen Coupland, Thomas A Mellan, Harrison Zhu, Tresnia Berah, Jeffrey W Eaton, Pablo N P Guzman, Nora Schmit, Lucia Callizo, Kylie E C Ainslie, Marc Baguelin, Isobel Blake, Adhiratha Boonyasiri, Olivia Boyd, Lorenzo Cattarino, Constanze Ciavarella, Laura Cooper, Zulma Cucunubá, Gina Cuomo-Dannenburg, Amy Dighe, Bimandra Djaafara, Ilaria Dorigatti, Sabine van Elsland, Rich FitzJohn, Han Fu, Katy Gaythorpe, Lily Geidelberg, Nicholas Grassly, Will Green, Timothy Hallett, Arran Hamlet, Wes Hinsley, Ben Jeffrey, David Jorgensen, Edward Knock, Daniel Laydon, Gemma Nedjati-Gilani, Pierre Nouvellet, Kris Parag, Igor Siveroni, Hayley Thompson, Robert Verity, Erik Volz, Patrick GT Walker, Caroline Walters, Haowei Wang, Yuanrong Wang, Oliver Watson, Charles Whittaker, Peter Winskill, Xiaoyue Xi, Azra Ghani, Christl A. Donnelly, Steven Riley, Lucy C Okell, Michaela A C Vollmer, Neil M. Ferguson1, and Samir Bhatt. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature*, pp. 1–8, 2020. doi: https://doi.org/10.1038/s41586-020-2405-7.

Carlos G. Grijalva, Nele Goeyvaerts, Hector Verastegui, Kathryn M. Edwards, Ana I. Gil, Claudio F. Lanata, and Niel Hens. Peruvian social contact data (Version 1.0) [Data set]. Zenodo. 2017. URL http://doi.org/10.5281/zenodo.1215891.

Iwona Hawryluk, Thomas A Mellan, Henrique Hoeltgebaum, Swapnil Mishra, Ricardo P Schnekenberg, Charles Whittaker, Harrison Zhu, Axel Gandy, Christl A Donnelly, Seth Flaxman, et al. Inference of COVID-19 epidemiological distributions from Brazilian hospital data. *Journal of the Royal Society Interface*, 17(172):20200596, 2020. doi: https://doi.org/10.1098/rsif.2020.0596.

Ministério da Saúde. SRAG 2020 - Banco de Dados de Síndrome Respiratória Aguda Grave, 2020. URL https://opendatasus.saude.gov.br/dataset/bd-srag-2020.

H Juliette Unwin, Swapnil Mishra, Valerie C Bradley, A Gandy, M Vollmer, T Mellan, H Coupland, J Ish-Horowicz, et al. State-level tracking of COVID-19 in the United States. *Nature Communications*, (6189), 2020. doi: https://doi.org/10.1038/s41467-020-19652-6.

Robert Verity, Lucy C Okell, Ilaria Dorigatti, Peter Winskill, Charles Whittaker, Natsuko Imai, Gina Cuomo-Dannenburg, Hayley Thompson, Patrick G T Walker, Han Fu, Amy Dighe, Jamie T Griffin, Marc Baguelin, Sangeeta Bhatia, Adhiratha Boonyasiri, Anne Cori, Zulma Cucunubá, Rich FitzJohn, Katy Gaythorpe, Will Green, Arran Hamlet, Wes Hinsley, Daniel Laydon, Gemma Nedjati-Gilani, Steven Riley, Sabine van Elsland, Erik Volz, Haowei Wang, Yuanrong Wang, Xiaoyue Xi, Christl A Donnelly, Azra C Ghani, and Neil M Ferguson. Estimates of the severity of coronavirus disease 2019: a model-based analysis. *Lancet Infectious Diseases*, 2020a. doi: https://doi.org/10.1016/S1473-3099(20)30243-7.

Robert Verity, Lucy C. Okell, Ilaria Dorigatti, Peter Winskill, Charles Whittaker, Natsuko Imai, Gina Cuomo-dannenburg, Hayley Thompson, Patrick G. T. Walker, Han Fu, Amy Dighe, T. Jamie, Katy Gaythorpe, Will Green, Arran Hamlet, Wes Hinsley, Daniel Laydon, and Gemma Nedjati. Estimates of the severity of COVID-19 disease. *Lancet Infect Dis*, in press, 2020b. doi: https://doi.org/10.1101/2020.03.09.20033357.

Patrick GT Walker, Charles Whittaker, Oliver J Watson, Marc Baguelin, Peter Winskill, Arran Hamlet, Bimandra A Djafaara, Zulma Cucunubá, Daniela Olivera Mesa, Will Green, et al. The impact of COVID-19 and strategies for mitigation and suppression in low-and middle-income countries. *Science*, 2020. doi: DOI:10.1126/science.abc0035.

WHO. WHO Coronavirus (COVID-19) Dashboard. 2023. URL `https://covid19.who.int/`.

# A Appendix

## A.1 Appendix A

Figures and Tables directly referenced in the main text.

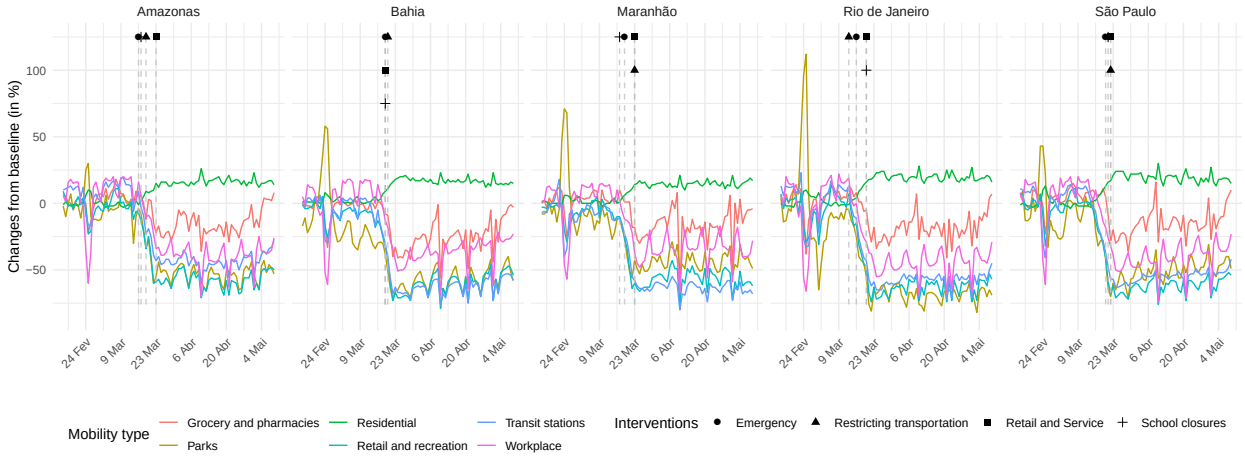


Figure 2: Mobility covariates from Google mobility reports for São Paulo (SP), Rio de Janeiro (RJ), Pernambuco (PE), Ceará (CE), Amazonas (AM). Data from `https://www.google.com/covid19/mobility/`

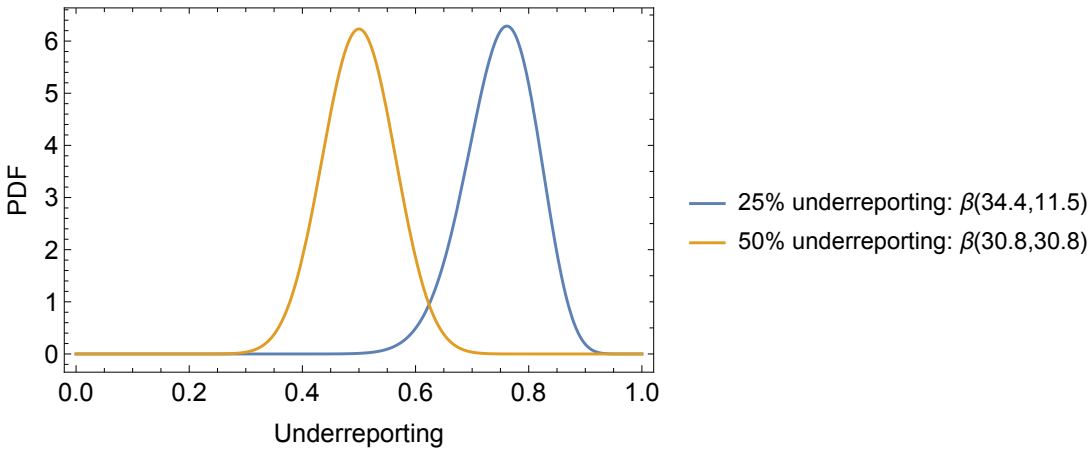


Figure 3: Prior distributions for death underreporting scenarios.

| State | IFR % | Population | Deaths | Deaths per million | Infections (thousands) | Attack rate % | $R_t$ |
|---|---|---|---|---|---|---|---|
| SP | 0.7 | 46,289,333 | 5,142 | 111 | 1,370 [950, 1,880] | 3.0 [2.1, 4.1] | 1.4 [1.1, 1.7] |
| RJ | 0.8 | 17,366,189 | 4,564 | 263 | 1,050 [777, 1,380] | 6.1 [4.5, 8.0] | 1.1 [0.9, 1.3] |
| PA | 0.9 | 8,690,745 | 1,752 | 202 | 529 [376, 724] | 6.1 [4.3, 8.3] | 1.3 [1.1, 1.5] |
| CE | 1.1 | 9,187,886 | 1,573 | 171 | 508 [381, 661] | 5.5 [4.1, 7.2] | 1.7 [1.5, 1.8] |
| AM | 0.8 | 4,207,714 | 1,446 | 344 | 337 [241, 454] | 8.0 [5.7, 10.8] | 1.2 [0.9, 1.4] |
| PE | 1.1 | 9,617,072 | 1,422 | 148 | 290 [203, 400] | 3.0 [2.1, 4.2] | 1.4 [1.1, 1.6] |
| MA | 1.0 | 7,114,598 | 745 | 105 | 121 [82, 172] | 1.7 [1.2, 2.4] | 0.8 [0.6, 1.0] |
| BA | 1.1 | 14,930,424 | 263 | 17 | 63 [39, 94] | 0.4 [0.3, 0.6] | 1.6 [1.3, 1.9] |
| ES | 0.9 | 4,064,052 | 202 | 50 | 57 [34, 89] | 1.4 [0.8, 2.2] | 1.5 [1.1, 1.8] |
| AL | 1.1 | 3,351,092 | 182 | 54 | 66 [38, 103] | 2.0 [1.1, 3.1] | 1.8 [1.4, 2.3] |
| PB | 1.2 | 4,039,277 | 161 | 40 | 59 [35, 92] | 1.5 [0.9, 2.3] | 2.1 [1.7, 2.5] |
| MG | 1.0 | 21,292,666 | 153 | 7 | 31 [15, 54] | 0.1 [0.1, 0.3] | 1.5 [1.0, 2.0] |
| PR | 0.9 | 11,516,840 | 113 | 10 | 21 [10, 37] | 0.2 [0.1, 0.3] | 1.2 [0.7, 1.7] |
| RS | 0.9 | 11,422,973 | 106 | 9 | 35 [17, 63] | 0.3 [0.2, 0.6] | 1.8 [1.3, 2.4] |
| RN | 1.1 | 3,534,165 | 90 | 26 | 21 [10, 38] | 0.6 [0.3, 1.1] | 1.6 [1.1, 2.1] |
| AP | 0.7 | 861,773 | 69 | 80 | 31 [15, 55] | 3.6 [1.7, 6.3] | 1.6 [1.2, 2.0] |
| SC | 0.8 | 7,252,502 | 68 | 9 | 17 [7, 33] | 0.2 [0.1, 0.5] | 1.4 [0.9, 2.0] |
| GO | 0.9 | 7,116,143 | 60 | 8 | 22 [9, 42] | 0.3 [0.1, 0.6] | 2.0 [1.4, 2.6] |

Table 2: Estimated infection fatality ratio (IFR), state population, reported deaths and deaths per million population, estimated number of infections in thousands, attack rate (AR), and time-varying reproduction number on 9th-May-2020 with 95% credible intervals, for São Paulo (SP), Rio de Janeiro (RJ), Pará (PA), Ceará (CE), Amazonas (AM), Pernambuco (PE), Maranhão (MA), Bahia (BA), Espírito Santo (ES), Alagoas (AL), Paraíba (PB), Minas Gerais (MG), Paraná (PR), Rio Grande do Sul (RS), Rio Grande do Norte (RN), Amapa (AP), Santa Catarina (SC), Goias (GO).

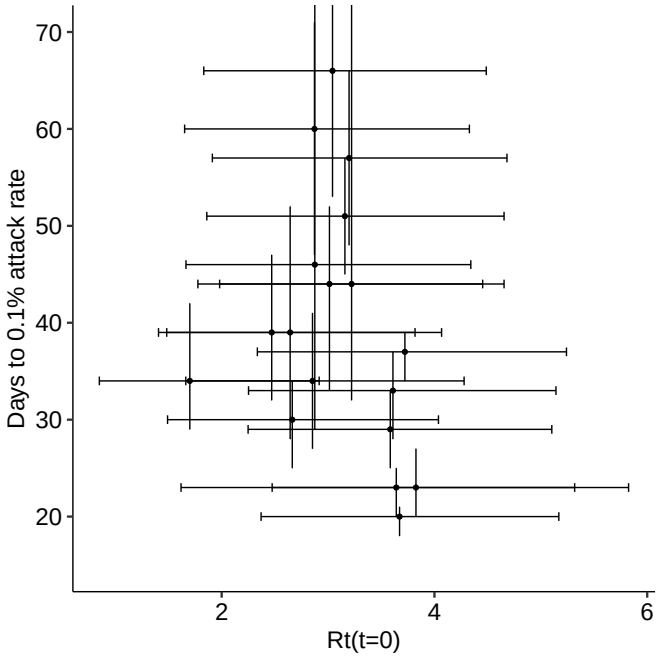


Figure 4: Time in days to 0.1% attack rate (measured from date of first case in Brazil) vs initial reproduction number $R_t(t = 0)$ for 18 states considered in the joint model: São Paulo (SP), Rio de Janeiro (RJ), Pará (PA), Ceará (CE), Amazonas (AM), Pernambuco (PE), Maranhão (MA), Bahia (BA), Espírito Santo (ES), Alagoas (AL), Paraíba (PB), Minas Gerais (MG), Paraná (PR), Rio Grande do Sul (RS), Rio Grande do Norte (RN), Amapa (AP), Santa Catarina (SC), Goias (GO).

### A.2 APPENDIX B

| State | 0% underreporting | 25% underreporting | 50% underreporting |
|---|---|---|---|
| SP | 2.95 [2.05, 4.07] | 4.09 [2.72, 5.81] | 6.07 [3.84, 9.43] |
| RJ | 6.06 [4.47, 7.96] | 8.10 [5.46, 11.80] | 12.30 [7.97, 17.70] |
| PA | 6.09 [4.33, 8.33] | 8.69 [5.25, 13.40] | 12.50 [7.91, 19.40] |
| CE | 5.53 [4.14, 7.20] | 7.33 [5.10, 10.70] | 10.80 [7.29, 15.80] |
| AM | 8.01 [5.73,10.80] | 10.80 [6.81, 15.40] | 16.50 [10.40, 25.40] |
| PE | 3.01 [2.12, 4.15] | 3.95 [2.54, 5.72] | 6.34 [3.82, 9.73] |
| MA | 1.70 [1.15, 2.41] | 2.35 [1.52, 3.56] | 3.62 [2.15, 5.59] |
| BA | 0.42 [0.26, 0.63] | 0.58 [0.33, 1.00] | 0.84 [0.49, 1.31] |
| ES | 1.41 [0.83, 2.18] | 1.85 [1.07, 3.02] | 2.74 [1.57, 4.54] |
| AL | 1.96 [1.13, 3.08] | 2.64 [1.51, 4.33] | 3.67 [2.06, 6.06] |
| PB | 1.46 [0.87, 2.29] | 2.09 [1.09, 3.67] | 2.87 [1.59, 4.71] |
| MG | 0.14 [0.07, 0.25] | 0.20 [0.10, 0.36] | 0.30 [0.15, 0.54] |
| PR | 0.18 [0.09, 0.32] | 0.23 [0.11, 0.41] | 0.35 [0.17, 0.63] |
| RS | 0.31 [0.15, 0.55] | 0.41 [0.19, 0.77] | 0.68 [0.30, 1.27] |
| RN | 0.61 [0.29, 1.09] | 0.82 [0.35, 1.64] | 1.18 [0.54, 2.14] |
| AP | 3.60 [1.74, 6.34] | 4.55 [2.32, 8.02] | 7.57 [3.57, 14.10] |
| SC | 0.23 [0.09, 0.45] | 0.34 [0.14, 0.68] | 0.41 [0.19, 0.79] |
| GO | 0.30 [0.12, 0.59] | 0.46 [0.19, 0.90] | 0.62 [0.26, 1.17] |

Table 3: Estimated attack rates for 0%, 25% and 50% death underreporting scenarios.

### A.3 CASES AND $R_t$ FOR 18 STATES

The estimated cases, deaths and $R_t$ for all 18 states considered in our joint model, São Paulo (SP), Rio de Janeiro (RJ), Ceará (CE), Amazonas (AM), Pernambuco (PE), Pará (PA), Maranhão (MA), Bahia (BA), Espírito Santo (ES), Alagoas (AL), Minas Gerais (MG), Paraíba (PB), Paraná (PR), Amapa (AP), Rio Grande do Sul (RS), Rio Grande do Norte (RN), Santa Catarina (SC), Goias (GO), are shown in Figure 5 considering mix of data sets. The same analysis taking into account each data set was also conducted, see Figures 7 and 6.

### A.4 ONSET-TO-DEATH SENSITIVITY

Onset-to-death sensitivity analysis is shown in Table 4. Attack rates are compared using an onset-to-death gamma distribution fitted specifically to Brazilian data,Ministério da Saúde (2020); Hawryluk et al. (2020) and an onset-to-death gamma distribution with a longer mean based on earlier estimates.Verity et al. (2020b) While estimates are quantitatively affected, changes are small in most states, and conclusions overall remained unaltered.

### A.5 IFR CALCULATION AND SENSITIVITY ANALYSIS

In order to derive an expected IFR across different states in Brazil, mixing patterns from Latin America Grijalva et al. (2017) and virus' transmissibility Flaxman et al. (2020) were used. Moreover, to account for the disease severity, data from Chinese epidemic Verity et al. (2020a) was modified to match data from the outbreak in UK Flaxman et al. (2020). Additionally we modified these estimates of disease severity (specifically the IFR) to account for the substantial heterogeneity we expect to observe in health outcomes across states due to variation in healthcare quality and capacity, the details of which are described below.

Across the states considered in this analysis, average income (in US dollars) varies from as high as $\sim$ \$300 in São Paulo to as low as $\sim$ \$100 in Maranhão.ibg (2019a) Such disparities in income are likely to result in significant differences in the quality and extent of available healthcare. Motivated by this, we modified the state-specific IFRs used in an income-dependent manner. Specifically, we assumed that the state with the highest income (São Paulo) has a quality of care identical to that observed in China (and thus motivated using the estimates presented in Verity et al.Verity et al. (2020a)), and that the state with the lowest income (Maranhão) had significantly worse healthcare outcomes -

| State | AR% $\Gamma(13.4, 0.7)$ onset-to-death | AR% $\Gamma(17.8, 0.45)$ onset-to-death |
|---|---|---|
| SP | 2.95 [2.05, 4.07] | 3.44 [2.06, 5.42] |
| RJ | 6.06 [4.47, 7.96] | 6.17 [4.09, 8.84] |
| PA | 6.09 [4.33, 8.33] | 6.16 [3.72, 9.60] |
| CE | 5.53 [4.14, 7.20] | 7.55 [4.95, 10.90] |
| AM | 8.01 [5.73, 10.80] | 8.10 [5.22, 12.00] |
| PE | 3.01 [2.12, 4.15] | 3.52 [2.10, 5.53] |
| MA | 1.70 [1.15, 2.41] | 1.63 [1.01, 2.47] |
| BA | 0.42 [0.26, 0.63] | 0.53 [0.27, 0.92] |
| ES | 1.41 [0.83, 2.18] | 1.69 [0.83, 3.00] |
| AL | 1.96 [1.13, 3.08] | 3.09 [1.33, 5.90] |
| PB | 1.46 [0.87, 2.29] | 2.71 [1.24, 5.11] |
| MG | 0.14 [0.07, 0.25] | 0.18 [0.07, 0.37] |
| PR | 0.18 [0.09, 0.32] | 0.19 [0.09, 0.37] |
| RS | 0.31 [0.15, 0.55] | 0.41 [0.15, 0.89] |
| RN | 0.61 [0.29, 1.09] | 0.81 [0.29, 1.78] |
| AP | 3.60 [1.74, 6.34] | 4.85 [1.88, 9.81] |
| SC | 0.23 [0.09, 0.45] | 0.29 [0.10, 0.67] |
| GO | 0.30 [0.12, 0.59] | 0.49 [0.14, 1.16] |

Table 4: Attack rate (AR) with onset-to-death distribution for Brazil specifically, $\Gamma(13.4, 0.7)$ fitted from SIVEP Gripe dataset, and $\Gamma(17.8, 0.45)$ from Verity et al. (2020b)

.

more similar to those that would be expected in a Lower Middle Income Country (see Walker et al. (2020) for further details on how differences in health quality across settings are likely to impact outcomes). For the other states where income lies somewhere between that of Maranhão and São Paulo, we linearly interpolate the age-specific infection fatality probabilities based on state-level average income.ibg (2019a) These age-specific infection fatality probabilities are then combined with predictions of the age-distribution of infections to produce an overall, state-specific IFR.

Substantial uncertainty still remains in these IFR calculations. Motivated by this we carried out a sensitivity analysis exploring the impacts of different choices of mixing matrix (Peru vs the United Kingdom) and of assumptions surrounding healthcare quality (namely the interpolation method described above or assuming that all states are able to provide a level of healthcare equal to that seen during the Chinese epidemic). The results of these sensitivity analyses are shown in Table 5 for different IFRs. Although assumptions surrounding healthcare quality impact the quantitative predictions of the IFR and associated predicted attack rates, they do not qualitatively change our conclusions surrounding herd immunity and the lack of infections sufficient to have reached it. https://www.overleaf.com/project/63c9469afbb8d34434eb1900

| State | (i) AR% UK contact matrix | (ii) AR% Peru contact matrix | (iii) AR% UK contact matrix, poorer outcomes | (iv) AR% Peru contact matrix, poorer outcomes |
|---|---|---|---|---|
| SP | 3.07 [2.14, 4.21] | 2.95 [2.06, 4.07] | 3.07 [2.14, 4.22] | 2.95 [2.05, 4.07] |
| RJ | 6.62 [4.89, 8.72] | 6.33 [4.68, 8.32] | 6.31 [4.68, 8.28] | 6.06 [4.47, 7.96] |
| PA | 12.60 [9.04, 17.10] | 12.30 [8.85, 16.70] | 6.11 [4.32, 8.32] | 6.09 [4.33, 8.33] |
| CE | 10.20 [7.63, 13.20] | 9.73 [7.34, 12.70] | 5.63 [4.19, 7.36] | 5.53 [4.14, 7.20] |
| AM | 16.70 [12.10, 22.30] | 16.50 [12.00, 22.10] | 7.97 [5.70, 10.80] | 8.01 [5.73, 10.80] |
| PE | 5.54 [3.92, 7.57] | 5.33 [3.74, 7.34] | 3.06 [2.16, 4.19] | 3.01 [2.12, 4.15] |
| MA | 3.70 [2.49, 5.23] | 3.57 [2.41, 5.04] | 1.73 [1.17, 2.44] | 1.70 [1.15, 2.41] |
| BA | 0.80 [0.49, 1.19] | 0.76 [0.48, 1.13] | 0.43 [0.27, 0.64] | 0.42 [0.26, 0.63] |
| ES | 2.00 [1.20, 3.09] | 1.92 [1.15, 2.97] | 1.44 [0.86, 2.24] | 1.41 [0.83, 2.18] |
| AL | 4.02 [2.37, 6.30] | 3.90 [2.28, 6.11] | 1.96 [1.15, 3.11] | 1.96 [1.13, 3.08] |
| PB | 2.68 [1.58, 4.17] | 2.55 [1.51, 3.97] | 1.49 [0.89, 2.33] | 1.46 [0.87, 2.29] |
| MG | 0.22 [0.11, 0.38] | 0.21 [0.10, 0.36] | 0.15 [0.07, 0.26] | 0.14 [0.07, 0.25] |
| PR | 0.24 [0.12, 0.42] | 0.23 [0.11, 0.40] | 0.19 [0.09, 0.33] | 0.18 [0.09, 0.32] |
| RS | 0.35 [0.17, 0.62] | 0.33 [0.16, 0.58] | 0.32 [0.16, 0.57] | 0.31 [0.15, 0.55] |
| RN | 1.08 [0.51, 1.94] | 1.03 [0.48, 1.87] | 0.63 [0.29, 1.12] | 0.61 [0.29, 1.09] |
| AP | 7.70 [3.82, 13.50] | 7.52 [3.70, 13.10] | 3.60 [1.75, 6.34] | 3.60 [1.74, 6.34] |
| SC | 0.28 [0.12, 0.53] | 0.27 [0.11, 0.51] | 0.24 [0.10, 0.46] | 0.23 [0.09, 0.45] |
| GO | 0.48 [0.20, 0.92] | 0.47 [0.20, 0.91] | 0.30 [0.12, 0.59] | 0.30 [0.12, 0.59] |

Table 5: Attack rates % (AR) estimated using different infection fatality ratios (IFR) with Brazilian state-level population weighting and using: i) UK contact matrix, ii) Peru contact matrix, iii) UK contact matrix with poorer hospitalisation outcomes, iv) Peru contact matrix with poorer hospitalisation outcomes.

| State | (i) IFR UK contact matrix | (ii) IFR Peru contact matrix | (iii) IFR UK contact matrix, poorer outcomes | (iv) IFR Peru contact matrix, poorer outcomes |
|---|---|---|---|---|
| AC | 0.38 | 0.39 | 0.78 | 0.78 |
| AL | 0.51 | 0.53 | 1.06 | 1.07 |
| AM | 0.37 | 0.38 | 0.79 | 0.79 |
| AP | 0.34 | 0.35 | 0.73 | 0.73 |
| BA | 0.59 | 0.62 | 1.10 | 1.12 |
| CE | 0.58 | 0.61 | 1.07 | 1.09 |
| ES | 0.63 | 0.65 | 0.87 | 0.89 |
| MA | 0.48 | 0.50 | 1.03 | 1.04 |
| MG | 0.69 | 0.72 | 1.01 | 1.04 |
| PA | 0.42 | 0.43 | 0.89 | 0.89 |
| PB | 0.62 | 0.65 | 1.13 | 1.16 |
| PE | 0.58 | 0.60 | 1.06 | 1.07 |
| PI | 0.57 | 0.59 | 1.10 | 1.11 |
| PR | 0.66 | 0.69 | 0.84 | 0.86 |
| RJ | 0.73 | 0.76 | 0.76 | 0.79 |
| RN | 0.60 | 0.62 | 1.04 | 1.06 |
| RO | 0.45 | 0.45 | 0.81 | 0.81 |
| RR | 0.33 | 0.34 | 0.67 | 0.67 |
| RS | 0.78 | 0.81 | 0.84 | 0.87 |
| SC | 0.65 | 0.67 | 0.74 | 0.76 |
| SE | 0.51 | 0.53 | 0.96 | 0.97 |
| SP | 0.67 | 0.70 | 0.67 | 0.70 |
| TO | 0.49 | 0.51 | 0.89 | 0.90 |

Table 6: Infection fatality ratios (IFR) with Brazilian state-level population weighting,, using: i) UK contact matrix, ii) Peru contact matrix, iii) UK contact matrix with poorer hospitalisation outcomes, iv) Peru contact matrix with poorer hospitalisation outcomes.

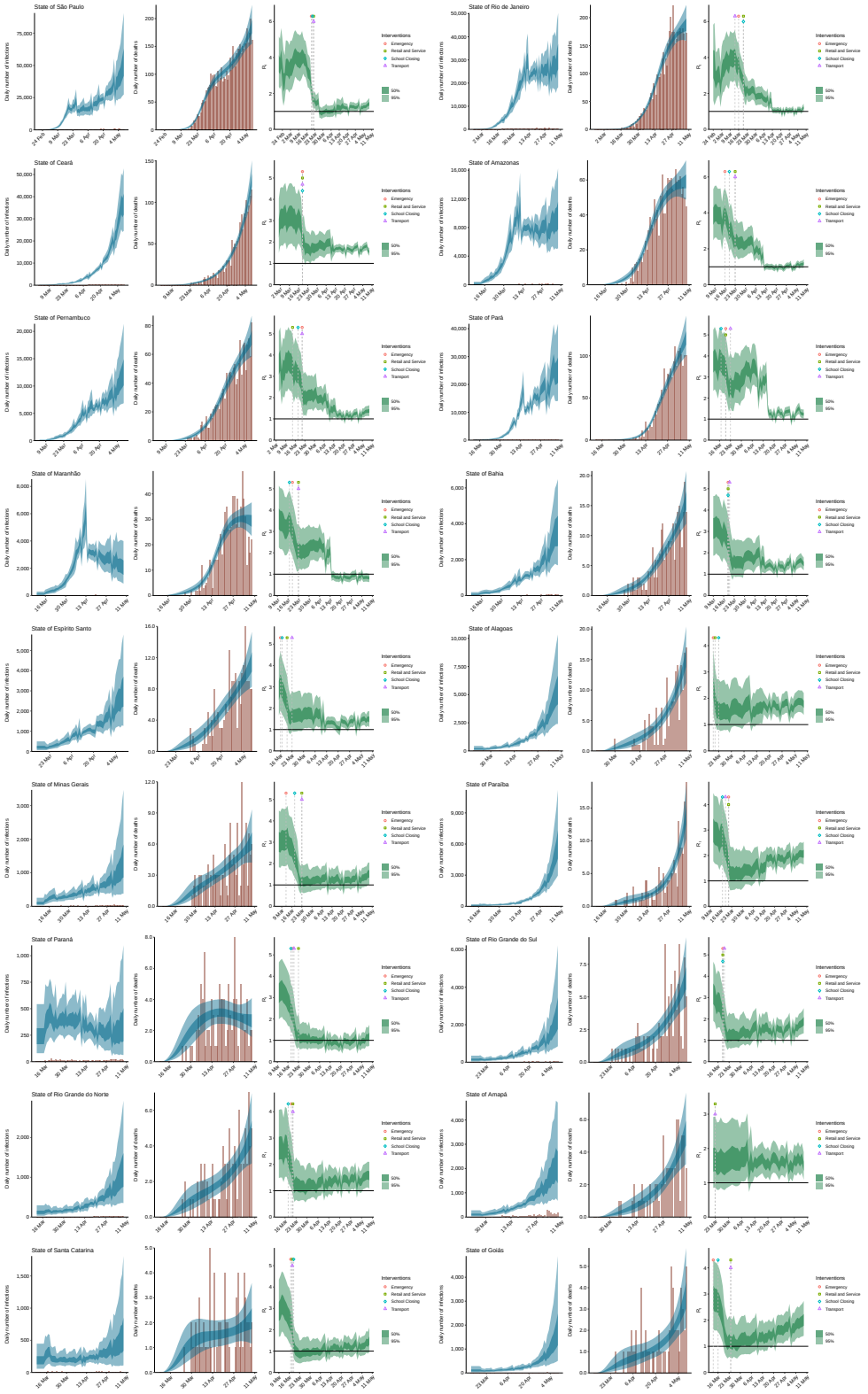


Figure 5: Estimates of infections, deaths and $R_t$ for all 18 states considered in the model considering the mix of both data sets (Painel Coronavírus cov (2020) and SIVEP GripeMinistério da Saúde (2020)).

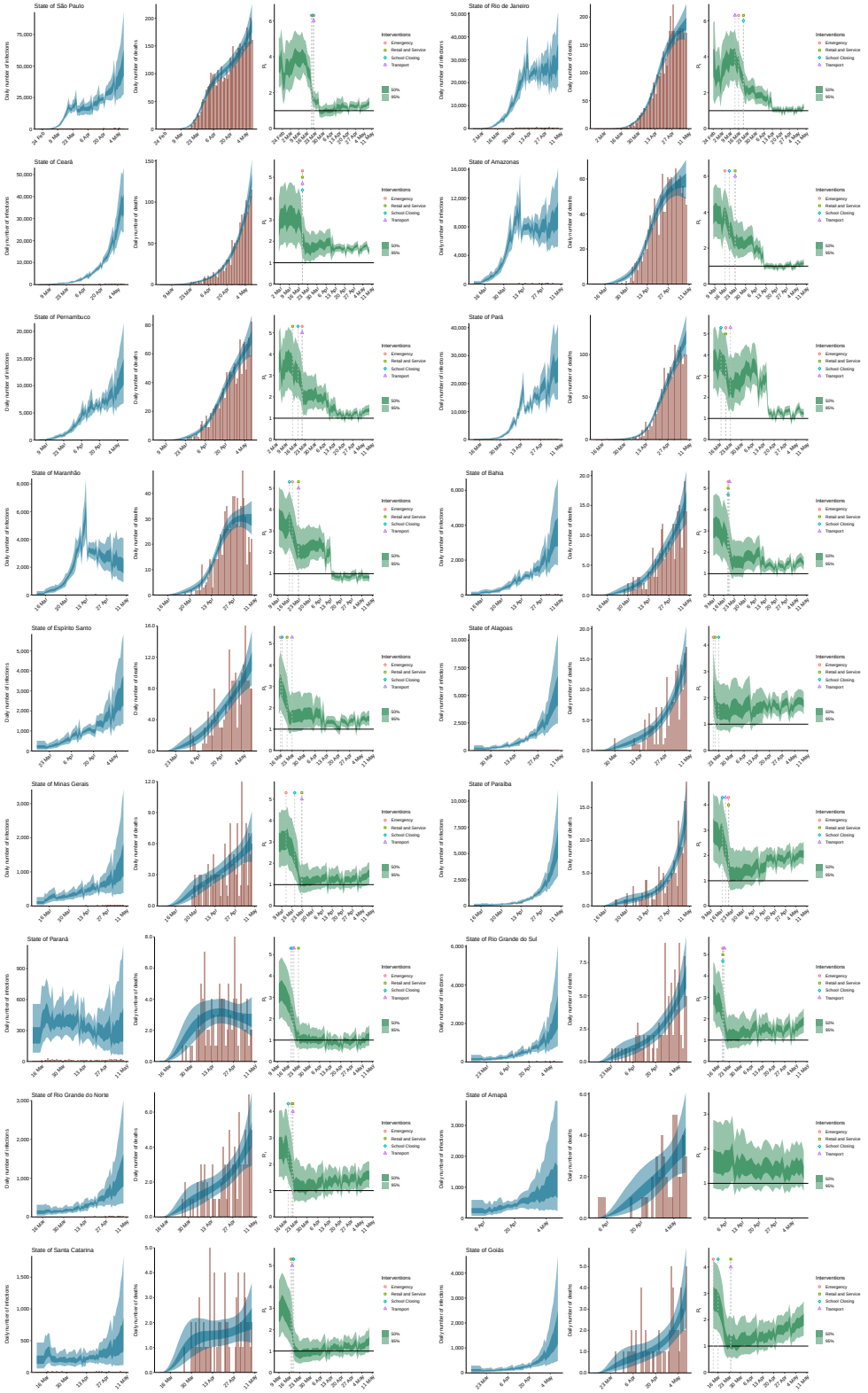


Figure 6: Estimates of infections, deaths and $R_t$ for all 18 states considered in the joint model based on SIVEP-Gripe Ministério da Saúde (2020) data.

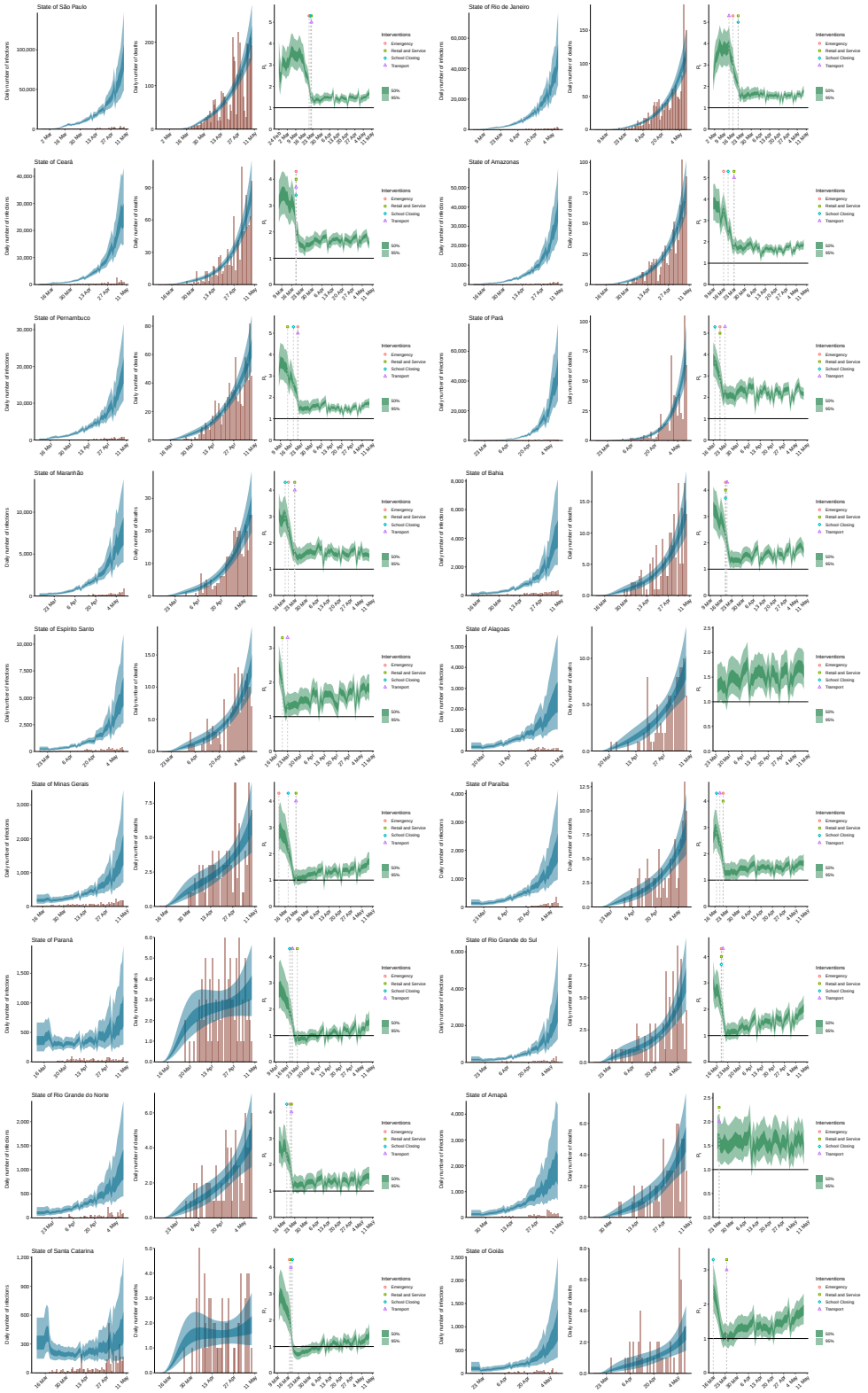


Figure 7: Estimates of infections, deaths and $R_t$ for a joint model of 18 states based on Painel Coronavírus cov (2020) data.

### A.6 DEATH DATA SOURCE INTERPRETATION AND AUXILIARY MODEL DATA

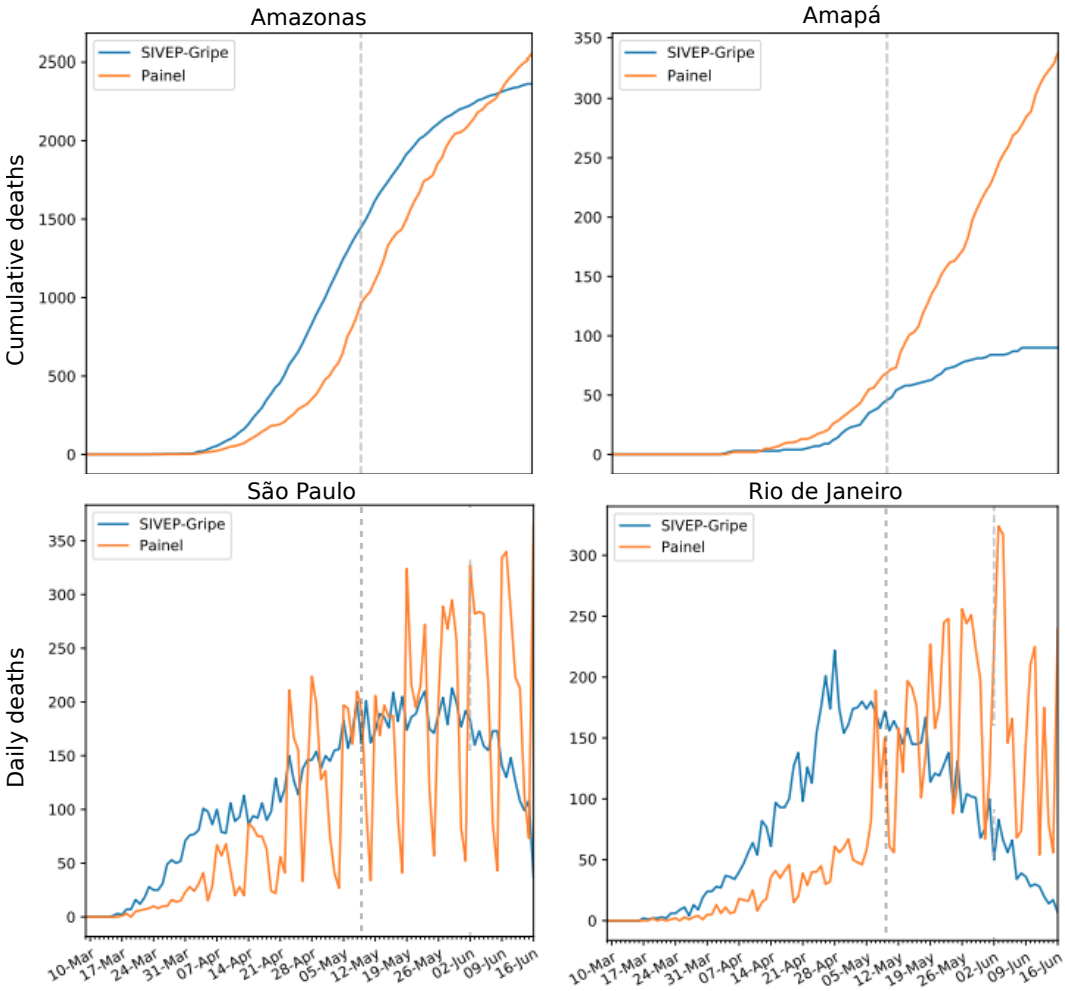


Figure 8: Comparison of the daily death records from Brazil's Ministry of Health *Painel Coronavírus* (date of notification - orange line)cov (2020) and SIVEP-Gripe (date of death - blue line)Ministério da Saúde (2020) datasets, from 10 March 2020 to 16 June 2020. Gray dashed vertical line represents 09 May, the last day in our analysis. Top row displays cumulative deaths plots. Top left, state of Amazonas (population 4.2M), shows a lag of approximately 10 days between the cumulative counts of deaths by date of death vs. date of notification. Top right, state of Amapá (population 861,000), has the cumulative count of deaths by date of notification higher than by date of death, indicating deaths are being reported to the Ministry of Health but not updated onto the SIVEP-Gripe database. Bottom row are daily death plots. Bottom left, state of São Paulo (population 46M), shows the effects of processing and testing delays on the official daily counts, and also the uncertain effect of right-censoring. Bottom right, state of Rio de Janeiro (population 17M), shows testing and processing lags. Important to caution that there is still substantial, but uncertain, right-censoring in the SIVEP-Gripe data for the state of Rio de Janeiro.

As deaths are only reported if SARS-CoV-2 is confirmed by a positive result on diagnostic testing, there can be considerable and variable lag between the date of death and the official reporting of COVID-19 deaths. Current Ministry of Health protocol guides that all hospitalizations and deaths due to suspected or confirmed COVID-19 must also be notified on an online system called SIVEP-Gripe.Ministério da Saúde (2020) After testing and investigation by health officials, cases and deaths receive a final COVID-19 classification and records are updated to include lab results

and the exact date of death or hospital discharge for each patient. Thus, SIVEP-Gripe allows for the epidemic to be tracked by the actual daily number of COVID-19 deaths. For instance, the official *Painel Coronavírus* daily death count on the 9th of May 2020 was 732, but there are at least 799 COVID19-confirmed deaths already registered on SIVEP-Gripe. In total, SIVEP-Gripe records 18,314 confirmed COVID-19 deaths, 72.3% more than the official 10,627 count for the same date. However, as not all states are following the official protocol of recording their patients on SIVEP-Gripe in a timely manner (see Figure 8), the true number of COVID-19 deaths on any given day can be significantly larger.

For population counts we used the 2020 projection by state published by *Instituto Brasileiro de Geografia e Estatística*.ibg (2019b)

Regarding intervention data, the values taken into account are the dates in which interventions were effectively applied, even though they were encouraged at earlier dates.

| State | Emergency declared | Retail and services closed | Transportation restricted | Schools closed |
|---|---|---|---|---|
| AC | 2020-03-20 | 2020-03-20 | 2020-03-20 | 2020-03-20 |
| AL | 2020-03-20 | 2020-03-21 | 2020-03-19 | 2020-03-23 |
| AM | 2020-03-16 | 2020-03-23 | 2020-03-23 | 2020-03-19 |
| AP | 2020-03-20 | 2020-03-23 | 2020-03-23 | 2020-03-17 |
| BA | 2020-03-19 | 2020-03-19 | 2020-03-20 | 2020-03-19 |
| CE | 2020-03-19 | 2020-03-19 | 2020-03-19 | 2020-03-19 |
| DF | 2020-02-28 | 2020-03-23 | 2020-03-18 | 2020-03-11 |
| ES | 2020-03-16 | 2020-03-20 | 2020-03-23 | 2020-03-17 |
| GO | 2020-03-13 | 2020-03-24 | 2020-03-24 | 2020-03-16 |
| MA | 2020-03-19 | 2020-03-23 | 2020-03-23 | 2020-03-17 |
| MG | 2020-03-12 | 2020-03-23 | 2020-03-23 | 2020-03-18 |
| MS | 2020-03-19 | 2020-03-19 | 2020-03-25 | 2020-03-24 |
| MT | 2020-03-23 | 2020-03-23 | 2020-03-18 | 2020-03-23 |
| PA | 2020-03-20 | 2020-03-20 | 2020-03-23 | 2020-03-17 |
| PB | 2020-03-21 | 2020-03-21 | 2020-03-19 | 2020-03-17 |
| PE | 2020-03-21 | 2020-03-14 | 2020-03-21 | 2020-03-18 |
| PI | 2020-03-19 | 2020-03-23 | 2020-03-23 | 2020-03-16 |
| PR | 2020-03-19 | 2020-03-23 | 2020-03-20 | 2020-03-18 |
| RJ | 2020-03-16 | 2020-03-20 | 2020-03-13 | 2020-03-20 |
| RN | 2020-03-20 | 2020-03-21 | 2020-03-21 | 2020-03-18 |
| RO | 2020-03-20 | 2020-03-21 | | 2020-03-17 |
| RR | 2020-03-23 | 2020-03-23 | 2020-02-20 | 2020-03-20 |
| RS | 2020-03-19 | 2020-03-19 | 2020-03-20 | 2020-03-19 |
| SC | 2020-03-17 | 2020-03-18 | 2020-03-18 | 2020-03-19 |
| SE | 2020-03-16 | 2020-03-20 | 2020-03-20 | 2020-03-16 |
| SP | 2020-03-20 | 2020-03-22 | 2020-03-22 | 2020-03-21 |
| TO | 2020-03-21 | 2020-03-21 | 2020-03-21 | 2020-03-16 |

Table 7: Non-pharmaceutical interventions by state, adapted from cep (2020).