Report 21: Estimating COVID-19 cases and reproduction number in Brazil

Thomas A Mellan[3*], Henrique H Hoeltgebaum[*], Swapnil Mishra[*], Charlie Whittaker[*], Ricardo P Schnekenberg[*], Axel Gandy[*], H Juliette T Unwin, Michaela A C Vollmer, Helen Coupland, Iwona Hawryluk, Nuno Rodrigues Faria, Juan Vesga, Harrison Zhu, Michael Hutchinson, Oliver Ratmann, Melodie Monod, Kylie Ainslie, Marc Baguelin, Sangeeta Bhatia, Adhiratha Boonyasiri, Nicholas Brazeau, Giovanni Charles, Laura V Cooper, Zulma Cucunuba, Gina Cuomo-Dannenburg, Amy Dighe, Bimandra Djaafara, Jeff Eaton, Sabine L van Elsland, Richard FitzJohn, Keith Fraser, Katy Gaythorpe, Will Green, Sarah Hayes, Natsuko Imai, Ben Jeffrey, Edward Knock, Daniel Laydon, John Lees, Tara Mangal, Andria Mousa, Gemma Nedjati-Gilani, Pierre Nouvellet, Daniela Olivera, Kris V Parag, Michael Pickles, Hayley A Thompson, Robert Verity, Caroline Walters, Haowei Wang, Yuanrong Wang, Oliver J Watson, Lilith Whittles, Xiaoyue Xi, Lucy Okell, Ilaria Dorigatti, Patrick Walker, Azra Ghani, Steven Riley, Neil M Ferguson[1], Christl A. Donnelly, Seth Flaxman[*] and Samir Bhatt[2*]

Department of Infectious Disease Epidemiology, Imperial College London

Department of Mathematics, Imperial College London

WHO Collaborating Centre for Infectious Disease Modelling

MRC Centre for Global Infectious Disease Analytics

Abdul Latif Jameel Institute for Disease and Emergency Analytics, Imperial College London

Department of Statistics, University of Oxford

Nuffield Department of Clinical Neurosciences, University of Oxford

[*]Contributed equally. Correspondence: [1]neil.ferguson@imperial.ac.uk, [2]s.bhatt@imperial.ac.uk,[3] t.mellan@imperial.ac.uk

SUGGESTED CITATION
Thomas A Mellan, Henrique H Hoeltgebaum, Swapnil Mishra *et al*. Estimating COVID-19 cases and reproduction number in Brazil. Imperial College London (08-05-2020), doi: https://doi.org/10.25561/78872.

1 Summary

Brazil is an epicentre for COVID-19 in Latin America. In this report we describe the Brazilian epidemic using three epidemiological measures: the number of infections, the number of deaths and the reproduction number. Our modelling framework requires sufficient death data to estimate trends, and we therefore limit our analysis to 16 states that have experienced a total of more than fifty deaths. The distribution of deaths among states is highly heterogeneous, with 5 states—São Paulo, Rio de Janeiro, Ceará, Pernambuco and Amazonas—accounting for 81% of deaths reported to date. In these states, we estimate that the percentage of people that have been infected with SARS-CoV-2 ranges from 3.3% (95% CI: 2.8%-3.7%) in São Paulo to 10.6% (95% CI: 8.8%-12.1%) in Amazonas. The reproduction number (a measure of transmission intensity) at the start of the epidemic meant that an infected individual would infect three or four others on average. Following non-pharmaceutical interventions such as school closures and decreases in population mobility, we show that the reproduction number has dropped substantially in each state. However, for all 16 states we study, we estimate with high confidence that the reproduction number remains above 1. A reproduction number above 1 means that the epidemic is not yet controlled and will continue to grow. These trends are in stark contrast to other major COVID-19 epidemics in Europe and Asia where enforced lockdowns have successfully driven the reproduction number below 1. While the Brazilian epidemic is still relatively nascent on a national scale, our results suggest that further action is needed to limit spread and prevent health system overload.

2 Introduction

The world faces an unprecedented public health emergency in the COVID-19 pandemic. Since the emergence of the novel coronavirus (SARS-CoV-2) in China in December 2019, global spread has been rapid, with over 3.5 million cases and almost 250 thousand deaths reported from 187 countries as of the 6th May [13]. Though transmission of the disease beyond Asia was initially centred around Western Europe and North America, significant spread is now seen in other parts of the world, including many countries across Sub-Saharan Africa and Latin America.

One such area of concern is Brazil - since report of its first case on 25th February, its epidemic has grown quickly, with the country now reporting over 135,000 cases and over 7,000 deaths [14].In response to significant spread and community transmission of the virus within the country, Brazilian state and city officials have mandated 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,

DOI: https://doi.org/10.25561/78872

restricting transportation, and closing schools. Specific packages of interventions have been decided at the state level, with substantial variation between states in the extent to which measures have been adopted, and their comparative timing [4]. Importantly, interventions employed to date remain short of the widespread and mandatory lockdowns implemented across parts of Asia and Europe and which have proved to be highly effective at containing spread of the virus [7, 19].

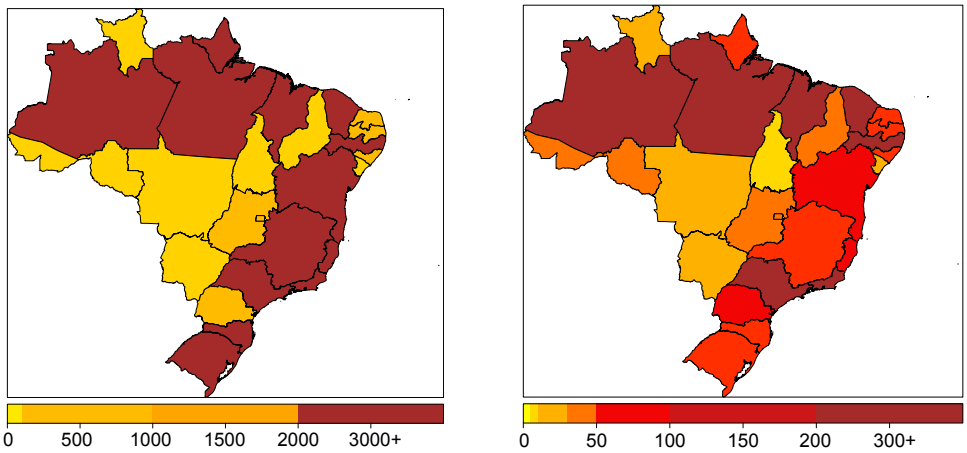

Figure 1: Map of cases (a) and deaths (b) in Brazil by state level. Data source: Painel Coronavirus at `https://covid.saude.gov.br`, accessed on 7th May 2020.

It remains unclear the extent to which these measures have been effective in reducing transmission across the country. Reported cases across Brazil have more than doubled over the past 10 days and show little sign of slowing. Given this rapid growth, a better understanding of the current epidemiological situation and the impact of interventions deployed to date is required to guide policy decisions aimed at preventing worsening of the public health emergency the country faces. Motivated by this, we extend a previously published semi-mechanistic Bayesian hierarchical model of COVID-19 epidemiological dynamics [7, 19] 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 detail at the state level and understand the spread of the virus across the country to date. These results include both estimates of the proportion of individuals infected so far as well as the impact of different control interventions, and provide insight into possible future epidemic trajectories should further control measures be employed, or current interventions relaxed.

DOI: https://doi.org/10.25561/78872

State	IFR %	Population	Deaths	Deaths per million	Infections (thousands)	Attack rate %	Reproduction number
SP	0.7	46,289,333	3,045	65.80	1,530 [1,310, 1,700]	3.30 [2.83, 3.68]	1.47 [1.34, 1.59]
RJ	0.8	17,366,189	1,205	69.40	582 [492, 657]	3.35 [2.83, 3.78]	1.44 [1.28, 1.60]
CE	1.1	9,187,886	848	92.30	410 [343, 464]	4.46 [3.73, 5.05]	1.61 [1.42, 1.81]
PE	1.1	9,617,072	803	83.50	288 [239, 328]	3.00 [2.49, 3.41]	1.32 [1.14, 1.53]
AM	0.8	4,207,714	751	178	448 [372, 509]	10.60 [8.84, 12.10]	1.58 [1.38, 1.81]
PA	0.9	8,690,745	392	45.10	439 [339, 513]	5.05 [3.90, 5.90]	1.90 [1.57, 2.31]
MA	1.0	7,114,598	291	40.90	147 [118, 170]	2.07 [1.65, 2.40]	1.55 [1.32, 1.80]
BA	1.1	14,930,424	160	10.70	59.5 [46.5, 69.6]	0.40 [0.31, 0.47]	1.37 [1.14, 1.63]
ES	0.9	4,064,052	145	35.70	91 [69.4, 107]	2.24 [1.71, 2.64]	1.57 [1.29, 1.90]
PR	0.9	11,516,840	101	8.77	28.4 [21.6, 33.6]	0.25 [0.19, 0.29]	1.16 [0.95, 1.39]
MG	1.0	21,292,666	97	4.56	28.1 [21, 33.4]	0.13 [0.10, 0.16]	1.30 [1.05, 1.57]
PB	1.2	4,039,277	92	22.80	25.7 [19.4, 30.4]	0.64 [0.48, 0.75]	1.23 [0.97, 1.52]
AL	1.1	3,351,092	89	26.60	40.1 [29, 48.1]	1.20 [0.87, 1.44]	1.27 [0.94, 1.67]
RS	0.9	11,422,973	87	7.62	48.2 [36.3, 57.1]	0.42 [0.32, 0.50]	1.44 [1.15, 1.77]
RN	1.1	3,534,165	72	20.40	19.9 [14.7, 23.7]	0.56 [0.42, 0.67]	1.18 [0.92, 1.45]
SC	0.8	7,252,502	59	8.14	16.5 [12.2, 19.7]	0.23 [0.17, 0.27]	1.14 [0.91, 1.38]

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 6 May 2020 with 95% credible intervals, for São Paulo (SP), Rio de Janeiro (RJ), Pernambuco (PE), Ceará (CE), Amazonas (AM), Pará (PA), Maranhão (MA), Bahia (BA), Espírito Santo (ES), Paraná (PR), Minas Gerais (MG), Paraíba (PB), Rio Grande do Sul (RS), Rio Grande do Norte (RN), Alagoas (AL), Santa Catarina (SC).

DOI: https://doi.org/10.25561/78872

3 Results

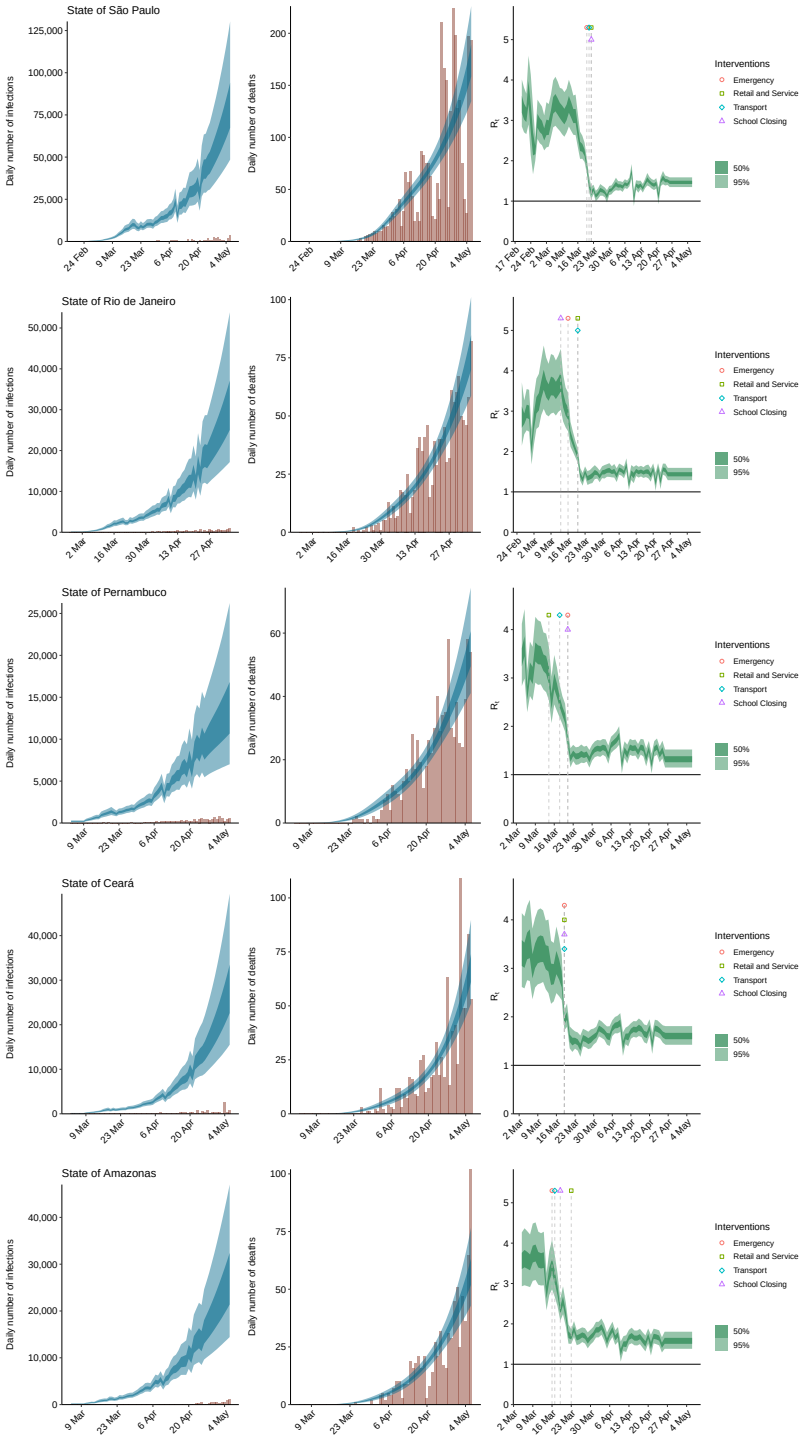

Figure 2: Estimates of infections, deaths and R_t. Left: daily number of infections, brown bars are reported cases, blue bands are predicted infections, dark blue 50% credible interval (CI), light blue 95% CI. Middle: daily number of deaths, brown bars are reported deaths, blue bands are predicted deaths,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

3.1 Attack Rates

Brazil has already reported almost twice as many COVID-19 deaths than China and over 100,000 confirmed cases. But despite these high numbers, we estimate that only a small proportion of individuals within each state has been infected to date (attack rate). The distribution of deaths among states is highly heterogeneous, with 5 states—São Paulo, Rio de Janeiro, Ceará, Pernambuco and Amazonas—accounting for 81% of reported deaths. In these states, we estimate that the percentage of people that have been infected with SARS-CoV-2 ranges from 3.3% (95% CI: 2.8%-3.7%) in São Paulo to 10.6% (95% CI: 8.8%-12.1%) in Amazonas. The remaining states all have attack rates below 2.3%, apart from Pará (PA), which at 5.05% (CI95: 3.9%-5.9%), is lower only than Amazonas. The full list of attack rate estimates for each state is shown in Table 1.

These results illustrate that the proportion of the population already infected and potentially immune remains low. Considering an R_0 of 3 and transmissibility similar to that observed across European[7] and Brazilian[16] settings, the estimated share of the population infected to date remains far short of the 70% herd immunity threshold required to prevent rapid resurgence of the virus if control measures are relaxed. [7] and in Brazil.[16]).

3.2 Sensitivity Analysis - Infection Fatality Rate & Underreporting

Substantial uncertainty remains in our understanding of the fundamental epidemiology of COVID-19 and the quality of surveillance systems across different settings. Consequently, estimates of the expected number of deaths and infections are sensitive to assumptions made within our modelling framework. In particular there is uncertainty surrounding the infection fatality ratio (IFR), which is the probability of an individual dying if infected with SARS-CoV-2. If this number is very low, more infections and a higher attack rate are to be expected for the same number of observed deaths, and vice versa. There is also considerable uncertainty in the observed death data, as little is known about the extent and nature of underreporting. In order to examine the effect of these assumptions on the conclusions described above, we undertook a series of sensitivity analyses (see Appendix) exploring different assumptions surrounding state-level IFR (relating to assumptions about how healthcare quality varies with state income) and the extent of death underreporting. The results of these sensitivity analysis, as expected, yield quantitative differences in the predicted attack rates - for example, assuming a 50% level of death underreporting changes our predicted attack rates from 3.30% (CI95: 2.83%-3.68%) and 10.60% (CI95: 8.84%-12.10%) to 6.49% (CI95: 5.44%-7.35%) and 19.90% (CI95:16.40%-22.80%) for São Paulo and Amazonas respectively. Similarly, assumptions surrounding the extent and variation of healthcare quality across states

DOI: https://doi.org/10.25561/78872

did not qualitatively alter the conclusions reached, namely that levels of infection in the population to date are significantly lower than that required for herd immunity.

3.3 Effectiveness of Control Measures

Attempts to contain the spread of SARS-Cov-2 in the community have centred around the deployment of various non-pharmaceutical interventions (NPIs) that involve reducing the number of contacts made between individuals [6]. Common examples are school closures, social distancing rules, banning of public gatherings and complete lockdown. Disrupting chains of transmission and bringing the reproduction number (R_t) below 1 is essential to control the virus and prevent exponential growth. Using our framework, we estimate the time-varying effective reproduction number across all Brazilian states. We parameterised R_t as a function of Google mobility data [2] - an implicit assumption within this framework is that changes in mobility patterns can be related to changes in transmission intensity, which is supported by previous research examining other respiratory viruses [18, 10].

We describe the time-varying effective reproduction number R_t for the 5 states with the current highest number of deaths: São Paulo (SP), Rio de Janeiro (RJ), Pernambuco (PE), Ceará (CE), and Amazonas (AM) (Figure 2). Estimates of the initial reproduction number (R_0) are consistently in the range of 3 - 4 across all states, in line with estimates of transmissibility derived from European data [7]. Our results also show that R_t has dropped dramatically following the implementation of public health interventions, with mobility declining by 29% on average across Brazil and R_t declining by 54% on average. However, in none of the states we considered did our results suggest that measures implemented to date have brought R_t below 1. By contrast, in previously published work examining Italy [19] where stringent measures including societal lockdowns have been implemented, mobility had reduced by 53% compared to baseline[2], reducing R_t by 85% compared to R_0 and bringing it significantly below 1. These results therefore suggest that, in the absence of additional major interventions, substantial further growth of the epidemic is expected across all 16 Brazilian states considered, leading to worsening of the COVID-19 public health crisis.

Changes in the effective reproduction number reflect alterations to patterns of mobility and contact stemming from both government-mandated interventions as well as changes in behaviour at the individual level. Within our framework, we explore the comparative impact of reductions in different types of mobility (in different settings, including the workplace, parks, residential and transit stations) on the effective reproduction number (see Figure 3). Our results support broad equivalency in the effect of reductions in different types of mobility and their corresponding impact on R_t (Figure 3). However,

DOI: https://doi.org/10.25561/78872

we note the large credible intervals associated with each estimate, which limits our ability to identify differences.

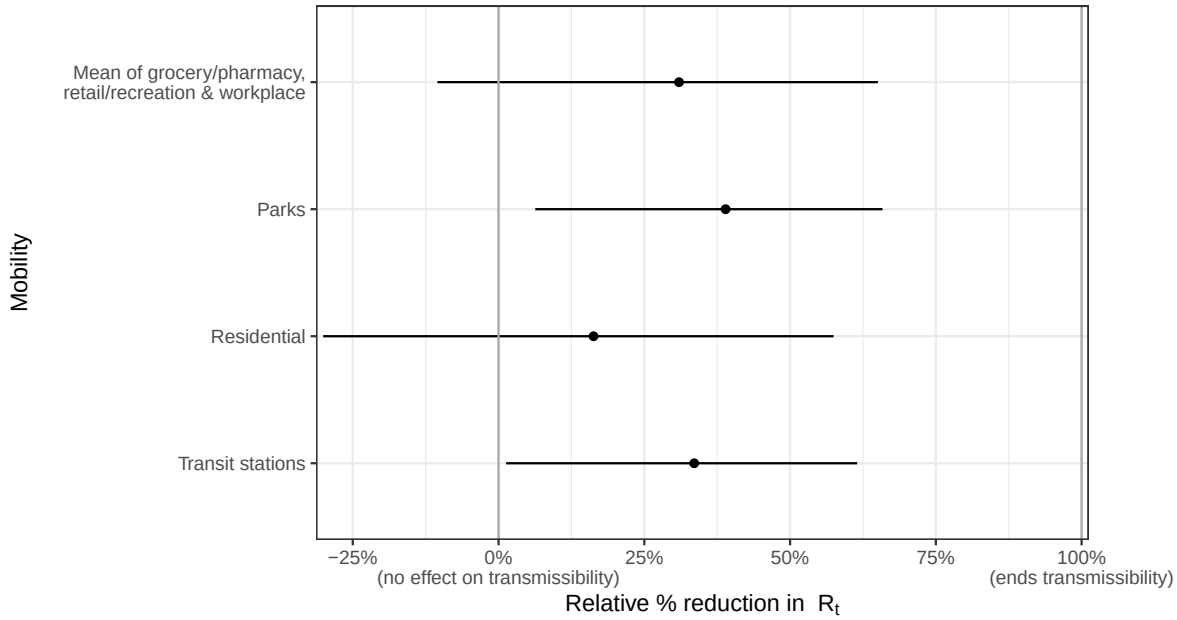

Figure 3: Effect sizes for mobility covariates in R_t.

4 Conclusion

In this report we utilise a semi-mechanistic Bayesian model of COVID-19 transmission, calibrated using data on reported deaths at the state level, to infer the epidemiological characteristics of the epidemic in Brazil to date. The results presented here suggest an ongoing epidemic in which substantial reductions in the average reproduction number have been achieved through non-pharmaceutical interventions. However our results also show that so far the changes in mobility have not been stringent enough to reduce the reproduction number below 1. Therefore we predict continued growth of the epidemic across Brazil and increases in the associated number of cases and deaths unless further actions are taken.

Our results reveal extensive heterogeneity in predicted attack rates between states, suggesting that the epidemic is at a far more advanced stage in some states compared to others. Despite this heterogeneity however, in no states do any of our results indicate that herd immunity is close to being reached, underscoring the early stage of the epidemic in Brazil currently, and the prospect of the situation worsening unless further control measures are implemented. The estimated attack rates are calculated from the reported number of deaths. We expect ascertainment of deaths to be higher than for cases, a phenomenon attributable to the large proportion of infected individuals who typically present as asymp-

tomatic e.g. [11]. Because of this extensive asymptomatic fraction and limitations surrounding population testing currently, we expect our estimates to be more accurate than naive attack rates estimated from the reported number of cases. Important to note however that estimating the extent to which deaths are underreported remains very difficult and we therefore consider a number of different death underreporting scenarios in our analyses. Although differences in these assumptions alter our quantitative estimates of attack rate, they do not alter our qualitative conclusions surrounding the impact of control interventions and the proportion of the population infected to date falling short of the threshold required for herd immunity. We additionally consider the possibility that the extent of COVID-19 death underreporting has been non-stationary over time - specifically, that increases in the ascertainment of COVID-19 deaths over time during the initial phase of the epidemic could lead to upwards biasing of our R_0 estimates. To test the sensitivity of this bias in R_0 we ran our model with a strong prior on a low R_0 but found that our conclusions regarding $R_t > 1$ were unchanged.

Our results surrounding attack rates are also sensitive to assumptions about the state-specific infection fatality ratio (IFR) used, a quantity driven by a diversity of different factors including demographic structure of the state's population, the pattern of social contacts between age-groups, and the quality/quantity of available healthcare (such as supportive oxygen therapy and mechanical ventilation). Our estimated IFRs for each of the Brazilian states range from 0.7% to 1.2%, reflecting substantial differences between states in their demographic structure and healthcare provision. As an example, the population of Amazonas is on average 7 years younger than São Paulo. Using previously published estimates of mortality risk [17], this difference would lead to naive estimates of the IFR that are lower in Amazonas compared to São Paulo (0.38% and 0.70% respectively. However, these previous studies assume a level of healthcare similar to that of China, accounting for the poorer health outcomes we expect in Brazil's least affluent states produces IFR estimates of 0.70% for São Paulo and 0.72% Amazonas. Whilst estimates of these IFRs will ultimately be sensitive to considerations of how healthcare quality shapes patterns of mortality, the presented results are robust to assumptions surrounding the extent of variation in healthcare quality (see Appendix). We would also note that across all states considered and across all scenarios, the IFR we predict is substantially higher than the value of 0.1% used in the recent work by Galluci-Neto and colleagues [5].

The semi-mechanistic Bayesian framework utilised also allows quantification of the impact of non-pharmaceutical interventions on the reproduction number, R_t. Our results suggest substantial reductions in the estimated value of R_t across all states following introduction of these interventions. This is predicated on the assumption that reductions in transmission can be accurately ascertained from reductions in mobility. Such an assumption has been borne out in previous work looking at the impact of reductions in

mobility on transmission of the virus during the Chinese epidemic [1]. However, it is important to note that the relationship linking mobility and transmission will likely be highly dynamic over the course of an epidemic, being modified by other factors such as behavioural changes surrounding physical distancing and routine mask wearing. .

Our results suggests that, despite reductions, these measures have only been partially successful in reducing transmission - across all states considered here, the value of R_t remains above 1, indicating that transmission remains uncontrolled and that in the absence of stricter measures leading to further reductions in mobility, that growth of the epidemic will continue. This is in contrast to European settings where societal lockdowns have been implemented, leading to estimated reductions in R_t below 1 [19]. Such reductions have been associated with substantial reductions in patterns of mobility - in countries such as Italy where a strict lockdown was mandated, mobility declined to far greater extent than has been observed to date in Brazil. For instance, patterns of mobility surrounding grocery/pharmacy in Lombardy, one of the most severely affected regions in Italy, dropped by almost 75% over when measures were introduced (yielding an estimated R_t of 0.58). By contrast, across Amazonas and São Paulo to date, the maximum observed reduction has only been 18% and 21% respectively, producing R_t estimates of 1.58 and 1.46 respectively. Overall whilst our work suggests that implemented measures to date have had an impact on transmission, they also highlight their insufficiency if transmission is to be controlled, and the need for further contact-limiting measures, beyond what is currently implemented, to reduce the reproduction number in Brazil to less than 1.

Overall, our results reveal that despite extensive spread and transmission of the virus across the country, the extent of infection in the general population remains low and far short of the level required for herd immunity. This result is robust to assumptions surrounding the IFR associated with each state, and the extent of underreporting we assume in the available deaths data. More broadly, our results suggest that in the absence of the introduction of further control measures that will more strongly curb transmission, Brazil faces the prospect of an epidemic that will continue to grow exponentially.

5 Acknowledgements

This work used the Cirrus UK National Tier-2 HPC Service at EPCC (http://www.cirrus.ac.uk) funded by the University of Edinburgh and EPSRC (EP/P020267/1). We would like to thank Amazon AWS and Microsoft Azure for computational credits. We would like to thank the Stan Development team for their constant support. We acknowledge the Medical Research Council and FAPESP (MR/S0195/1).

Figure 4: Mobility covariates from Google mobility reports for São Paulo (SP), Rio de Janeiro (RJ), Pernambuco (PE), Ceará (CE), Amazonas (AM).

6 Appendix

6.1 Model

We adopt the Bayesian semi-mechanistic model from [7] to estimate transmission intensity and attack rates of COVID-19 conditional on the reported number of deaths. The code base for our work can be found at `https://github.com/ImperialCollegeLondon/covid19model`. For the Brazilian model at state level, four covariates related to Google mobility were included. These describe the reduction or increase in mobility in residential areas ($k = 1$), transit stations ($k = 2$), parks ($k = 3$) and the average between groceries and pharmacies, retail and recreational areas and workplaces ($k = 4$), which are averaged due to collinearity.

As adopted in the recent published report, a subnational analysis for Italy [19], the time-varying reproduction number R_t is modelled as function of Google mobility data. Parameters are jointly estimated for 16 Brazilian states to evaluate if interventions taken so far were able to reduce R_t below to 1. Partial and full pooling were adopted, but produced almost identical results.

Denote $I_{k,t,m}$ as the k-th Google mobility indicator, at time t for Brazilian state m. The time-varying reproduction number for Brazilian state m, $R_{t,m}$, is modeled by:

$$R_{t,m} = R_{0,m} \left(2\lambda^{-1}(-\sum_{k=1}^{4}(\alpha_k + \beta_{m,k})I_{k,t,m}) \right)$$

where λ^{-1} denotes the logistic function, α_k the effects shared between M states and $\beta_{m,k}$ state-specific

effects. Prior distributions for the partial pooling model were set as

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

while the prior distribution for $R_{0,m}$ was chosen to be

$$R_{0,m} \sim \mathcal{N}(3.28, |\kappa|)$$
$$\kappa \sim \mathcal{N}(0, 0.5)$$

with κ being the same among all states to share information about the variability of $R_{0,m}$. The value of 3.28 was already used in [7, 19] based on [12].

6.2 Death underreporting scenarios

In this work, an extension of the semi-mechanistic Bayesian hierarchical model from [7] 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 5.

As in the original model [7], 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 ψ. 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 ϕ follows a normal distribution, i.e.

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

in which $\mathcal{N}(\mu, \sigma)$ denotes a normal distribution with mean μ and standard deviation σ. The rest of the mathematical model follows the original manuscript [7] introducing the new feature of underreporting death rate ψ on daily deaths.

The effect of death underreporting on the attack rate is shown in Table 2 for three additional scenarios: 33% and 50% and 67% underreporting. The underreporting scenarios are implemented by scaling reported death data by beta distributions with means (0.67, 0.5, 0.33) and in each instance variance 0.004. The distributions are shown in Figure 5.

State	0% underreporting	33% underreporting	50% underreporting	67% underreporting
SP	3.30 [2.83, 3.68]	4.90 [4.13, 5.52]	6.49 [5.44, 7.35]	9.58 [7.74, 11.00]
RJ	3.35 [2.83, 3.78]	5.05 [4.19, 5.74]	6.75 [5.50, 7.74]	10.20 [8.09, 11.90]
CE	4.46 [3.73, 5.05]	6.66 [5.46, 7.61]	8.76 [7.14, 10.10]	12.90 [10.30, 14.90]
PE	3.00 [2.49, 3.41]	4.50 [3.67, 5.15]	5.94 [4.79, 6.84]	8.86 [6.93, 10.30]
AM	10.60 [8.84, 12.10]	15.40 [12.80, 17.60]	19.90 [16.40, 22.80]	27.70 [22.80, 31.80]
PA	5.05 [3.90, 5.90]	7.63 [5.83, 8.95]	9.95 [7.50, 11.80]	14.90 [11.10, 17.70]
MA	2.07 [1.65, 2.40]	3.12 [2.44, 3.63]	4.15 [3.24, 4.83]	6.23 [4.71, 7.37]
BA	0.40 [0.31, 0.47]	0.61 [0.47, 0.71]	0.81 [0.62, 0.95]	1.26 [0.93, 1.49]
ES	2.24 [1.71, 2.64]	3.36 [2.52, 3.99]	4.45 [3.34, 5.25]	6.69 [4.82, 8.04]
PR	0.25 [0.19, 0.29]	0.37 [0.28, 0.44]	0.50 [0.37, 0.59]	0.74 [0.53, 0.89]
MG	0.13 [0.10, 0.16]	0.20 [0.15, 0.24]	0.27 [0.19, 0.33]	0.41 [0.29, 0.49]
PB	0.64 [0.48, 0.75]	0.97 [0.72, 1.15]	1.31 [0.96, 1.56]	1.98 [1.41, 2.39]
AL	1.20 [0.87, 1.44]	1.81 [1.27, 2.18]	2.41 [1.70, 2.89]	3.66 [2.51, 4.42]
RS	0.42 [0.32, 0.50]	0.65 [0.48, 0.77]	0.87 [0.64, 1.05]	1.35 [0.95, 1.62]
RN	0.56 [0.42, 0.67]	0.85 [0.62, 1.02]	1.16 [0.83, 1.39]	1.77 [1.23, 2.14]
SC	0.23 [0.17, 0.27]	0.34 [0.25, 0.41]	0.46 [0.33, 0.55]	0.69 [0.48, 0.84]

Table 2: Estimated attack rates for 0%, 33%, 50% and 67% death underreporting scenarios.

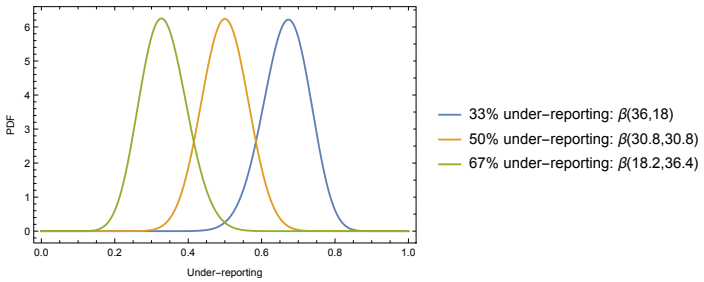

Figure 5: Underreporting prior distributions.

6.3 Cases and R_t for 16 states

The estimated cases, deaths and R_t for all 16 states considered in our joint model, São Paulo (SP), Rio de Janeiro (RJ), Pernambuco (PE), Ceará (CE), Amazonas (AM), Pará (PA), Maranhão (MA), Bahia (BA), Espírito Santo (ES), Paraná (PR), Minas Gerais (MG), Paraíba (PB), Rio Grande do Sul (RS), Rio Grande do Norte (RN), Alagoas (AL), Santa Catarina (SC), are shown in Figure 6.

DOI: https://doi.org/10.25561/78872

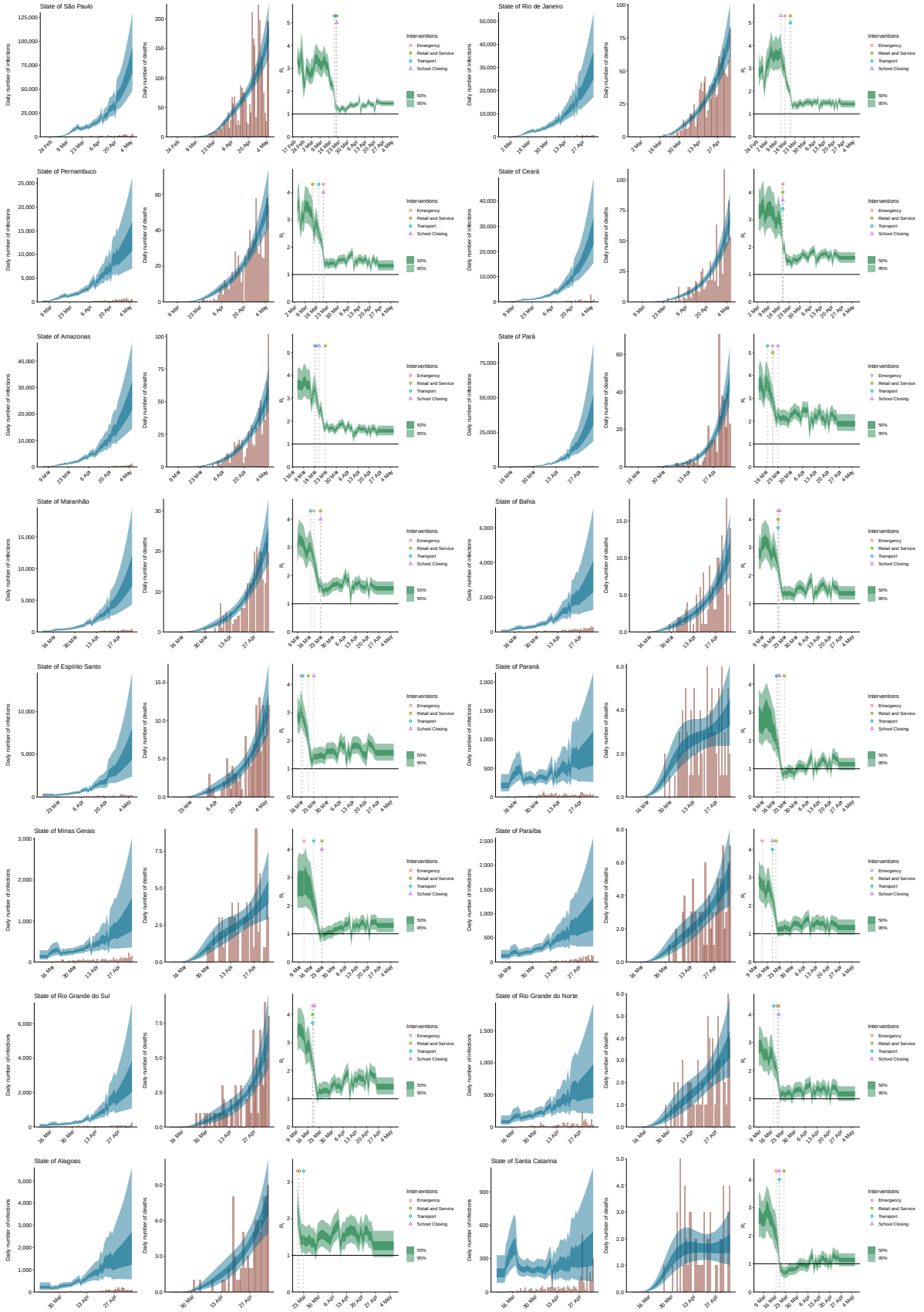

Figure 6: Estimates of infections, deaths and R_t for all 16 states considered in the model.

DOI: https://doi.org/10.25561/78872

6.4 Full and partial-pooling sensitivity

The results in this work have been produced with partial pooling of covariate coefficients. Full pooling results are shown in Table 3. There is no substantial difference.

State	AR% half pooling	AR% full pooling
SP	3.30 [2.83, 3.68]	3.32 [2.86, 3.69]
RJ	3.35 [2.83, 3.78]	3.33 [2.81, 3.74]
CE	4.46 [3.73, 5.05]	4.31 [3.62, 4.88]
PE	3.00 [2.49, 3.41]	3.00 [2.51, 3.41]
AM	10.60 [8.84, 12.10]	10.60 [8.94, 12.00]
PA	5.05 [3.90, 5.90]	4.67 [3.72, 5.41]
MA	2.07 [1.65, 2.40]	2.07 [1.65, 2.40]
BA	0.40 [0.31, 0.47]	0.39 [0.31, 0.45]
ES	2.24 [1.71, 2.64]	2.17 [1.69, 2.55]
PR	0.25 [0.19, 0.29]	0.24 [0.18, 0.29]
MG	0.13 [0.10, 0.16]	0.13 [0.10, 0.15]
PB	0.64 [0.48, 0.75]	0.65 [0.49, 0.77]
AL	1.20 [0.87, 1.44]	1.15 [0.86, 1.37]
RS	0.42 [0.32, 0.50]	0.42 [0.32, 0.50]
RN	0.56 [0.42, 0.67]	0.57 [0.43, 0.68]
SC	0.23 [0.17, 0.27]	0.23 [0.17, 0.28]

Table 3: Attack rate (AR) by state, with half and full-pooling of mobility covariates between states.

6.5 Onset-to-death sensitivity

Onset-to-death distribution sensitivity analysis for attack rate is shown in Table 4. Outcomes are not substantially affected by perturbation of onset-to-death distribution mean by plus or minus 10%.

6.6 IFR Calculation and Sensitivity Analysis

Estimates of the expected IFR across different states are derived from previously published estimates of mixing patterns in a Latin America setting [3] alongside estimates of the virus' transmissibility (the basic reproduction number, R_0) derived from European settings [7] and from estimates of disease sever-

State	AR% onset-to-death mean decreased by 10%	AR% onset-to-death mean increased by 10%
SP	3.15 [2.72, 3.51]	3.38 [2.88, 3.79]
RJ	3.17 [2.68, 3.58]	3.54 [2.96, 4.01]
CE	4.04 [3.38, 4.57]	4.96 [4.10, 5.66]
PE	2.77 [2.32, 3.13]	3.21 [2.65, 3.65]
AM	9.80 [8.22, 11.10]	11.50 [9.56, 13.10]
PA	4.38 [3.38, 5.12]	5.91 [4.51, 6.97]
MA	1.86 [1.50, 2.14]	2.29 [1.80, 2.67]
BA	0.36 [0.29, 0.42]	0.44 [0.34, 0.51]
ES	1.98 [1.52, 2.32]	2.51 [1.88, 2.98]
PR	0.24 [0.19, 0.28]	0.25 [0.19, 0.30]
MG	0.13 [0.09, 0.15]	0.14 [0.10, 0.17]
PB	0.60 [0.46, 0.70]	0.68 [0.50, 0.81]
AL	1.09 [0.79, 1.29]	1.32 [0.94, 1.59]
RS	0.37 [0.28, 0.44]	0.48 [0.35, 0.57]
RN	0.53 [0.39, 0.63]	0.59 [0.43, 0.71]
SC	0.22 [0.17, 0.26]	0.23 [0.17, 0.28]

Table 4: Attack rate (AR) with onset-to-death distribution mean decreased by 10% to 16.9 days and increased by 10% to 20.7 days.

ity derived from the Chinese epidemic [15] and subsequently modified to match data emerging from the epidemic in the United Kingdom [7]. We additionally modified these estimates of disease severity (specifically the Infection Fatality Ratio or 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 varies from as high as $\sim \$300$ in São Paulo to as low as $\sim \$100$ in Maranhão.[8] Such disparities in income are likely to result in significant disparities 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.[15]), and that the state with the lowest income (Maranhnão) had significantly worse healthcare outcomes - more similar to those that would be expected in a Lower Middle Income Country (see Walker et al., [20] for further details on how differences in health quality across set-

tings are likely to impact outcomes). Details on the age-specific infection fatality probabilities for these two states are provided in Table 5. 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.[8] These age-specific infection fatality probabilities are then combined with predictions of the age-distribution of infections to produce an overall, state-specific IFR.

Ages	IFR% São Paulo	IFR% Maranhão
0-4	0.0028	0.021
5-9	0.0024	0.018
10-14	0.0044	0.033
15-19	0.0091	0.067
20-24	0.020	0.15
25-29	0.039	0.29
30-34	0.062	0.45
35-39	0.094	0.66
40-44	0.13	0.82
45-49	0.22	1.13
50-54	0.45	1.77
55-59	0.82	2.38
60-64	1.72	3.70
65-70	2.71	4.73
75-80	4.25	6.47
80-84	6.15	8.47
85+	9.63	12.57

Table 5: IFR by age for São Paulo and Maranhão.

Substantial uncertainty still remains in these IFR calculations however, and 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 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 6 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 to date sufficient to have reached it.

DOI: https://doi.org/10.25561/78872

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	3.42 [2.93, 3.82]	3.28 [2.82,3.67]	3.41 [2.93, 3.80]	3.30 [2.83, 3.68]
RJ	3.67 [3.08, 4.13]	3.52 [2.96,3.98]	3.49 [2.92, 3.94]	3.35 [2.83, 3.78]
CE	8.17 [6.86, 9.24]	7.83 [6.58,8.85]	4.55 [3.79, 5.15]	4.46 [3.73, 5.05]
PE	5.53 [4.61, 6.28]	5.31 [4.43,6.03]	3.04 [2.52, 3.45]	3.00 [2.49, 3.41]
AM	21.00 [17.90,23.70]	20.70 [17.60,23.40]	10.60 [8.85, 12.00]	10.60 [8.84, 12.10]
PA	10.40 [8.13, 12.20]	10.30 [8.02,11.90]	5.06 [3.90, 5.94]	5.05 [3.90, 5.90]
MA	4.53 [3.60, 5.24]	4.39 [3.51,5.08]	2.08 [1.65, 2.41]	2.07 [1.65, 2.40]
BA	0.76 [0.59, 0.89]	0.73 [0.57,0.85]	0.40 [0.32, 0.47]	0.40 [0.31, 0.47]
ES	3.14 [2.39, 3.71]	3.04 [2.32,3.58]	2.27 [1.72, 2.67]	2.24 [1.71, 2.64]
PR	0.32 [0.24, 0.37]	0.30 [0.23,0.36]	0.25 [0.19, 0.30]	0.25 [0.19, 0.29]
MG	0.20 [0.15, 0.23]	0.19 [0.14,0.23]	0.14 [0.10, 0.16]	0.13 [0.10, 0.16]
PB	1.21 [0.91, 1.44]	1.14 [0.86,1.36]	0.65 [0.49, 0.77]	0.64 [0.48, 0.75]
AL	2.61 [1.89, 3.12]	2.52 [1.82,3.01]	1.21 [0.87, 1.45]	1.20 [0.87, 1.44]
RS	0.47 [0.36, 0.56]	0.45 [0.34,0.54]	0.44 [0.33, 0.52]	0.42 [0.32, 0.50]
RN	1.01 [0.75, 1.20]	0.96 [0.71,1.16]	0.57 [0.42, 0.68]	0.56 [0.42, 0.67]
SC	0.27 [0.20, 0.32]	0.26 [0.19,0.31]	0.23 [0.17, 0.28]	0.23 [0.17, 0.27]

Table 6: 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.

DOI: https://doi.org/10.25561/78872

State	(i) IFR UK contact matrix	(ii) IFR Peru contact matrix	(iii) IFR UK contact matrix, poorer out-comes	(iv) IFR Peru contact matrix, poorer out-comes
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 7: 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.

DOI: https://doi.org/10.25561/78872

6.7 Data

As input of deaths and reported cases, our model uses daily updates from a government initiative funded by Brazil's ministry of Health called *Painel Coronavírus*, available at `https://covid.saude.gov.br`. In this data the number of deaths attributable to COVID-19 is segmented by state level. Possible under-reporting in death data is addressed in the mathematical model described in Section 6.1.

For population counts we used the 2020 projection by state published by *Instituto Brasileiro de Geografia e Estatística* (IBGE).[9]

Mobility report data from Google (`https://www.google.com/covid19/mobility/`) 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 and recreation, groceries and pharmacies, parks, transit stations, workplaces, and residential comparing to a baseline. Such baseline corresponds to the median value of each day of the week, using data of January 3rd to February 6th, 2020. More details can be found in Figure 7.

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.

DOI: https://doi.org/10.25561/78872

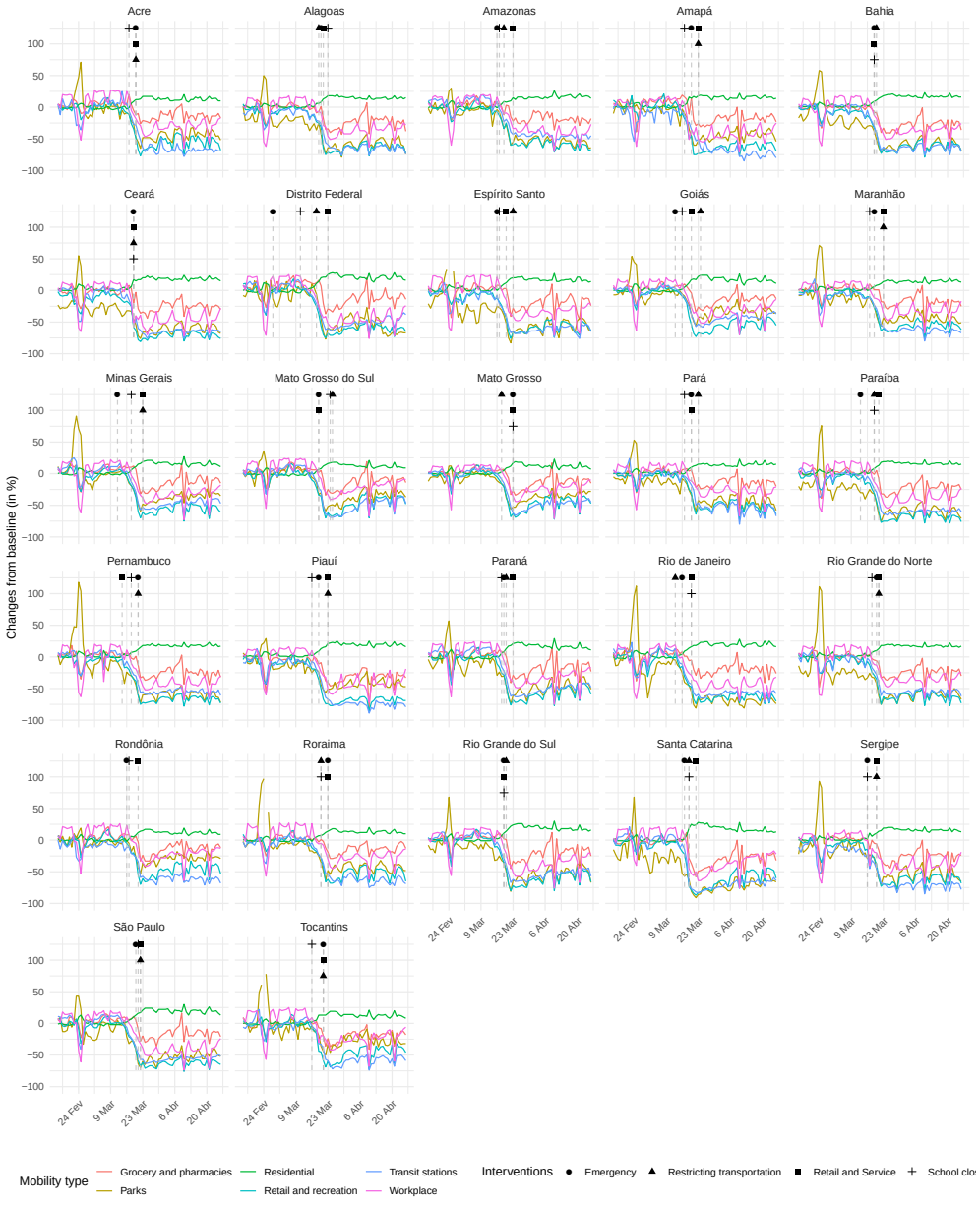

Figure 7: Google mobility time series for all Brazilian states with their respective interventions from 15th February to 26th April 2020.

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 8: Non-pharmaceutical interventions by state, adapted from [4].

DOI: https://doi.org/10.25561/78872

References

[1] KEC Ainslie et al. “Evidence of initial success for China exiting COVID-19 social distancing policy after achieving containment [version 1; peer review: awaiting peer review]”. In: *Wellcome Open Research* 5.81 (2020).

[2] Ahmet Aktay et al. “Google COVID-19 Community Mobility Reports: Anonymization Process Description (version 1.0)”. In: *ArXiv* abs/2004.0 (2020).

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

[4] “COVID-19 Observatory in Latin America and the Caribbean”. In: URL: https://www.cepal.org/en/topics/covid-19.

[5] Samy Dana et al. “Brazilian Modeling of COVID-19 (BRAM-COD): a Bayesian Monte Carlo approach for COVID-19 spread in a limited data set context”. In: *medRxiv* (2020). eprint: https://www.medrxiv.org/content/early/2020/05/03/2020.04.29.20081174.full.pdf. URL: https://www.medrxiv.org/content/early/2020/05/03/2020.04.29.20081174.

[6] NM Ferguson et al. *Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand (Report 9)*. Tech. rep. URL: https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/news--wuhan-coronavirus/ (visited on 03/25/2020).

[7] Seth Flaxman et al. “Report 13: Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries”. In: (2020). URL: https://doi.org/10.25561/77731.

[8] “IBGE divulga o rendimento domiciliar per capita 2019”. In: (). URL: https://agenciadenoticias.ibge.gov.br/agencia-sala-de-imprensa/2013-agencia-de-noticias/releases/26956-ibge-divulga-o-rendimento-domiciliar-per-capita-2019.

[9] “IBGE Projeções da População”. In: (). URL: https://www.ibge.gov.br/estatisticas/sociais/populacao/9109-projecao-da-populacao.html?=&t=resultados.

[10] Moritz U. G. Kraemer et al. “The effect of human mobility and control measures on the COVID-19 epidemic in China”. In: *Science* 368.4690 (2020).

[11] Enrico Lavezzo et al. “Suppression of COVID-19 outbreak in the municipality of Vo, Italy”. In: *medRxiv* (2020).

[12] Ying Liu et al. "The reproductive number of COVID-19 is higher compared to SARS coronavirus". In: *Journal of Travel Medicine* 27.2 (Feb. 2020). taaa021. ISSN: 1195-1982. eprint: https://academic.oup.com/jtm/article-pdf/27/2/taaa021/32902430/taaa021.pdf. URL: https://doi.org/10.1093/jtm/taaa021.

[13] World Health Organisation. "WHO, Coronavirus disease 2019 (COVID-19) Situation Report – 107". In: (2020). URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports.

[14] "Painel Coronavírus". In: (). URL: https://covid.saude.gov.br/.

[15] 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 Neil M Ferguson Robert Verity Lucy C Okell. "Estimates of the severity of coronavirus disease 2019: a model-based analysis". In: *Lancet Infectious Diseases* (2020).

[16] William Marciel de Souza et al. "Epidemiological and clinical characteristics of the early phase of the COVID-19 epidemic in Brazil". In: *medRxiv* (2020). eprint: https://www.medrxiv.org/content/early/2020/04/29/2020.04.25.20077396.full.pdf. URL: https://www.medrxiv.org/content/early/2020/04/29/2020.04.25.20077396.

[17] Robert Verity et al. "Estimates of the severity of COVID-19 disease". In: *Lancet Infect Dis* in press (2020).

[18] Julia Gog Ottar N. Bjørnstad Stephen Kissler Lone Simonsen Bryan T. Grenfell Cécile Viboud Vivek Charu Scott Zeger. "Human mobility and the spatial transmission of influenza in the United States". In: *PLOS Computational Biology* (2017).

[19] Michaela A. C. Vollmer et al. "Report 20: Using mobility to es mate the transmission intensity of COVID-19 in Italy: A subnational analysis with future scenarios". In: (2020).

[20] P. G. T. Walker et al. *Report 12: The Global Impact of COVID-19 and Strategies for Mitigation and Suppression*. URL: https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/news--wuhan-coronavirus/.