medRxiv preprint doi: https://doi.org/10.1101/2020.07.28.20163493; this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

The effective reproductive number (Rt) of COVID-19 and its relationship with social distancing

José Alexandre Felizola Diniz-Filho[1], Lucas Jardim[2], Cristiana M. Toscano[3], Thiago Fernando Rangel[1]

[1] Departamento de Ecologia, Instituto de Ciências Biológicas, Universidade Federal de Goiás (UFG), Brasil

[2] Instituto Nacional de Ciência & Tecnologia em Ecologia, Evolução e Conservação da Biodiversidade, UFG, Programa DTI/CNPq, Brasil;

[3] Departamento de Saúde Coletiva, Instituto de Patologia Tropical e Saúde Pública (IPTSP), UFG, Brasil.

José Alexandre F. Diniz-Filho (https://orcid.org/0000-0002-0967-9684)

Lucas Jardim (https://orcid.org/0000-0003-2602-5575)

Cristiana M. Toscano (https://orcid.org/0000-0002-9453-2643)

Thiago F. Rangel (https://orcid.org/0000-0002-2001-7382)

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.28.20163493; this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

ABSTRACT

The expansion of the new coronavirus disease (COVID-19) triggered a renewed interest in epidemiological models and on how parameters can be estimated from observed data. Here we investigated the relationship between average number of transmissions though time, the reproductive number *Rt*, and social distancing index as reported by mobile phone data service *inloco*, for Goiás State, Brazil, between March and June 2020. We calculated *Rt* values using *EpiEstim* package in R-plataform for confirmed cases incidence curve. We found a correlation equal to -0.72 between *Rt* values and isolation index at a time lag of 8 days. As the *Rt* values were paired with center of the moving window of 7 days, the delay matches the mean incubation period of the virus. Our findings reinforce that isolation index can be an effective surrogate for modeling and epidemiological analyses and, more importantly, helpful for anticipating the need for early interventions, a critical issue in public health.

Key words: Social isolation – time series analysis – correlation – effective R - COVID-19.

INTRODUCTION

The global expansion of the new coronavirus disease (COVID-19) triggered a great and renewed public interest in epidemiological models to understand temporal and geographical patterns of expansion of the pandemics and to use such models to guide decision making processes regarding public health measures to mitigate its spread[1,2]. As so, it is of utmost importance to better understand clinic, immunologic and epidemiologic characteristics of the new coronavirus (SARS-CoV-2) that causes COVID-19 and how these can be translated into statistical parameters that can be estimated from observed data and used in predictive models. In the core of such discussions is the basic reproduction number (R_0), which reflects the average number of new infections generated from an initial infection in a susceptible population[3]. This parameter synthetizes, under a more realistic view, incubation and transmission periods of the infectious agent[4,5], but actually it is not constant though time. During an epidemic, changes in the effective or realized infection transmission through time (here denoted as *Rt*) also reflect the demographics of the host-pathogen interaction, which in turn is regulated by several

medRxiv preprint doi: https://doi.org/10.1101/2020.07.28.20163493; this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

extrinsic environmental effects (mainly population density and social contact)[6,7]. Thus, in the particular case of COVID-19 pandemic, for which no specific therapeutic or preventive strategy is yet available, the downward shifts in *Rt* is the expected outcome of non-pharmaceutical interventions, including social distancing measures to reduce the number of transmissions[1]. This result in a flattening of the epidemic curve, allowing healthcare systems to prepare and manage the expected cases of disease, while waiting for a vaccine to be developed and made available.

Although it is relatively simple to estimate the infection *Rt* based on incidence data and estimates of serial interval[8], it is much more difficult to explain its continuous temporal variation given the implementation of interventions, in particular social distancing measures[6]. Given the characteristics of COVID-19, among them a large proportion of asymptomatic cases, the ability to be transmitted by asymptomatic individuals, and its rapid spread, it is impossible to accurately identify the number of infected individuals early on in the epidemic. As such, it is important to have surrogates for continuous tracking of the *Rt* dynamics. The broad-scale real time monitoring of mobility derived from mobile phones has been reported to significantly correlate with decrease in the number of COVID-19 cases and increased social distance in the population[6]. Particularly in extremes of the epidemic, social distancing estimated by mobility indicators have been one surrogate of the *Rt* dynamics[7, 9, 10]. However, in addition to these extremes, it is interesting to assess whether a more continuous relationship between mobility and *Rt* over time, as this would allow closer and timely monitoring of disease transmission which may in turn guide public health responses tailored at the local level[6].

Here we investigate the relation between *Rt*, estimated from incidence curves of confirmed casos and deaths, and social distancing measured by mobility indicators from mobile phone data in Goiás State, in Central Brazil. The first COVID-19 cases were identified in Goiás in mid-March, mostly clearly imported from São Paulo and foreigner countries. Shortly after the identification of these few imported cases notified to the local surveillance system, the State Government implemented strict social distancing measures, including cancellation of events, school and workplace closures, closure of commercial establishments, and services except for essential services[11] (see also https://medidas-covidbr-iptsp.shinyapps.io/painel/). After that social distancing was relaxed and isolation

medRxiv preprint doi: https://doi.org/10.1101/2020.07.28.20163493; this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

indicators gradually decreased, thus providing an opportunity to evaluate the continuous relationship between *Rt* and social distancing.

METHODS

Goiás state has approximately 7 million inhabitants living in its 246 municipalities[12]. We considered epidemiologic and mobility data from all municipalities in the state. We here present the mobility indicator as its inverse denominated isolation indicator (resulting from 1-mobility and presented as %), as reported by mobile phone data service *inloco* (https://mapabrasileirodacovid.inloco.com.br/pt/). Isolation indicador was obtained for each municipality in Goiás state. Epidemiologic data for COVID-19 including confirmed cases and deaths were obtained from the Goias State Health Department (SES-GO, available at http://covid19.saude.go.gov.br/). The study period ranged from March 16 to June 6, 2020.

We estimated the *Rt* values based on the incidence curve of confirmed COVID-19 symptomatic cases reported in the study period representing a total of 14,751 cases. We also obtained data from deaths (summing 500 confirmed deaths by July 01, by date of death) and used the package *EpiEstim* of the *R* plataform[13, 14] to analyze incidence by dates of symptom. Despite the significant delay for case confirmation in the system, of approximately 7-10 days in Goiás state, as we considered cases by date of symptom onset, we have enough replicatons to test the main effects of change in the isolation indicator. We estimated the posterior distribution of *Rt* modeling the COVID-19 disease incidences through time as a Poisson distribution with mean $Rt \sum_s^t I_{t-s} w_s$,, where *ws* is the probability of secondary infection time[14]. A gamma distribution (a = 1, b = 5) prior distribution is assumed for *Rt* with discretized gamma distribution with generation time $\mu = 5.2$ and $\sigma = 0.6$ for *ws*, approximating the one used by other authors[7, 8, 15]. The values of *Rt* were calculated using a moving average considering 7-day window.

Although more sophisticated methods are available to account for such delays and eventually sub notifications[16], here we simply truncated the end of the distribution of cases as our goal is not to have a real-time estimate of *Rt*, but rather to correlate the shifts in the time series with the isolation indicator and to evaluate the impact more rigid or flexible social distancing measures in place in the State throughout the study period, as well as the observed decreasing levels of isolation from middle April on.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.28.20163493; this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

The population-weighted mean isolation indicators for the entire state were correlated with the *Rt* values, initially using standard Pearson correlations. We paired the *Rt* values to isolation assuming that each *Rt* reflects the middle of each time window. However, considering date of symptom onset and the fact that infection transmission take place in average 5 days earlier[4, 5], we also calculated these correlations considering different time lags to evaluate its behavior. Because of the strong autocorrelation in both series, especially in the *Rt* values (as they are calculated using overlapping moving windows), it is not appropriate to apply a standard statistical test it due to inflated Type I error. We then used Dutilleul's approach based on temporal Moran's *I* correlograms to reduce the degrees of freedom taking into account positive autocorrelation in both time series[17].

RESULTS

These interventions resulted in a significant increase in the isolation indicator from about 25% to ca. 55% within a one-week period (Figure 1A). As expected, the COVID-19 epidemic curve grew very slowly in the state, as well as in surroundings states encompassing the entire Central region of Brazil, from mid-March to mid-May 2020. Up to early June, a total of 172 confirmed COVID-19 deaths were reported in Goiás, representing a mortality rate of 2.45/100,000 inhabitants. After this 2 month-period, isolation indicator decreased steadily together with observed increase case counts and disease transmission from mid-May onwards.

The time-series of *Rt* (Figure 1B) shows clearly the effects of the early and strict social distancing measures in Goiás State in mid-March 2020, reducing the *R* (which we here call R_0, although by definition disease transmission began occurring locally when measures were already in place) from around 2.0, to slightly higher than 1.0 in middle April 2020. However, it is possible to observe that *Rt* increases slowly from this point on, reaching 1.4 - 1.5 in late May/early June 2020.

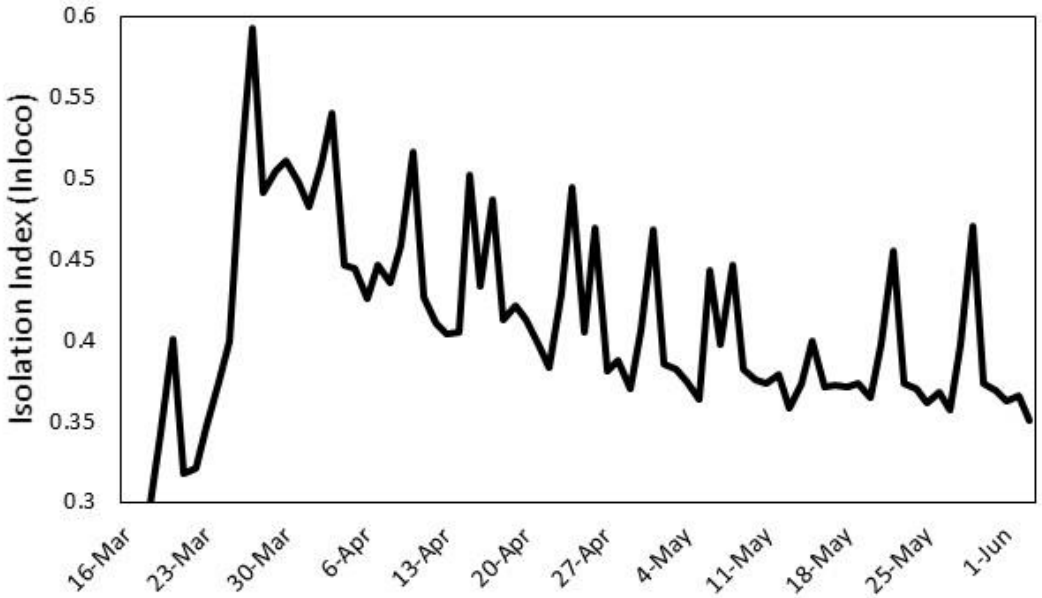
Even so, despite the similarity of the overall pattern of the time-series, when we correlate *Rt* and the isolation indicator at a fixed data, there is a relatively low correlation (r = -0.281). However, it is indeed necessary to consider a time-lag to reveal a relatively high negative correlation, and it is possible to find a much higher correlation of -0.696 at time lag 8 (so that higher number of transmissions in the middle of the time window is

medRxiv preprint doi: https://doi.org/10.1101/2020.07.28.20163493; this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

indeed associated with low isolation 8 days before; as we used a 7 day time window, it closely matches the 5 average days of incubation of the virus). The relationship may be slightly non-linear (Figure 2), although Spearman rank correlation is not that different from linear correlation (-0.645). O power fit (i.e., log-transformation in both variables) increased only slightly the correlation to -0.72. The first class autocorrelation using Moran's I autocorrelation coefficient in both variables is around 0.5, decreasing to zero after ca. 21 days, and thus Dutileull's approach suggest that correlation between the two series should be effectively tested with about 20 degrees of freedom (and even in this case these correlations are significant at $P < 0.01$).

medRxiv preprint doi: https://doi.org/10.1101/2020.07.28.20163493; this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

A)



B)

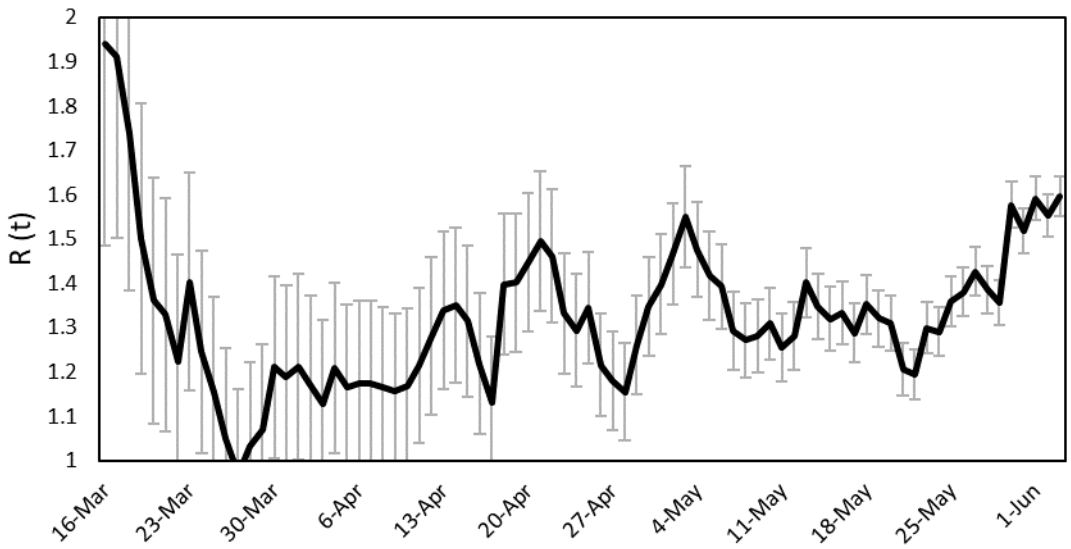


Figure 1. Times series of Isolation index (*inloco)* (A) and mean reproduction number (*Rt*) through time in Goiás State (B) between March – June 2020, with respective 95% confidence intervals.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.28.20163493; this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

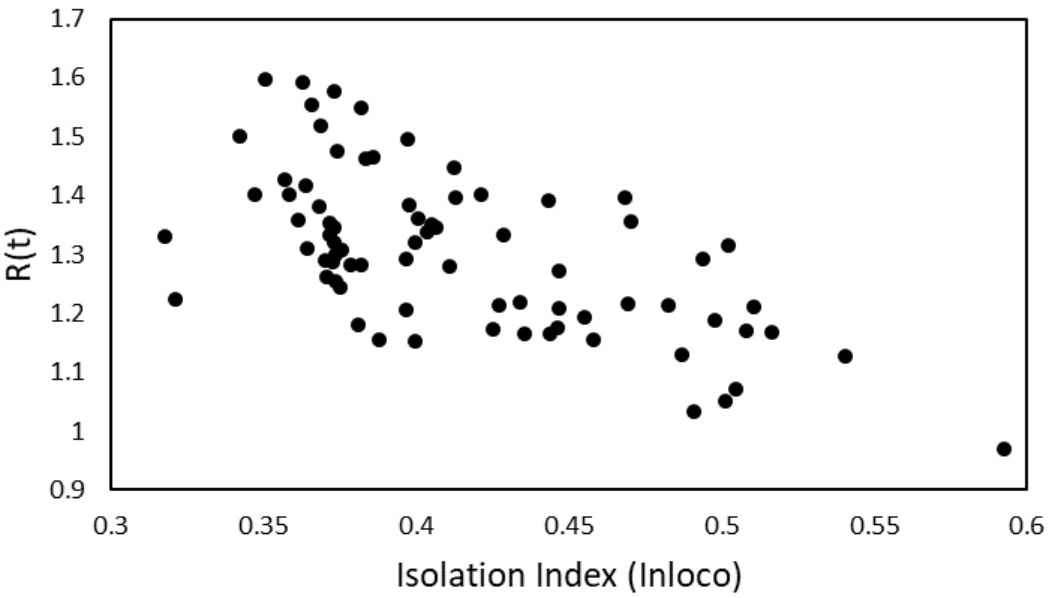


Figure 2. Negative relationship between Reproduction number *Rt* and isolation index measured by *inloco*, Goiás State, between March and late May. The correlation is equal to 0.696 at original scale and, due to a slightly non-linearity, it increases to 0.721 after log-transforming both axis).

In the scatterplot shown in Figure 2 we were not able to identify any systematic effects or trends in regression residuals, although there was a reinforcement of the behavioral measures from middle April onwards (although we are not aware of the effectiveness of such measures and do not have a quantitative account of their effects).

DISCUSSION

There is a real interest to evaluate the local dynamics of disease transmission, as this is the main parameter that drives the spread of COVID-19 in a region. Several analyses have shown, as expected, that social distancing measures effectively reduce the reproduction number (*Rt*) parameter[6, 7]. However, it is actually difficult to evaluate the true effects of such measures in a more realistic fashion, as there may be several confounding and interactions between broad-scale isolation related to mobility and more local and "behavioral" aspects at very local an individual scale, including hygiene measures, mask use and social awareness as a result of the pandemic[18, 19], as well as a more effective contact tracing and early detection of cases[20].

medRxiv preprint doi: https://doi.org/10.1101/2020.07.28.20163493; this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

We identified for Goiás State a relatively high negative correlation of -0.696 (increasing to -0.72 at log-scale) between the estimated *Rt* and isolation indicator obtained from mobile phone data. Our empirical results match those reported for São Paulo State, where the *Rt* decreases from about 2 to 1 after strict social distancing interventions were implemented in mid-March 2020. Similar findings were also reported for broad-scale analysis of other countries[6, 9]. However, our results show that this high correlation between *Rt* and isolation index suggests a more continuous relationship between these variables, as recently calculated for United States[6], rather than a more discrete variation of extremes (i.e, if isolation increases to 50% the *Rt* decreases to 1.0, as shown for São Paulo state[21], France[22] and Europe[7], and used for initial modeling in Brazil[10]).

Even more interestingly, the negative correlation between isolation and *Rt* appears only if a time-lag is created, which is indeed expected considering current knowledge of COVID-19 disease and in particularly its incubation time. This correlation may seems clear to our case in Goiás perhaps due to the rich behavior of the time series, with an initial increase in the isolation indicators shortly after the start disease transmission locally, followed by a more continuous reduction in the isolation indicator, and finally by a stabilization at around 38% in late May. We expect an even richer dynamics as the State has now implemented intermittent lockdown with cycles of 14 days, although not yet followed by all municipalities. Even so, it is also important to note that this correlation between *Rt* and isolation indicator must be better detected in early phases of the COVID-19 expansion, and it will tend to disappear as the epidemic size increases and the peak of transmissions occurs because *Rt* will be also more strongly affected by the reduction of number of susceptible individuals in the population.

In conclusion, about 50% of the variation in the mean number of transmissions expressed by *Rt* in the Goiás state is explained by broad-scale isolation indicator measured by mobile phones. Although there is a relatively large amount of unexplained variation, we do not have additional variables related to more local measures that could improve our understanding of disease transmission. Considering the amount of variation explained and the randomness of the residuals of the relationship between isolation and *Rt*, we understand that using isolation indicators obtained from mobile phones can be a good surrogate for COVID-19 transmission in a given population, and can be used for modeling purposes as well as routine epidemiological analyses. More importantly, it and can be

medRxiv preprint doi: https://doi.org/10.1101/2020.07.28.20163493; this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

quite useful for anticipating the need for early interventions, an important and critical issues in public health.

ACKNOWLEDGMENTS

Authors thank National Council for Scientific and Technological Development (CNPq) for productivity fellowships and grants. Our work on simulation and population modeling has been continuously supported by National Institute of Science and Technology (INCT) in Ecology, Evolution and Biodiversity Conservation (J. A. F. D.-F. and T.F.R), supported by CNPq and FAPEG. CMT is a researcher of the Institutes for Health Technology Assessment (IATS), also supported as a INCT by CNPq. We thank Goias State and Goiania city Health Departments for support and access to original data.

AUTHORS' CONTRIBUTION

J.A.F.D.-F., C. M.T. and T.F.R designed the work and J. A. F. D.-F. did the preliminary analyses; LJ implemented the code for *EpiEstim* and did the first analyses, whereas T. F. R. processed all data on social index; all authors contributed to the writing and reviewing of the final submitted version.

REFERENCES

1. Adam, D. Modeling the pandemic. Nature 2020; 580: 316-318.

2. Jewell NP, Lewnard JA, Jewell BL. Predictive Mathematical Models of the COVID-19 Pandemic: Underlying Principles and Value of Projections. JAMA 2020; 323(19): 1893-1894.

3. Ridenhour B, Kowalik JM, Shay DK. Unraveling R0: considerations for public health applications. Am. J. Public Health 2014; 104: e32-241.

4. Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, et al. The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. Ann Intern Med. 2020 ; 172(9): 577-582.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.28.20163493; this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

5. He X, Lau EHY, Wu P, Deng X, Wang J, Hao X, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. Nature Medicine 2020; 26: 672–675.

6. Badr HS, Du H, Marshall M, Dong E, Squire MM, Gardner LM. Association between mobility patters and COVID-19 transmission in the USA: a mathematical modelling study. Lancet Infect Dis 2020 Jul 1.

7. Flaxman S, Mishra S, Gandy A, Unwin HJT, Mellan TA, Coupland H, et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature 2020 Jun 8.

8. Abbott S, Hellewell J, Thompson RN, Sherratt K, Gibbs HP, Bosse NI, et al. Estimating the time-varying reproduction number of SARS-CoV-2 using national and subnational case counts. Wellcome Open Research 2020 5:112. https://doi.org/10.12688/wellcomeopenres.16006.1.

9. Kupferschmidt, K. The lockdowns worked—but what comes next? Science 2020; 368: 218-219

10. Mellan TA, Hoeltgebaum HH, Mishra S, Whittaker C, Schnekenberg RP, Gandy A, et al. Report 21: Estimating COVID-19 cases and reproduction number in Brazil. Imperial College London 2020 May 05. https://doi.org/10.25561/78872.

11. Silva LLS, Lima AFR, Polli DA, Razia PFS, Pavão LFA, Cavalcanti MAFH, et al. Medidas de distanciamento social para o enfrentamento da COVID-19 no Brasil: Caracterização e análise epidemiológica por Estados. Cadernos de Saúde Pública 2020 (in press).

12. SEGPLAN, IMB – Instituto Mauro Borges de Estatística de Estudos. https://www.imb.go.gov.br/bde (acesso em abril de 2020).

14. Cori A, Ferguson NM, Fraser C, Cauchemez S. A new framework and software to estimate time-varying reproduction numbers during epidemics. Amer. J. Epidemiol. 2013; 178: 1505-1512.

13. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing 2020. Available from https://www.R-project.org.

15. Ganyani T, Kremer C, Chen D, Torneri A, Faes C, Wallinga J, et al. Estimating the generation interval for coronavirus disease (COVID-19) based on symptom onset

medRxiv preprint doi: https://doi.org/10.1101/2020.07.28.20163493; this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

data. Euro Surveill. 2020; 25. https://doi.org/10.2807/1560-7917.ES.2020.25.17.2000257.

16. Gostic KM, McGough L, Baskerville E, Abbott S, Joshi K, Tedijanto C, et al. Practical considerations for measuring the effective reproductive number, Rt. MedRxiv 2020. https://doi.org/10.1101/2020.06.18.20134858.

17. Legendre, P. and Legendre, L. Numerical Ecology, Elsevier, Amsterdan. 2013.

18. Chu DK, Akl EA, Duda S, Solo K, Yaacoub S, Schünemann HJ, et al. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. The Lancet 2020; 395: 1973-1987.

19. Kucharski AJ, Klepac P, Conlan AJK, Kissler SM, Tang ML, Fry H, et al. Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study. The Lancet Infectious Diseases 2020 Jun 26. https://doi.org/10.1016/S1473-3099(20)30457-6.

20. Peto J, Carpenter J, Smith GD, Duffy S, Houlston R, Hunter DJ, et al. Weekly COVID-19 testing with household quarantine and contact tracing is feasible and would probably end the epidemic. Proc. R. Soc. 2020 Jun 24. https://doi.org/10.1098/rsos.200915.

21. Ganem F, Mendes FM, Oliveira SB, Porto VBG, Araujo W, Nakaya H, et al. The impact of early social distancing at COVID-19 Outbreak in the largest metropolitan area of Brazil. MedRxiv 2020. https://doi.org/10.1101/2020.04.06.20055103.

22. Roques L, Klein EK, Papaïx J, Sar A. Soubeyrand S. Impact of Lockdown on the Epidemic Dynamics of COVID-19 in France. Frontiers in Medicine 2020; 7: 274. https://doi.org/10.3389/fmed.2020.00274.