EMERGING INFECTIOUS DISEASES®

Seroprevalence of IgG anti-SARS-CoV-2 in blood donors as an important measure to monitor the COVID-19 epidemic progress

Journal:	*Emerging Infectious Diseases*
Manuscript ID	EID-21-1961.R2
Manuscript Type:	Research
Date Submitted by the Author:	n/a
Complete List of Authors:	Chaves, Daniel; Fundação Hemominas, Serviço de Pesquisa Takahashi, Ricardo; Universidade Federal de Minas Gerais Campelo, Felipe; Universidade Federal de Minas Gerais; Aston University Silva-Malta, Maria Clara; Fundação Hemominas Oliveira, Isabelle; Universidade Federal de Minas Gerais Figueiredo Barbosa-Stancioli, Edel; Universidade Federal de Minas Gerais Ribeiro, Maísa; Fundação Hemominas Lobato Martins, Marina; Fundação Hemominas
Keywords:	SARS-CoV-2, COVID-19, Blood donor, Epidemic model, Seroprevalence
Abstract:	During the epidemics such as SARS-CoV-2, data from different sources can provide information on different aspects of the epidemic process. These may include records from serology-based epidemiologic surveys to compose a consistent scenario. This study assessed the seroprevalence of IgG anti-SARS-CoV-2 in 7,837 samples from blood donors in seven Brazilian cities from March to December 2020. Based on the results this work proposes a modification in a compartmental model that receives as inputs the reported number of COVID-19 cases and the serology results of blood donors and delivers estimates of hidden variables such as daily values of transmission rate and cumulative incidence rate of reported and non-reported cases. The study concludes that the information about cumulative incidence of the disease in the cities´ populations can be obtained simply by carrying out tests on samples collected from blood donors. The same methodology can be extended to surveillance of other infectious diseases.

Seroprevalence of IgG anti-SARS-CoV-2 in blood donors as an important measure to monitor the COVID-19 epidemic progress

Daniel Gonçalves Chaves[1], Ricardo Hiroshi Caldeira Takahashi[1], Felipe Campelo, Maria Clara Fernandes da Silva Malta, Isabelle Rocha de Oliveira, Edel Figueiredo Barbosa-Stancioli, Maísa Aparecida Ribeiro, Marina Lobato Martins

Author affiliations

Fundação Hemominas, Belo Horizonte, Minas Gerais, Brazil (D.G. Chaves, M.C.F. Silva-Malta, M.A. Ribeiro, M.L. Martins)

Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil (R.H.C. Takahashi, F. Campelo, I.R. de Oliveira, E.F. Barbosa-Stancioli)

Aston University, Birmingham, UK (F. Campelo)

[1]These authors contributed equally to this article.

Text word counts: 3,637

Abstract word counts: 150

Figures: 3

Tables: 2 + 1

References: 33

Article Summary Line: This work tested a large number of blood donors for anti-SARS-CoV-2 antibodies and used the data to improve a mathematical model of epidemic prediction for COVID-19.

Running Title: Blood donors SARS-CoV-2 IgG in epidemic monitoring

Keywords: SARS-CoV-2; COVID-19; Blood donor; Epidemic model; Seroprevalence.

Abstract

During the epidemics such as SARS-CoV-2, data from different sources can provide information on different aspects of the epidemic process. These may include records from serology-based epidemiologic surveys to compose a consistent scenario. This study assessed the seroprevalence of IgG anti-SARS-CoV-2 in 7,837 samples from blood donors in seven Brazilian cities from March to December 2020. Based on the results this work proposes a modification in a compartmental model that receives as inputs the reported number of COVID-19 cases and the serology results of blood donors and delivers estimates of hidden variables such as daily values of transmission rate and cumulative incidence rate of reported and non-reported cases. The study concludes that the information about cumulative incidence of the disease in the cities´ populations can be obtained simply by carrying out tests on samples collected from blood donors. The same methodology can be extended to surveillance of other infectious diseases.

Introduction

Coronavirus disease 2019 (COVID-19), caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has emerged in China in 2019 and spread out around the world in 2020. As of August 2021, over 200 million people were infected by SARS-CoV-2 and more than four million died (*1*). A significant proportion of infected individuals infected remain asymptomatic or present only mild symptoms (*2,3*). It is important to consider the role of all infected individuals in the maintenance of disease transmission, especially when asymptomatic/mildly symptomatic ones are often not tested or reported to public health authorities (*4*). Additionally, due to economic, political and social difficulties, availability of molecular tests for virus detection is often limited, particularly in developing countries (*5,6*). In this context, alternative measures

must be explored to generate reliable data that allow government decisions to contain viral spread.

A rapid immune response that culminates in the production of antibodies anti-SARS-CoV-2 in the first weeks after infection has been reported (*5-8*). Assessment of serological IgG anti-SARS-CoV-2 can be an important tool to assess the dynamics of virus transmission.

Some authors hypothesized that serosurveillance in blood donors can be important to monitor the evolution of SARS-CoV-2 infections (*9-11*). In this large longitudinal study, the records of reported COVID-19 cases and SARS-CoV-2 serology results of blood donors are used as inputs to a modified SEIR (Susceptible-Exposed-Infected-Recovered) epidemic model (*12*). The proposed model delivers daily estimates of relevant variables that usually stay hidden, including the infection transmission rate and the cumulative number of reported and non-reported cases of infections. The continuous change of the transmission rate is considered, making the estimated variables compatible with the shifting conditions of disease transmission. The monthly estimates of cumulative incidence provided by serology analysis of blood samples are used to estimate the proportion of reported and non-reported cases. The purpose of the model is to constitute a platform for the integration and interpretation of information coming from different sources.

This study also presents evidence supporting the possibility of using serology of blood samples collected in blood centers to estimate the accumulated incidence of the disease. The results refer to blood samples collected in seven blood centers. The proposed methodology provides a consistent picture of the evolution of COVID-19 in the

respective cities, describing the cumulative incidence, daily transmission rate and proportion between reported and non-reported cases.

Materials and methods

Study population

This study enrolled blood donors at seven Brazilian blood centers (Fundação HEMOMINAS) from March 1st to December 31, 2020. These blood centers account for approximately 60% of the blood collections performed by HEMOMINAS and are located in the cities of Belo Horizonte (BH), Governador Valadares (GV), Juiz de Fora (JF), Montes Claros (MC), Pouso Alegre (PA), Uberaba (UB) and Uberlândia (UD) in the Brazilian state of Minas Gerais. The number of samples to be included monthly in each selected blood center was calculated using the Epitools - Epidemiological Calculators (https://epitools.ausvet.com.au/) based on the prevalence of cases of COVID-19 reported by the municipal health secretaries. The number of samples analyzed was above the quantity defined by that calculation in all periods and centers. The study was approved by the Institutional Ethics Committee (CAAE 31087720.2.0000.5118). Data from all donors were collected from the records of each blood center.

Sample collection and testing

Samples collected for serological screening in the seven blood centers (n=7,854) were randomly selected, aliquoted and frozen at -80°C until IgG anti-SARS-CoV-2 testing was performed. The SARS-CoV-2 IgG kit (Abbott, Sligo, Ireland) was used in accordance with the manufacturer's manual for determination of IgG anti-nucleocapsid protein of SARS-CoV-2 positive and negative samples. Among the samples tested, 17

(0.2%) were excluded from the study because they corresponded to different donations from the same blood donors that were positive for IgG anti-SARS-CoV-2. Testing results in different donations from the same individual (n=17) always showed the same result. Only the first donation from each IgG anti-SARS-CoV-2 positive repeat donor was kept in the analyses.

Epidemic Model

Studies that present dynamic models of COVID-19 epidemics usually employ compartmental models with SEIR (Susceptible-Exposed-Infected-Removed) structure (*13*). This study proposes a SEIR model that employs the same compartments as defined in (*12*), with individuals being susceptible, exposed (those in the latent period of infection), reported infected (those that can propagate the virus and are reported in public health statistics), non-reported infected, and removed (those who have either recovered and become immune, at least temporarily, or have died). In addition to the number of individuals in each compartment, the model also defines the transmission rate, β, as a variable that changes with time. The time evolution of variables is described by differential equations. The model version presented here does not yet consider vaccination or the possibility of reinfection. While the inclusion of vaccination may be performed by moving the vaccinated individuals to the Removed compartment, the usage of the model for larger time horizons will require further studies concerning the loss of immunity both in recovered and in vaccinated individuals. Other issues that should be studied are the response of serology tests in detecting vaccine antibodies and the effect of antibody waning in test results. The proposed model is described in the supplemental material.

The model has two parameters that are mainly biologically determined: the average time an individual stay in the compartment of *exposed* before changing to *infected*, and the average duration of infection. Other parameters depend on both biological and social factors.

Most of the studies related to the dynamic modeling of COVID-19 epidemics consider either a constant or a piecewise constant β, which changes as governments enact or remove social distancing and other containment measures (*14*). However, the actual dynamics of COVID-19 epidemics varies in a faster way, due to the shifting response of populations to virus containment measures.

In this work, it is assumed that β has a daily value which is estimated by minimizing the error between the number of reported infected individuals delivered by the model and the one informed by public health services. This creates an implicit feedback loop which forces the model internal variables to adapt such that they mirror the corresponding real hidden variables of the epidemic process. A model with this capability of delivering estimates of hidden internal variables of a system is called a *state observer* (*13,15*).

After β is estimated, the remaining model parameter to be found is the fraction α of infected individuals that are detected by testing, becoming reported cases. This value is determined by comparing the accumulated number of reported cases with the accumulated incidence indicated by seroprevalence in blood samples.

The proposed methodology was assessed using data of apparent lethality (deaths divided by reported cases) of COVID-19. This was possible due to the testing policy

used in the state of Minas Gerais, which defined the eligibility of patients for being tested based on severity of symptoms. Those criteria were very restrictive in the first moment and were relaxed after the infrastructure for testing was expanded on July 2020. Therefore, α changed from a fixed value to another fixed value, leading to a change in the apparent lethality by the same factor of the change in α. As the data relative to deaths, reported cases and incidence in blood samples are mutually independent, the proposed model may be assessed by checking its simultaneous compatibility with those data in all cities. For this purpose, the apparent lethality is used to infer α, instead of using the proportion between the accumulated number of reported cases and incidence in blood samples.

Statistical analysis

The number of occurrences of each outcome and their frequencies were calculated for the categorical variables. Comparisons were made using Fisher's exact test. Medians and interquartile ranges (IQR) were calculated for the continuous variables and the comparisons were performed using two-sided Mann–Whitney tests. All tests and confidence intervals were calculated at the 95% confidence level. The proportion of positive IgG tests for each blood center was estimated by aggregating the number of tests and positive results of each month, removing repeated donors with positive tests already recorded in previous visits. Epitool (https://epitools.ausvet.com.au/) was used to calculate the unadjusted and test-adjusted seroprevalence for sensitivity (90%) and specificity (99%) (*16,17*), using Wilson's confidence interval for apparent prevalence and Blaker's confidence interval for true prevalence.

Results

IgG anti-SARS-CoV-2 seroprevalence in blood donors

This study included data from 7,837 blood donors at seven Brazilian blood centers from March 1st to December 31, 2020. The total number of samples included in the study represents 6.4% of blood donations carried out during that period in the selected centers. Serological testing for identification of IgG anti-SARS-CoV-2 revealed 441 (5.63%) positive blood donors throughout the period. When adjusted for sensitivity and specificity of the test, the overall rate of positivity was 5.20% (95% CI 4.65-5.80). Male donors had 1.35 (1.12 - 1.63) times the odds of being seropositive than female. The type of donor (first-time vs. repeat) did not present significant differences between the positive and negative groups for IgG anti-SARS-CoV-2. Age also did not present statistically significant differences, either across age groups (after adjusting for multiple hypotheses tests using the Holm correction) or when regressing the rate of positivity on age, using simple linear regression (Table 1). The evolving seroprevalence over the months of 2020 in each blood center and their geographic location in the State of Minas Gerais are represented in Figure 1. To most blood centers, the increase in seroprevalence rates was slower in the first months, accelerating from August, but becoming faster from October. The monthly data of each blood center is presented in the Appendix Table 1.

Modeling SARS-CoV-2 infection cases in general population

The seroprevalence rates of IgG anti-SARS-CoV-2 in blood donors were used to infer the proportion of infected general population in the cities of blood centers, according to the statistical model established. The parameter α is chosen such that the accumulated incidence rate delivered by the model, including the reported and non-reported cases,

fits the prevalence of COVID-19 in the blood donors in each blood center for each month.

Model assessment

The evolution of apparent lethality of COVID-19 (deaths divided by reported cases) in the cities that host blood centers (except UD) suggests that the proportion of infected individuals that are tested changes around day 122 (July 16th), increasing by nearly 70%, as shown in Figure 2. That moment coincides with the time when additional laboratories were integrated to the testing infrastructure for SARS-CoV-2 provided by the State government. In the case of UD, there was a relevant testing infrastructure provided by the municipality, in addition to the infrastructure provided by the State government that was available in all cities. Those data provide independent information about the relative changes in the value of α which can be used for assessing the model consistency.

This assessment was performed by running the model with α taking the same value in all cities except UD, assuming a fixed value that is increased by 70% on t=122 days (July 16th), remaining fixed on this new value from that date up to December 31st. The initial value of α which leads to the best fitting of the observed values of IgG rate of positivity in the blood centers was found, leading to the value of 0.18 as the proportion of reported cases up to t=122 days and of 0.31 for t >122 days for the cities BH, PA, MC, JF, GV and UB. In the case of UD, the values 0.37 for t <122 and 0.41 for t >122 were found.

Figure 3 shows the seroprevalences in the blood centers on each month, superimposed to the estimated curves of accumulated number of cases (reported and non-reported), as

predicted by the model in the respective cities. In five cities (BH, JF, MC, UB, UD), such parameter values resulted in reasonable matches with almost all data points. In two cities (GV, PA), the seroprevalence data in the last period (October, November, and December) were not well-adjusted to the model. This suggests that the pattern of variation of α in those cities could be different from the other ones. However, as those are small cities, the separate analysis of the change in the apparent lethality is not possible, which prevents the possibility of applying the same methodology for refining the estimates in those cases.

Discussion

In this study we evaluated the rate of blood donors positive to IgG anti-SARS-CoV-2 who donated blood in seven cities from Minas Gerais, Brazil during March to December 2020. The data was used to estimate the rate of infection in general population, which was then used within a dynamic model with SEIR structure.

The higher rate of positivity found in male individuals did not agree with the reported COVID-19 cases in Minas Gerais in the same period (49.2% for male). The higher proportion of positive tests in males suggests that the epidemiological profile of infection may change when more asymptomatic or mild COVID-19 individuals are tested, such as expected for blood donors. Rate of positivity associated with sex has been previously observed (*18*), but different works did not identify this association in blood donors (*19-21*) or in general population (*27*).

Concerning differences of positivity between age groups, no statistically significant difference was found in this study. This is a controversial issue, with some studies

reporting higher seroprevalence in younger individuals (*19,23*), while others indicate greater seroprevalence in older individuals (*24*) or not find significant associations between seroprevalence and age (*25*). A study carried out in 133 Brazilian cities identified that individuals aged 20-59 years are more likely to be infected, an age group that corresponds to most blood donors included in this study (*26*). Those differences between studies may be partly explained by cultural and populational issues, making it difficult to consolidate a general conclusion. Loss of statistical power due to corrections for testing multiple hypothesis may also play a role in the observed differences not achieving statistical significance, particularly if the effect size is moderate.

Seroprevalence in the blood centers showed the proportion of positive individuals increasing slowly in the first six months, with higher proportions of positivity from August onwards, with regional variations. In Minas Gerais an increase in COVID-19 cases occurred in June, peaking in August, decreasing slowly until October and then reaching the highest value in December 2020. Our results agree with this scenario, suggesting that seroprevalence rates in blood donors correlated with reported COVID-19 case rates. An important feature of the rate of positivity indicated by serological testing in the blood centers is that it is much greater than the prevalence that would be obtained by the accumulated number of reported cases of COVID-19. This difference was expected due to under-reporting. Notwithstanding, public communication about COVID-19 pandemics is commonly articulated based on reported cases, which strongly underestimates the actual spread of the disease. This indicates the convenience of using a model-based approach as proposed in this work, enabling the use of measured data for estimating hidden variables such as the total number of infected individuals.

Although all the cities evaluated had shown increased values of rate of positivity in December, GV showed the highest values. It is in consonance with the fact that there was an accumulated incidence per 100,000 habitants of COVID-19 in GV higher than that seen in Minas Gerais State and Brazil (GV-4,227.8; MG-2,270.1; BR-3,383.6, respectively) and the same was observed in relation to mortality (GV-143; MG-51.3; BR-88 deaths per 100,000 inhabitants, respectively).

Serial SARS-CoV-2 serological surveillance studies using blood donors is being implemented in several countries (*19,23,27*). These provide relevant results to complement population seroprevalence data (*19*) and valuable information for decision-making in countries where such data are not available. However, some issues should be considered, including the appropriate test to assess seroprevalence as well as the threshold for identifying positive and negative samples. It should be considered that the available automated serological tests were validated using samples from symptomatic COVID-19 patients with a confirmed diagnosis by RT-PCR (*28*). Results obtained in other studies using the same chemiluminescence test indicated a lower sensitivity to detect IgG anti-SARS-CoV-2 in newly infected individuals (29), which may affect the extrapolation of seroprevalence data to the population when using blood donors' samples (*11,30*). These data reinforce the importance of choosing serological assays with high sensitivity, specificity and durable antibody detection even months after infection (*31-32*).

Blood donor-based estimates of the seroprevalence of COVID-19 may deviate from the seroprevalence in the general population for several reasons, including the exclusion of populations who cannot donate, e.g. people younger than 16, older than 70, or residents

of nursing homes or prisons. The proportion of different groups (e.g. males and females, or different age groups) in the samples may differ from their respective proportions in population. Additionally, the recruitment and eligibility criteria for blood donations recommended by the Brazilian Ministry of Health during the COVID-19 pandemic exclude for 14 days from the date of attendance at the blood center those asymptomatic candidates that had contact with infected persons in the past 30 days. The recommendation also excludes for 30 days from the disappearance of symptoms those blood donors previously diagnosed with COVID-19 (*33*). These guidelines may result in decreased SARS-CoV-2 rate of positivity among donors.

Notwithstanding, the results of this study revealed that prevalence estimates obtained using the SEIR model, when compared to actual health system notification data, suggest that blood donor serosurveillance data can provide valuable information for monitoring the epidemic and evaluating the effectiveness of measures to fight the virus spread in the cities that have blood centers. This study also showed that the evolution of the epidemic can be considerably different from city to city, even considering cities within the same state in Brazil, suggesting that the application of the proposed SEIR model in other cities would require some strategy of periodic collection of blood samples for serological analysis, in a sufficient number of individuals spread across the population.

Some aspects of the proposed modeling approach should be highlighted. First, the procedure for estimating the time-varying transmission rate β for the SEIR model allows a reasonable automatic estimation of that parameter, in this way circumventing an important difficulty in COVID-19 modeling (*12,14*). As a by-product, this procedure also eliminates the difficulty usually encountered in determining adequate initial

conditions. In fact, the SEIR model, once endowed with the estimation procedure for β, becomes equivalent to a state observer (*13,15*), producing estimates of the model hidden variables that will approximate the real non-measured variables regardless of initial conditions, provided that the model parameters are reasonable approximations of the actual ones.

The estimated hidden variables may be quite useful in practice. For instance, β(t) provides information that is not contained in the reproduction number R_t, since β(t) does not vary with the number of recovered individuals, representing a better descriptor of social isolation intensity. Perhaps counterintuitively, the cumulative incidence estimate provided by the model can be considered more reliable than the monthly point estimates derived from raw data of serological analysis in blood centers, since the model performs a filtering of the random variation in data that results from sampling.

Concerning the assessment of the proposed model, it would be possible to choose different values for the α parameter for each city and for each month, according to the outcomes of serological tests in the respective blood centers. If this was done, the accumulated incidence of cases estimated by the model would be enforced to follow the trajectory of serology results, and this would not mean a confirmation of model validity. The procedure of model assessment adopted here used the same trajectories for α in six cities, getting the changes in α from an independent source. The consistency of the model outcomes with the serology results in most of data points, considering cities with rather different trajectories of the epidemics, provides corroboration of the proposed model.

Some limitations of this study should be mentioned. First, the stratification of the blood donors by sex or by age would allow the correction of the seroprevalence estimates according to the demographic composition of the general population, leading to more precise results. The observation of Figure 3, in which the prevalence in some cities presents a systematic tendency to remain below the values predicted by the model in the last three months also suggests that there may exist a relevant process of seroreversion, with IgG waning. The modeling of such a decay process may be important for the correct interpretation of the data in the last months of the experiment. Finally, it should be mentioned that some of the blood centers considered in this study are relatively small (PA and UB), which increases the uncertainty associated to the data collected in those centers not only by reducing the size of the sample, but also by reducing the robustness to skewed data.

Conclusion

The results suggest that blood centers could be incorporated to the COVID-19 surveillance systems with the role of regularly providing quantitative estimates of prevalence of the disease in population. For this purpose, an epidemic model with state observer property (which performs a track of some measured variable, producing outputs that converge to the system hidden variables) should be used. In this work, a specific SEIR epidemic model that performs the adjustment of the transmission rate β such that the model tracks the measured number of reported cases of COVID-19 is proposed. The seroprevalence data collected in blood centers are employed to perform the adjustment of the proportion of reported cases considered in the model. This model was shown to provide consistent estimates of relevant variables that otherwise would not be accessible, in this way supporting a well-informed decision-making process. The

proposed methodology can be adapted to the surveillance of other infectious diseases, using other kinds of input information from sentinel surveillance systems combined with serosurveillance data gathered in blood centers.

Acknowledgments

We thank all of the professionals who contributed to the enrollment of blood donors in the study. We also thank Secretaria Estadual de Saúde de Minas Gerais (SES/MG) and Fundação Hemominas for the financial support. EFB-S, RHCT and IRO are fellows from CNPq.

Author Bio

DGC has a degree in Biological Sciences, with a Masters and PhD in Biochemistry and Immunology. DGC is currently head of the Research Service at Fundação Hemominas, where he works as a researcher in the areas of Hematology and Infectious Diseases.

References

1. World Health Organization. WHO Coronavirus (COVID-19) Dashboard. 2021; https://covid19.who.int/

2. Huang L, Zhang X, Zhang X, Wei Z, Zhang L, Xu J, et al.; Rapid asymptomatic transmission of COVID-19 during the incubation period demonstrating strong infectivity in a cluster of youngsters aged 16-23 years outside Wuhan and characteristics of young patients with COVID-19: A prospective contact-tracing study. J Infect. 2020;80:e1-e13.

3. Lai CC, Liu YH, Wang CY, Wang YH, Hsueh SC, Yen MY, et al.; Asymptomatic carrier state, acute respiratory disease, and pneumonia due to severe acute respiratory

syndrome coronavirus 2 (SARS-CoV-2): Facts and myths. J Microbiol Infect. 2020;53:404-12.

4. Buitrago-Garcia D, Egli-Gany D, Counotte MJ, Hossmann S, Imeri H, Ipekci AM, et al.; Occurrence and transmission potential of asymptomatic and presymptomatic SARS-CoV-2 infections: A living systematic review and meta-analysis. PLoS Med. 2020;17:e1003346.

5. Zhou W, Xu X, Chang Z, Wang H, Zhong X, Tong X, et al.; The dynamic changes of serum IgM and IgG against SARS-CoV-2 in patients with COVID-19. J Med Virol. 2021;93:924-33.

6. Sun B, Feng Y, Mo X, Zheng P, Wang Q, Li P, et al.; Kinetics of SARS-CoV-2 specific IgM and IgG responses in COVID-19 patients. Emerg Microbes Infect. 2020;9:940-8.

7. Lee YL, Liao CH, Liu PY, Cheng CY, Chung MY, Liu CE, et al.; Dynamics of anti-SARS-Cov-2 IgM and IgG antibodies among COVID-19 patients. J Infect. 2020;81:e55-e58.

8. Wellinghausen N, Plonné D, Voss M, Ivanova R, Frodl R, Deininger S. SARS-CoV-2-IgG response is different in COVID-19 outpatients and asymptomatic contact persons. J Clin Virol. 2020;130:104542.

9. Busch MP, Stone M. Serosurveillance for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) incidence using global blood donor populations. Clin Infect Dis. 2021;72:254-6.

10. Kadkhoda K. Letter to the editor: COVID-19: how accurate are seroprevalence studies? Euro Surveill. 2020;25:2001374.

426 11. Buss LF, Prete CA Jr, Abrahim CMM, Mendrone A Jr, Salomon T, de Almeida-
427 Neto C, et al.; Three-quarters attack rate of SARS-CoV-2 in the Brazilian Amazon
428 during a largely unmitigated epidemic. Science. 2021;371:288-92.
429 12. Li R, Pei S, Chen B, Song Y, Zhang T, Yang W, et al.; Substantial undocumented
430 infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2).
431 Science. 2020;368:489-93.
432 13. Takahashi RHC, Peres PLD. Unknown input observers: a unifying approach. Eur J
433 Control. 1999;5:261-75.
434 14. Centers for Disease Control and Prevention. COVID-19 Forecasts: Cases. 2021;
435 https://www.cdc.gov/coronavirus/2019-ncov/science/forecasting/forecasts-cases.html
436 15. Luenberger DG. Observing the state of a linear system. IEEE Transactions on
437 Military Electronics. 1964;8:74-80.
438 16. Bryan A, Pepper G, Wener MH, Fink SL, Morishima C, Chaudhary A, et al.;
439 Performance characteristics of the Abbott Architect SARS-CoV-2 IgG Assay and
440 seroprevalence in Boise, Idaho. J Clin Microbiol. 2020;58:e00941-20.
441 17. Ng DL, Goldgof GM, Shy BR, Levine AG, Balcerek J, Bapat SP, et al.; SARS-
442 CoV-2 seroprevalence and neutralizing activity in donor and patient blood. Nat
443 Commun. 2020;11:4698.
444 18. Vassallo RR, Dumont LJ, Bravo MD, Hazegh K, Kamel H. Progression and
445 predictors of SARS-CoV-2 antibody seroreactivity in US blood donors. Transfus Med
446 Rev. 2021;S0887-7963(21)00028-6.
447 19. Stone M, Di Germanio C, Wright DJ, Sulaeman H, Dave H, Fink RV, et al.; Use of
448 U.S. blood donors for national serosurveillance of SARS-CoV-2 antibodies: basis for an
449 expanded national donor serosurveillance program. Clin Infect Dis. 2021;ciab537.

20. Slot E, Hogema BM, Reusken CBEM, Reimerink JH, Molier M, Karregat JHM, et al.; Low SARS-CoV-2 seroprevalence in blood donors in the early COVID-19 epidemic in the Netherlands. Nat Commun. 2020;11:5744.

21. Chang L, Hou W, Zhao L, Zhang Y, Wang Y, Wu L, et al.; The prevalence of antibodies to SARS-CoV-2 among blood donors in China. Nat Commun. 2021;12:1383.

22. Peckham H, de Gruijter NM, Raine C, Radziszewska A, Ciurtin C, Wedderburn LR, et al.; Male sex identified by global COVID-19 meta-analysis as a risk factor for death and ITU admission. Nat Commun. 2020; 11:6317.

23. Public Health England. Sero-surveillance of COVID-19. 2021; https://www.gov.uk/government/publications/national-covid-19-surveillance-reports/sero-surveillance-of-covid-19

24. Pagani G, Conti F, Giacomelli A, Bernacchia D, Rondanin R, Prina A, et al. Seroprevalence of SARS-CoV-2 significantly varies with age: Preliminary results from a mass population screening. J Infect. 2020;81:e10-e12.

25. Nwosu K, Fokam J, Wanda F, Mama L, Orel E, Ray N, et al. SARS-CoV-2 antibody seroprevalence and associated risk factors in an urban district in Cameroon. Nat Commun. 2021;12:5851.

26. Hallal PC, Hartwig FP, Horta BL, Silveira MF, Struchiner CJ, Vidaletti LP, et al. SARS-CoV-2 antibody prevalence in Brazil: results from two successive nationwide serological household surveys. Lancet Glob Health. 2020;8:e1390-e1398.

27. Erikstrup C, Hother CE, Pedersen OBV, Molbak K, Skov RL, Holm DK, et al.; Estimation of SARS-CoV-2 infection fatality rate by real-time antibody screening of blood donors. Clin Infect Dis. 2020;72:249–53.

28. Eyre DW, Lumley SF, O'Donnell D, Stoesser NE, Matthews PC, Howarth A, et al. Stringent thresholds for SARS-CoV-2 IgG assays result in under-detection of cases reporting loss of taste/smell. medRxiv 2020; 20159038.

29. Hamilton, F., Muir, P., Attwood, M., Vipond, A., Hopes, R., Moran, E., et al. Kinetics and performance of the Abbott architect SARS-CoV-2 IgG antibody assay. J Infect 2020;81:e7–e9.

30. Sabino EC, Buss LF, Carvalho MPS, Prete CA Jr, Crispim MAE, Fraiji NA, et al. Resurgence of COVID-19 in Manaus, Brazil, despite high seroprevalence. Lancet 2021;397:452-5.

31. Peluso MJ, Takahashi S, Hakim J, Kelly JD, Torres L, Iyer NS, et al. SARS-CoV-2 antibody magnitude and detectability are driven by disease severity, timing, and assay. Sci Adv 2021;7:eabh3409.

32. Stone M, Grebe E, Sulaeman H, Di Germano C, Dave H, Kelly K, et al. Evaluation of commercially available high-throughput SARS-CoV-2 serological assays for serosurveillance and related applications. medRxiv 2021;09.04.21262414.

33. Nota Técnica nº 13/2020-CGSH/DAET/SAES/MS. Atualização dos critérios técnicos contidos na Nota Técnica nº 5/2020-CGSH/DAET/SAES/MS para triagem clínica dos candidatos à doação de sangue relacionados ao risco de infecção pelo SARS-CoV-2 (vírus causador da COVID-19). http://antigo.anvisa.gov.br/documents/2857848/5624592/SEI_MS+-+0014052636+-+Nota+T%C3%A9cnica+13.pdf/eb3aad9b-2ddb-4c15-b979-8aec2a6e331b. Accessed December 03, 2021.

Address for correspondence**:** Daniel Gonçalves Chaves, Fundação Centro de Hematologia e Hemoterapia de Minas Gerais (Fundação HEMOMINAS), Alameda Ezequiel Dias, 321, Belo Horizonte, Minas Gerais 30.130-110, Brazil.

E-mail: daniel.chaves@hemominas.mg.gov.br

Phone: 55-31-37684587

Fax: 55-31-37682002

Table 1. SARS-CoV-2 IgG seroprevalence in blood donors according to donor's characteristics from March to December 2020.

Characteristics	Total samples	Positive samples	Unadjusted seroprevalence, % (95% CI)	Test-adjusted seroprevalence*, % (95% CI)	OR (95% CI)
Total	7,837	441	5.6 (5.1-6.2)	5.2 (4.7-5.8)	
Gender					
Males	4,284	273	6.4 (5.7-7.1)	6.0 (5.3-6.9)	1.4
Females	3,553	168	4.7 (4.1-5.5)	4.2 (3.5-5.0)	(1.1-1.6)
Age (range)					
16-30	2,895	153	5.3 (4.5-6.2)	4.8 (4.0-5.8)	
31-40	2,330	135	5.8 (4.9-6.8)	5.4 (4.4-6.5)	
41-50	1,701	115	6.8 (5.7-8.1)	6.5 (5.3-7.9)	
51-60	829	32	3.9 (2.8-5.4)	3.2 (2.0-4.9)	
61-70	82	6	7.3 (3.4-15.1)	7.1 (2.7-15.8)	
Blood donors type					
First-time donors	1,483	72	4.9 (3.9-6.1)	4.3 (3.2-5.7)	0.8
Repeat donors	6,353	369	5.8 (5.3-6.4)	5.4 (4.8-6.1)	(0.7-1.1)

*Considering sensitivity of 90% and specificity of 99% (analysis performed by Epitool, using Wilson's confidence interval for apparent rate of positivity and Blaker's confidence interval for true rate of positivity).

515 **Table 2.** SARS-CoV-2 IgG seroprevalence in blood donors from cities of Minas Gerais,
516 Brazil

	Local						
IgG	PA	UB	JF	BH	MC	GV	UD
Reactive	22	37	42	108	36	81	115
Non-reactive	839	827	931	1785	572	1044	1399
Total	861	864	973	1893	608	1125	1514
Donor rate of positivity (%)							
Unadjusted (95% CI)	2.6 (1.7-3.8)	4.3 (3.1-5.9)	4.3 (3.2-5.8)	5.7 (4.8-6.8)	5.9 (4.3-8.1)	7.2 (5.8-8.9)	7.6 (6.4-9.0)
Test-adjusted (95% CI)	1.8 (0.8-3.2)	3.7 (2.4-5.5)	3.7 (2.5-5.4)	5.3 (4.2-6.6)	5.5 (3.7-8.00)	7.0 (5.4-8.8)	7.4 (6.0-9.0)
City population (15-69 y)*	96,803	220,925	385,568	1,842,210	262,875	187,691	451,581

517 PA: Pouso Alegre; UB: Uberaba; JF: Juiz de Fora; BH: Belo Horizonte; MC: Montes
518 Claros; GV: Governador Valadares; UD: Uberlândia. *Census 2010, IBGE.

519

520 Figure 1. Legend.

521 Temporal evolving of cumulative SARS-CoV-2 seroprevalence in the blood centers
522 from Fundação HEMOMINAS. Individuals eligible to blood donate were tested to IgG
523 anti-SARS-CoV-2 from March to December 2020. The figure shows the data in the
524 indicated months.

525

526 Figure 2. Legend.

527 Apparent lethality of COVID-19 in six cities (BH, PA, MC, JF, GV, UB) from April to
528 December 2020. Up to day 60 (shaded area), was the beginning of the epidemics in

Minas Gerais, with few cases, and the testing infrastructure was still being organized. The apparent lethality from day 60 to day 120 is nearly 5.2%, and from day 120 to 290 it is nearly 3.0%. This change corresponds to an increase of nearly 70% in the value of α (proportion of infected individuals that are reported), assuming that the actual lethality has not changed. Data coming from the city of Uberlândia was not included in this estimation due to local legislation regulating the testing, which resulted in a much larger proportion of people being tested in that municipality than in the other cities.

Figure 3. Legend.

Proportion of positive IgG anti-SARS-CoV-2 blood donors in each blood center by the last day of each month (blue dots, with vertical segments showing the 95% confidence intervals). Black squares indicate the official cumulative prevalence of reported cases for each city, and the red line represents model estimates of the number of infected individuals (including reported and non-reported cases) in each city, as a proportion of the city population. Vertical dashed lines indicate national holidays.

Appendix Table. SARS-CoV-2 IgG seroprevalence in blood donors according to cities of Minas Gerais

Month/2020	City																		
	PA			UB			JF			BH			MC			GV			UD		
	Serology results and prevalence of IgG anti-SARS-CoV-2																				
	Neg	Pos	%	Neg	Pos	%	Neg	Pos	%	Neg	Pos	%	Neg	Pos	%	Neg	Pos	%	Neg	Pos	%
March	44	0	0.0	44	0	0.0	44	0	0.0	44	0	0.0	44	0	0.0	44	0	0.0	43	1	2.3
April	44	0	0.0	42	2	4.6	44	0	0.0	44	0	0.0	44	0	0.0	44	0	0.0	44	0	0.0
May	58	1	1.7	60	0	0.0	59	1	1.7	59	1	1.7	59	1	1.7	60	0	0.0	60	0	0.0
June	65	0	0.0	64	1	1.5	64	0	0.0	156	2	1.3	50	0	0.0	63	2	3.1	70	0	0.0
July	67	3	4.3	67	3	4.3	68	2	2.9	186	5	2.6	50	0	0.0	66	4	5.7	67	2	2.9
August	76	1	1.3	51	1	1.9	83	4	4.6	189	12	6.0	60	2	3.2	97	4	4.0	64	2	3.0
September	66	3	4.4	66	4	5.7	76	1	1.3	204	8	3.8	47	4	7.8	77	12	13.5	136	12	8.1
October	95	3	3.1	99	3	2.9	111	4	3.5	157	9	5.4	53	8	13.1	126	11	8.0	219	30	12.1
November	136	4	2.9	138	6	4.2	160	9	5.3	377	36	8.7	62	9	12.7	192	18	8.6	324	38	10.5
December	188	7	3.6	196	17	8.0	222	21	8.6	369	35	8.7	103	12	10.4	274	30	9.9	372	30	7.5
Total	839	22	2.6	827	37	4.3	931	42	4.3	1,785	108	5.7	572	36	5.9	1,043	81	7.2	1,399	115	7.6

PA: Pouso Alegre; UB: Uberaba; JF: Juiz de Fora; BH: Belo Horizonte; MC: Montes Claros; GV: Governador Valadares; UD: Uberlândia.

Seroprevalence of IgG anti-SARS-CoV-2 in blood donors as an important measure to monitor the COVID-19 epidemic progress

– Supplemental Material –

State observer for SEIR models with time-varying transmission rate

1 Epidemics dynamic model

Most studies that present dynamic models of Covid-19 epidemics employ compartmental models with SEIR (Susceptible-Exposed-Infected-Removed) structure. This model structure is a variation of the traditional SIR (Susceptible-Infected-Removed) model, with the inclusion of a compartment of Exposed individuals, which accounts for the latent period of the infection. A key parameter in those models is the *transmission rate*, β, which aggregates the effects of some social behaviors in a population such as the mean number of interpersonal contacts of individuals, the strength of protection measures in contact situations (for instance, usage of masks, physical distancing during a contact, and others) and the selective isolation of individuals with symptoms, and also the relevant biological features that determine the ability of the virus to be transmitted when a contact occurs (for instance the mean exhaled viral load, the viral pathogenic mechanisms, and others).

The present study employs a model that follows the SEIR structure:

$$\begin{aligned}
\frac{dS}{dt} &= -\frac{\beta}{N} S\left(I^r + I^n\right) \\
\frac{dE}{dt} &= \frac{\beta}{N} S\left(I^r + I^n\right) - \frac{1}{Z} E \\
\frac{dI^r}{dt} &= \frac{\alpha}{Z} E - \frac{1}{D} I^r \\
\frac{dI^n}{dt} &= \frac{(1-\alpha)}{Z} E - \frac{1}{D} I^n \\
\frac{dR}{dt} &= \frac{1}{D}(I^r + I^n)
\end{aligned} \tag{1}$$

This model is similar to the one presented in [Li, 2020]. In this model, the compartment $S(t)$ represents the number of susceptible individuals in population, $E(t)$ represents the number of exposed individuals (the individuals

which are in the latent period of infection, in which they are not able to propagate the virus yet), $I^r(t)$ represents the number of infected individuals (the ones which will propagate the virus if they contact a susceptible individual) that have been reported in public health statistics, and $I^n(t)$ represents the number of infected individuals that have not been reported. The compartment $R(t)$ represents removed individuals (the individuals that have recovered from the disease and consequently have become immune, at least temporarily, or which have died). In this equation, N represents the initial number of individuals in the population. In addition, the following equation performs the computation of the cumulative number of reported infected individuals, represented by C^r:

$$\frac{dC^r}{dt} = \alpha \frac{E}{Z} \tag{2}$$

The equation (1) has some parameters that are mainly biologically determined, $Z = 3.69$ (the average time an individual stays in the compartment of exposed individuals before becoming infected), and $D = 7.0$ (the average duration of infection).

Most of the studies that have been published concerning the dynamic modeling of COVID-19 epidemics either consider a constant value of β or a piecewise constant value, which changes as social distancing measures are changed by governments. However, the actual dynamics of COVID-19 epidemics varies in a much faster way, due to the varying response of populations to virus containment measures – as can be inferred from the growth of infection rates just after holidays or other dates of social events. In addition to β, the α parameter also depends on social factors, representing the fraction of infected individuals that are detected by testing and become reported cases.

It should be clear that, for performing a simulation of an actual scenario, estimates for the values of β and α are necessary, as well as estimates for the initial values of all model variables, $S(0), E(0), I^r(0), I^n(0), R(0)$. The issues related to the assignment of values to those parameters are discussed next, jointly with the introduction of a modification in the model (1) that transforms it in a *state observer*, endowing the model with the capability of auto-adapting to parameter changes while performing a fitting of the accumulated number of reported cases, C^r, represented within the model, to the corresponding number reported by the public health services.

2 State observer for the epidemics dynamic model

State observers are important tools that have been developed for monitoring the internal variables of dynamic systems, usually for the purpose of assisting the system control. There is a large number of reported applications of those tools, mainly in the monitoring and control of complex technological systems

such as in aerospace artifacts, chemical industry, and so forth. In this section, a general discussion of the idea of state observers is presented first. Then, the specific state observer that was developed in this study for the monitoring of epidemic processes is introduced.

Consider a dynamic system that is described by the following system of differential equations:

$$\begin{aligned} \dot{\mathbf{x}} &= f(\mathbf{x}) \\ \mathbf{y} &= g(\mathbf{x}) \end{aligned} \tag{3}$$

In this system, $f(\cdot)$ represents the system dynamic function, $g(\cdot)$ represents the output measurement function, the vector $\mathbf{x} \in \mathbb{R}^n$ represents the system internal variables (the system states), and the vector $\mathbf{y} \in \mathbb{R}^m$ represents the vector of signals that are directly measured on the system. State observers are models that represent dynamic systems that are intended to provide estimates of the system internal signals. It is assumed that the exact representation of the system, as described in (3), is not available to the analyst.

A state observer for system (3) can be represented as:

$$\begin{aligned} \dot{\hat{\mathbf{x}}} &= \hat{f}(\hat{\mathbf{x}}, \hat{\mathbf{e}}) \\ \hat{\mathbf{y}} &= \hat{g}(\hat{\mathbf{x}}) \\ \hat{\mathbf{e}} &= \hat{\mathbf{y}} - \mathbf{y} \end{aligned} \tag{4}$$

In this equation, the functions $\hat{f}(\cdot)$ and $\hat{g}(\cdot)$ are approximated representations of functions $f(\cdot)$ and $g(\cdot)$, $\hat{\mathbf{x}}$ represents the vector of estimates of the system internal variables, $\hat{\mathbf{y}}$ represents the estimate of output measurement vector, and $\mathbf{e}$ is the error between the estimated output vector $\hat{\mathbf{y}}$ and the actual measurement vector $\mathbf{y}$. The working principle of the state observers is that the error signal $\mathbf{e}$ is fed back into the observer, with this feedback loop designed such that the difference between the system state vector $\mathbf{x}$ and the estimate $\hat{\mathbf{x}}$ of the state vector provided by the observer converges to zero. After this convergence, the state observer provides estimates of all system signals, including the system internal signals that are not measured directly. The exact convergence may be achieved when $\hat{f} = f$ and $\hat{g} = g$. When the differences between the model $(\hat{f}, \hat{g})$ represented in the observer and the actual system dynamics (f, g) are small, the observer state vector $\hat{\mathbf{x}}$ is expected to represent a good estimate of the system internal variables $\mathbf{x}$.

Most of the state observers that have been studied up to now employ an additive feedback of the measurement error, which makes the observer dynamic equation become:

$$\dot{\hat{\mathbf{x}}} = \hat{f}(\hat{\mathbf{x}}) + \mathbf{K}\hat{\mathbf{e}} \tag{5}$$

in which $\mathbf{K} \in \mathbb{R}^{n \times m}$ is a matrix of constant feedback coefficients.

The feedback structure (5) has been employed in some published works that propose state observers for SIR-like epidemic models [Degue and Ny, 2019,

Iggidr and Souza, 2019]. A main drawback of those approaches is that they depend on the function $\hat{f}(\cdot)$ being a reasonable approximation of the function $f(\cdot)$ in the actual system. As discussed in section 1, in the case of COVID-19 the parameter β presents strong and fast variations, which makes the usage of those observers very difficult, since they could be used for very short time horizons in which estimates of β could be considered reasonable approximations of the actual disease transmission rate. In addition, those observers would have no role in the estimation of β values, thus failing to provide the estimate of the variable that would likely be the most important to be estimated.

In this work, a new structure of state observer for SIR-like models in which the infection transmission rate β continuously varies along the time is presented. In the proposed technique, the actual accumulated number of COVID-19 cases, C^r, is measured as reported by public health services, and the error between this number and the number $\hat{C}^r$ estimated by the observer is calculated. This error is fed back to the estimator in a rather unusual way. First, it is assumed that β is a time-varying parameter, which becomes represented by $\beta(t)$. An optimization procedure is run, searching for a time-varying estimate $\hat{\beta}(t)$ which minimizes that error on each day. When the optimal sequence $\hat{\beta}^*(t)$ is found, the estimates of the other system internal variables appear as by-products of the optimization procedure that result from the simulation of the model with that optimal values of the transmission rate. More specifically, the following cost function is defined:

$$J(\hat{\beta}, k) = \sum_{i=k-d}^{k+d} (\log(C^r(i)) - \log(\hat{C}^r(i, \hat{\beta})))^2 \tag{6}$$

with $C^r(i)$ representing the accumulated number of actual reported cases in the city on day i and $\hat{C}^r(i, \hat{\beta})$ representing the accumulated number of reported cases calculated by the model from time $t = 1$ to $t = k$, using $\beta = \hat{\beta}$ in a time window of length $2d + 1$ centered in $t = k$. The estimated values of the daily disease transmission rate $\beta^*(t)$ are given by:

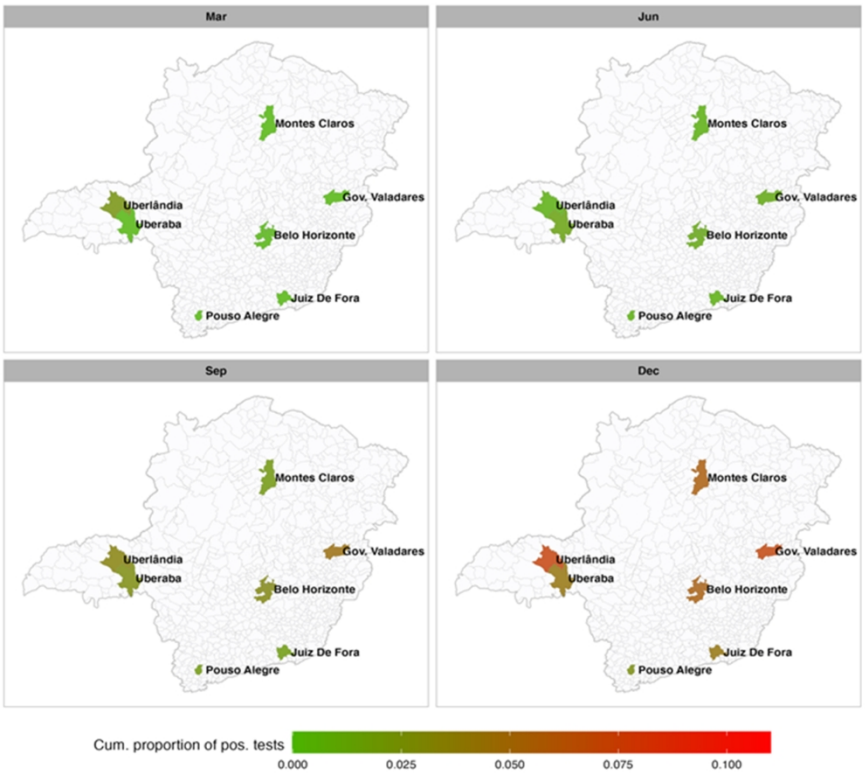
$$\begin{aligned} &\beta^*(t) = \arg\min_{\beta} \ J(\beta, t) \\ &\text{subject to: } \{(1), (2) \end{aligned} \tag{7}$$

References

[Degue and Ny, 2019] Degue, K. H. and Ny, J. L. (2019). Estimation and outbreak detection with interval observers for uncertain discrete-time SEIR epidemic models. *International Journal of Control*.

[Iggidr and Souza, 2019] Iggidr, A. and Souza, M. O. (2019). State estimators for some epidemiological systems. *Journal of Mathematical Biology*, 78:225–256.

[Li, 2020] Li, R. (2020). Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). *Science*.

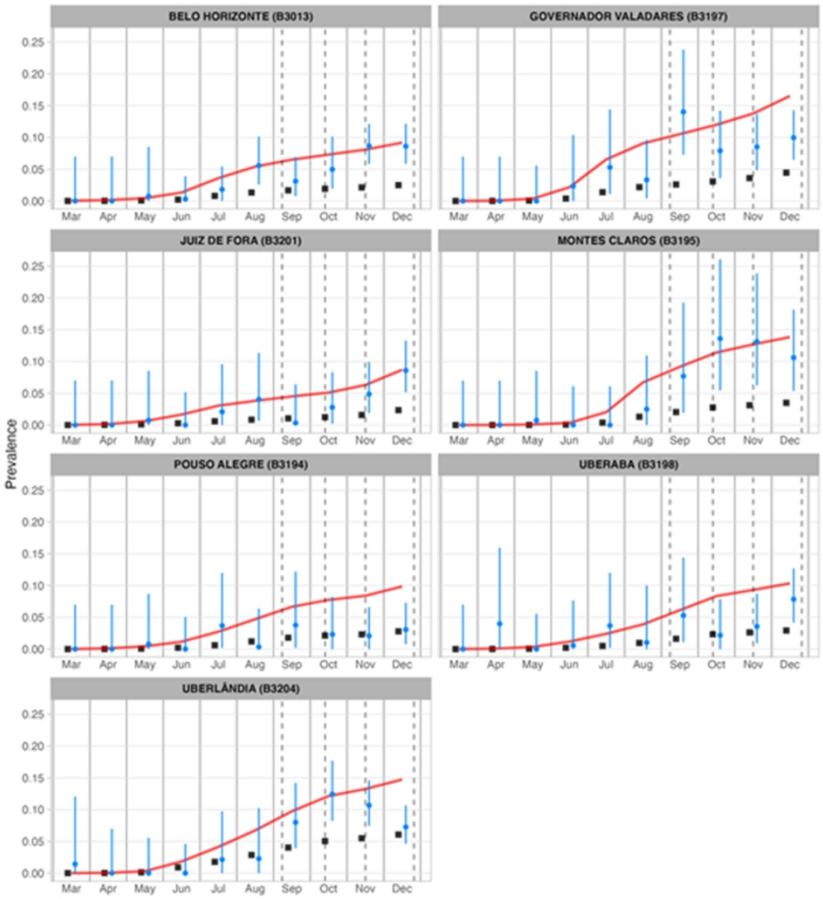


Temporal evolving of cumulative SARS-CoV-2 seroprevalence in the blood centers from Fundação HEMOMINAS. Individuals eligible to blood donate were tested to IgG anti-SARS-CoV-2 from March to December 2020. The figure shows the data in the indicated months.

215x200mm (300 x 300 DPI)

Apparent lethality of COVID-19 in six cities (BH, PA, MC, JF, GV, UB). Each interval between vertical green lines corresponds to a month, from April to December 2020. Up to day 60 (shaded area), was the beginning of the epidemics in Minas Gerais, with few cases, and the testing infrastructure was still being organized. The apparent lethality from day 60 to day 120 is nearly 5.2%, and from day 120 to 290 it is nearly 3.0%. This change corresponds to an increase of nearly 70% in the value of □ (proportion of infected individuals that are reported), assuming that the actual lethality has not changed. Data coming from the city of Uberlândia was not included in this estimation due to local legislation regulating the testing, which resulted in a much larger proportion of people being tested in that municipality than in the other cities.

209x106mm (300 x 300 DPI)



Proportion of positive IgG anti-SARS-CoV-2 blood donors in each blood center by the last day of each month (blue dots, with vertical segments showing the 95% confidence intervals). Black squares indicate the official cumulative prevalence of reported cases for each city, and the red line represents model estimates of the number of infected individuals (including reported and non-reported cases) in each city, as a proportion of the city population. Vertical dashed lines indicate national holidays.

174x192mm (300 x 300 DPI)