

## Delay in death reporting affects timely monitoring and modeling of the COVID-19 pandemic

O atraso na notificação de óbitos afeta o monitoramento oportuno e modelagem da pandemia da COVID-19

La demora en la notificación de muertes afecta el seguimiento puntual y el modelaje de la pandemia de COVID-19

Carolina Abreu de Carvalho <sup>1</sup>  
Vitória Abreu de Carvalho <sup>1</sup>  
Marcos Adriano Garcia Campos <sup>1</sup>  
Bruno Luciano Carneiro Alves de Oliveira <sup>1</sup>  
Eduardo Moraes Diniz <sup>1</sup>  
Alcione Miranda dos Santos <sup>1</sup>  
Bruno Feres de Souza <sup>1</sup>  
Antônio Augusto Moura da Silva <sup>1</sup>

doi: 10.1590/0102-311X00292320

### Abstract

*This study describes the COVID-19 death reporting delay in the city of São Luís, Maranhão State, Brazil, and shows its impact on timely monitoring and modeling of the COVID-19 pandemic, while seeking to ascertain how nowcasting can improve death reporting delay. We analyzed COVID-19 death data reported daily in the Epidemiological Bulletin of the State Health Secretariat of Maranhão and calculated the reporting delay from March 23 to August 29, 2020. A semi-mechanistic Bayesian hierarchical model was fitted to illustrate the impact of death reporting delay and test the effectiveness of a Bayesian Nowcasting in improving data quality. Only 17.8% of deaths were reported without delay or the day after, while 40.5% were reported more than 30 days late. Following an initial underestimation due to reporting delay, 644 deaths were reported from June 7 to August 29, although only 116 deaths occurred during this period. Using the Bayesian nowcasting technique partially improved the quality of mortality data during the peak of the pandemic, providing estimates that better matched the observed scenario in the city, becoming unusable nearly two months after the peak. As delay in death reporting can directly interfere with assertive and timely decision-making regarding the COVID-19 pandemic, the Brazilian epidemiological surveillance system must be urgently revised and notifying the date of death must be mandatory. Nowcasting has proven somewhat effective in improving the quality of mortality data, but only at the peak of the pandemic.*

COVID-19; SARS-CoV-2; Mortality Registries; Data Accuracy

### Correspondence

M. A. G. Campos  
Programa de Pós-graduação em Saúde Coletiva, Universidade Federal do Maranhão.  
Rua Barão de Itapari 155, São Luís, MA 65020-070, Brasil.  
marcos.adrianogc@gmail.com

<sup>1</sup> Universidade Federal do Maranhão, São Luís, Brasil.



## Introduction

Surveillance of mortality data is a key tool for public health, as it allows to monitor the dynamics and impact of health-related events <sup>1</sup>, gaining greater relevance in contexts of pandemics and new diseases. But for mortality data to fulfill its role in helping the follow-up of an outbreak, the notification of deaths must be updated <sup>2</sup>.

Since the onset of COVID-19 transmission in Brazil, the country has shown a notorious difficulty in identifying new cases. Few tests were performed, with the case reporting rate being estimated at only 9.2% <sup>2,3,4,5</sup>. A seroprevalence study conducted in São Luís, State of Maranhão, showed that only 3.4% of infections were reported <sup>6</sup>. Another issue regarding data on the number of cases is the delay in case reporting <sup>7</sup>. Modelling of infectious diseases has greater accuracy when done using case reporting based on date of symptom onset-data that is rarely available. Moreover, most case notifications were based on rapid antibody tests (lateral flow immunoassays) rather than molecular ones, with only 38% of cases being identified by RT-PCR (reverse transcription polymerase chain reaction) in Brazil. This reflects a low capacity to timely diagnose active cases of SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2) <sup>8</sup>.

High underreporting, unavailability of separate RT-PCR-based case reporting, and lack of data on date of symptom onset hinders COVID-19 modelling based on case reporting. By using cases in modelling, we are dealing with an outdated picture of the disease, looking at the transmission dynamics via the rear-view mirror. The use of mortality data attributed to COVID-19 tends, therefore, to be a more accurate indicator for monitoring the pandemic, since the underreporting of deaths is expected to be lower than that of cases <sup>9,10</sup>.

But mortality data is not without its problems. An analysis conducted by the COVID-19 BR Observatory estimated that 61% of deaths in Brazil took more than 10 days to be reported, only 3% were reported one day after occurring, and very rarely on the date of death (0.17%) <sup>11</sup>. Death reporting delay in Brazil is therefore high and the notified data reflect an already outdated scenario. Some studies point to RT-PCR shortage, death attributed to other causes with similar clinical manifestations, and the occurrence of false negatives due to quality control problems in nasal swab collection for testing as possible causes for COVID-19 death underreporting <sup>2</sup>. Due to these issues one must be careful and cautious when using reported data from information systems or epidemiological bulletins to analyze the pandemic scenario.

One way to improve the quality of this information and allow a less biased use of the reported data is nowcasting. This approach seeks to estimate, at a given point in time, the number of events that have occurred but have not yet been reported <sup>12</sup>. It generates a distribution of the reporting delay from observations where the occurrence and reporting date of the event of interest are known. Given this distribution and the number of events reported at a given time, one can infer the actual number of events that occurred. The result is a pandemic curve closer to the current state of the outbreak <sup>13</sup>.

Besides Brazil, delays in death reporting have also been described in high-income countries such as Sweden, Germany and the United States <sup>13,14</sup>. Seeking ways to minimize delay in official data reporting is relevant in places where reporting information is of poor quality, to more reliably follow the pandemic dynamics and implement public health measures. Thus, this study describes the COVID-19 death reporting delay in the city of São Luís and shows its impact on timely monitoring and modelling the COVID-19 pandemic, while seeking to ascertain how nowcasting can improve death reporting delay.

## Methods

This is a descriptive study conducted using data on COVID-19 death reported daily in the Epidemiological Bulletin of the State Health Secretariat of Maranhão (SES-MA in Portuguese), Brazil. Epidemiologica Bulletin contains, among other data, variables such as city of residence, date of death and date of notification. Inclusion criteria consisted of deaths of residents in São Luís city reported from March 29 to December 24, 2020. All data are anonymized and publicly available at SES-MA

official website, thus its use complies with ethical precepts and does not require approval by an ethics committee.

São Luís is the capital of the State of Maranhão and has an estimated population of 1,101,884 inhabitants, according to the Brazilian Institute of Geography and Statistics (IBGE. <http://www.ibge.gov.br>) in 2019. By January 25, 2021 it had 27,738 confirmed cases and 1,327 deaths attributable to COVID-19<sup>15</sup>.

Death reporting delay was calculated by the difference in days between occurrence date and reporting date, from March 23 to August 29, 2020 (when the death reporting delay was significantly reduced). This variable was presented as median and interquartile range (IQ) and in boxplots according to the Epidemiological Week of reporting. We outlined a graph to show the weekly number of deaths by date of reporting and of occurrence, and the nowcasting estimates by date of occurrence.

To adjust for incomplete notification data in recent weeks, we considered McGough et al.'s approach<sup>12</sup>. Implementation was done by R package NobBS (<https://www.r-project.org/>), using a negative binomial model for death reporting delay, with both adaptation phase and burn-in of 5000 iterations, and the software's default values for the remaining parameters.

To illustrate the impact of death reporting delay and to test the effectiveness of nowcasting in improving data quality, we fitted a semi-mechanistic Bayesian hierarchical model to death data by reporting date and to nowcasted death data by occurrence date. This model uses death data (not cases), which we consider to be less susceptible to underreporting and, therefore, are more appropriate for studying the magnitude and trend of the pandemic. It was originally proposed by Flaxman et al.<sup>16</sup> to describe the number of infections, number of deaths and time-varying effective reproduction number ( $R_t$ ) of COVID-19 based on death counts. This Bayesian smoothing nowcasting method models the delay distribution based on the available data of onset date (occurrence date) and reporting date of cases or deaths. Such approach consists in modeling daily deaths  $D_t$  as a negative binomial distribution<sup>16</sup>:

$$D_t \sim \text{Negative Binomial}(d_t, d_t + d_t^2/\psi)$$

Where  $d_t$  represents the expected number of deaths attributed to COVID-19 on day  $t$  and  $\psi$ , given by a half-normal distribution  $\psi \sim N(0.5)$ <sup>16</sup>. The calculation of  $d_t$  takes into account the number of new infections on day  $t$ , namely  $C_t$ , and the probability of death  $\varpi$ , given by a weighted sum:

$$d_t = \sum_{\tau=0}^{t-1} C_{\tau} \pi_{t-\tau}$$

i.e., the expected number of deaths on day  $t$  is a sum of previous infections weighted by their probability of death<sup>16</sup>. By considering the temporal relation between the death and its reporting, the model relies on a more realistic estimation process; it was then adapted to incorporate Google mobility data, assuming that population mobility patterns are linked to transmission intensity, and used to analyze the pandemic in 16 Brazilian states<sup>17</sup>. Google mobility data were described by four covariates: mobility in residential areas, transit stations, parks and the average between grocery, pharmacy, retail and workplace areas. For this study, we used data for the city of São Luís publicly available at <https://www.google.com/covid19/mobility/>.

We fitted this model in two moments: first, with data until May 09, 2020 and then with information until June 30, 2020. May 09, 2020 was chosen for being the date when deaths by COVID-19 in the city of São Luís reached their peak; June 30, 2020, in turn, corresponds to when the peak had already passed, but the number of death reports continued to increase. Each analysis compared estimates of the daily number of infections and deaths with the time-varying reproduction number ( $R_t$ ), considering deaths by notification date, by occurrence date, by occurrence date correcting notification delay by nowcasting, and by occurrence date including deaths notified until December 24, 2020 – date when most deaths in the first half of 2020 had already been reported, without further delay. Besides São Luís, we also included data on the other 15 state capitals from the original report<sup>17</sup> to fulfill the partial-pooling of covariates coefficients of the model.

All analyses were performed using R version 4.0.218 (<https://www.r-project.org/>) and the graphics were plotted using Ggplot2 package<sup>18</sup>.

## Results

Figure 1 presents the boxplots of COVID-19 death reporting delay in days by reporting week. We can observe delays from the first death report, on March 29. In the weeks from June 14 to August 29 deaths were reported with greater delay, reaching a median of 103 days between August 16 and 22 and a maximum value of 126 days late on August 25.

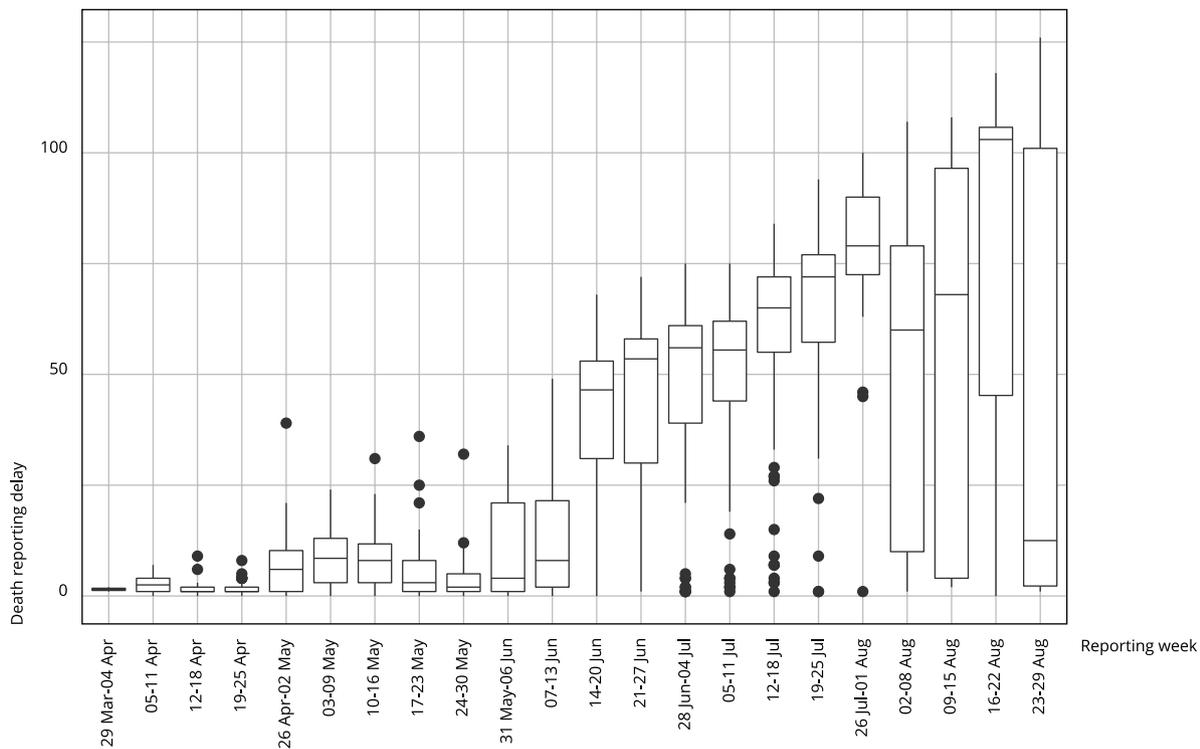
Only 17.8% of deaths were reported without delay or the day after, while 40.5% were reported more than 30 days late, with a median of 14 days and IQ of 53 days. From April 19 to June 13, 2020, period with the highest number of deaths due to COVID-19 in São Luís, the median delay was 5 days (IQ = 10). After this period, the median delay increased to 57 days (IQ = 28), with 81.6% showing a delay greater than 30 days (Table 1).

Figure 2 shows the weekly number of observed deaths by occurrence and reporting date. The notification curve is, in general, delayed when compared to the observed deaths by occurrence date. As a result, we have an initial underestimation of deaths, whose peak in the week from May 3rd to 9th, 2020 reached 214 deaths; while the peak of notifications, between May 10 and May 16, had 106 deaths reported. From June 7 to August 29, 2020, in turn, we observed a high reporting of deaths that occurred before this period: 644 deaths were reported, but only 116 deaths occurred in these weeks.

Figure 3 presents the daily number of infections, deaths and time-varying reproduction number ( $R_t$ ) until May 9, estimated by a semi-mechanistic Bayesian hierarchical model of COVID-19. Considering estimates adjusted for deaths by notification date, the daily number of infections and deaths showed an increasing pattern and  $R_t$  was above 1. By May 9, the model estimated 20 deaths per day (Figure 3a). Regarding estimates by occurrence date, the daily number of infections and deaths rose early in the pandemic and almost levelled off by May 9, with an estimation of 10 deaths per day and

**Figure 1**

COVID-19 death reporting delay in days by reporting Epidemiological Week. São Luís, Maranhão State, Brazil, 2020.



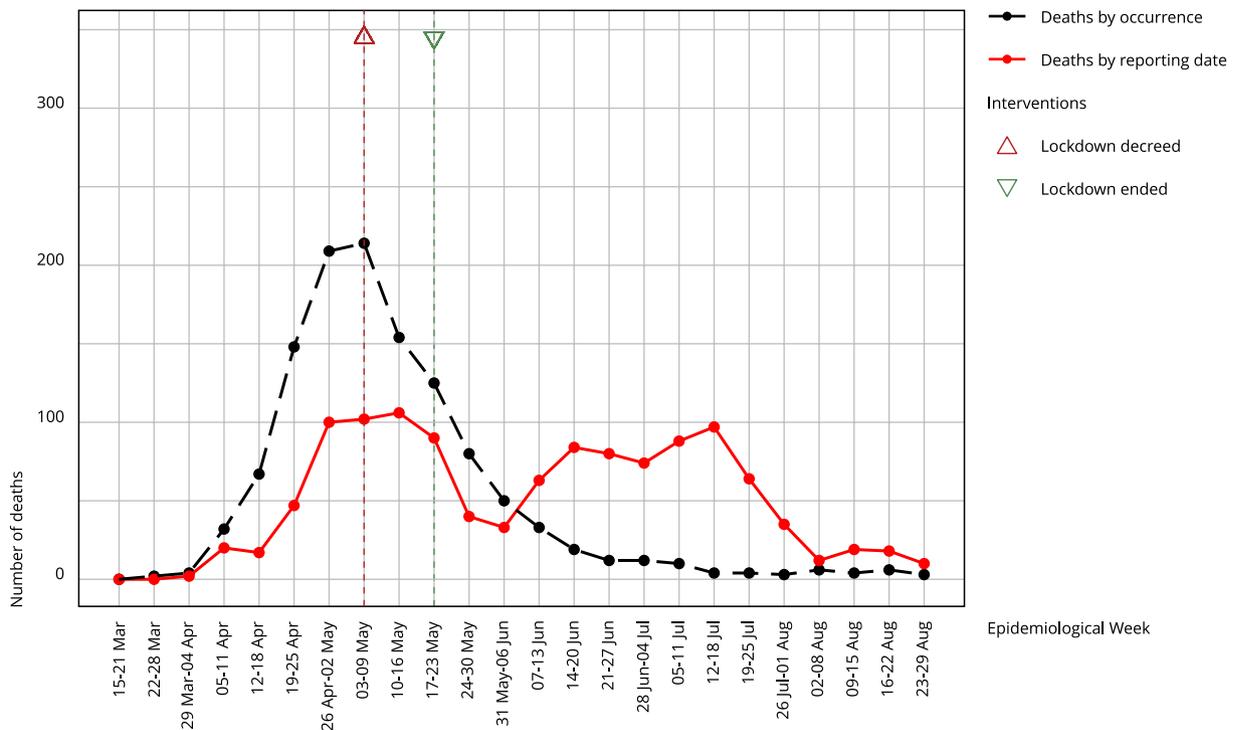
**Table 1**

COVID-19 death reporting delay in days presented as median, interquartile range and delay by reporting period. São Luís, Maranhão State, Brazil, 2020.

Reporting period	COVID-19 death reporting delay								Total reported deaths
	Median (days)	Interquartile range (days)	Up to 1 day		Up to 7 days		More than 30 days		
			n	%	n	%	n	%	
March 29 to August 29	14	53	214	17.8	455	37.9	486	40.5	1,201
March 29 to April 18	2	3	18	46.2	38	97.4	0	0.0	39
April 19 to June 13	5	10	173	29.8	350	60.2	12	2.1	581
June 14 to August 29	57	28	23	4.0	67	11.5	474	81.6	581

**Figure 2**

Weekly number of observed deaths by occurrence and reporting date and Bayesian nowcasting estimated deaths by occurrence date. São Luís, Maranhão State, Brazil, 2020.

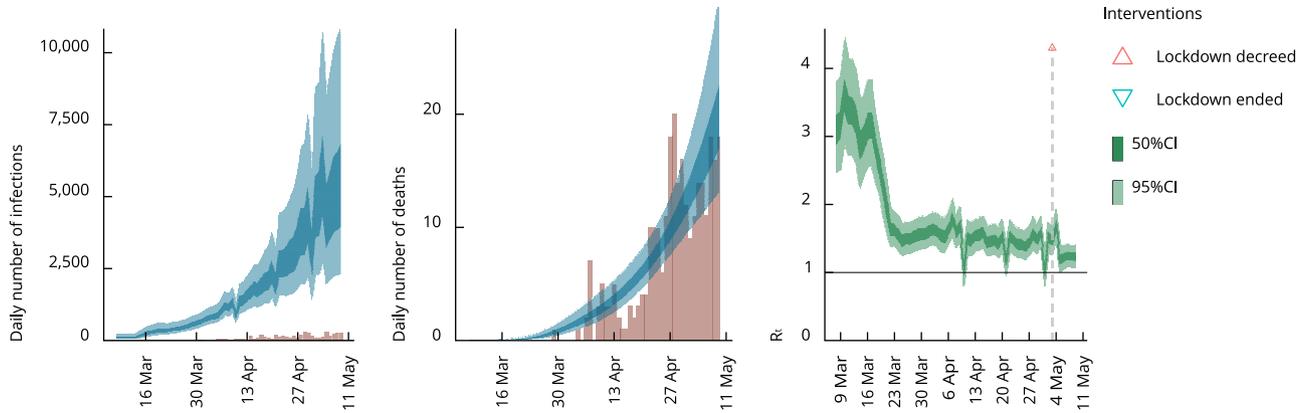


the 50% credible interval for the  $R_t$  below 1, suggesting a deceleration of the pandemic (Figure 3b). The estimates by occurrence date correcting death reporting delay by nowcasting, showed an increasing daily number of infections and deaths and  $R_t$  above 1; by May 9, however, the model estimated nearly 30 deaths per day, 3 times higher than the estimate obtained without nowcasting (Figure 3c). As for estimates by occurrence date including all deaths reported until December 2020, we found an increasing daily number of infections and deaths and  $R_t$  above 1. However, the model estimated 45 deaths per day by May 9, 50% higher than the number obtained by nowcasting and 4.5 times higher than the uncorrected estimates (Figure 3d).

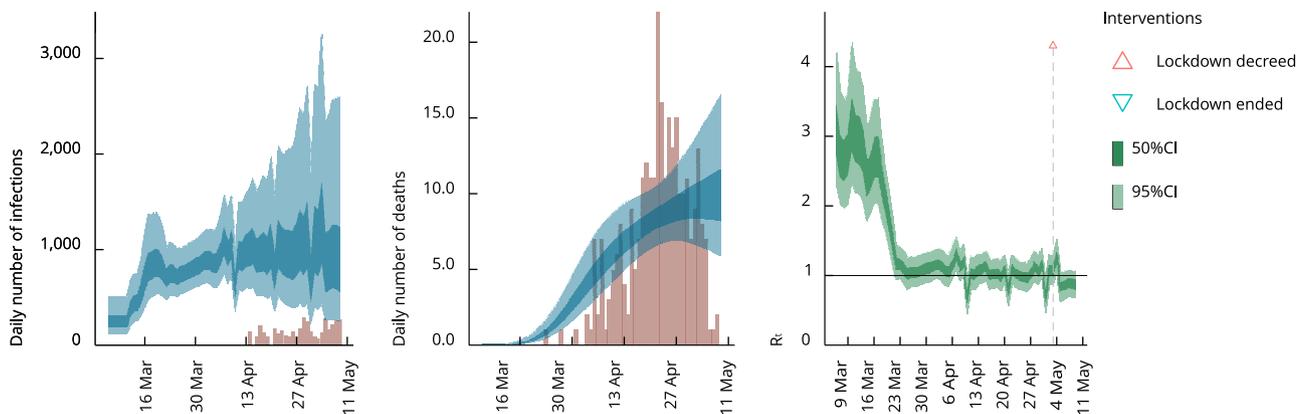
**Figure 3**

Daily number of infections, deaths and time-varying reproduction number ( $R_t$ ) until May 9, 2020 estimated by a semi-mechanistic Bayesian hierarchical model of COVID-19. São Luís, Maranhão State, Brazil.

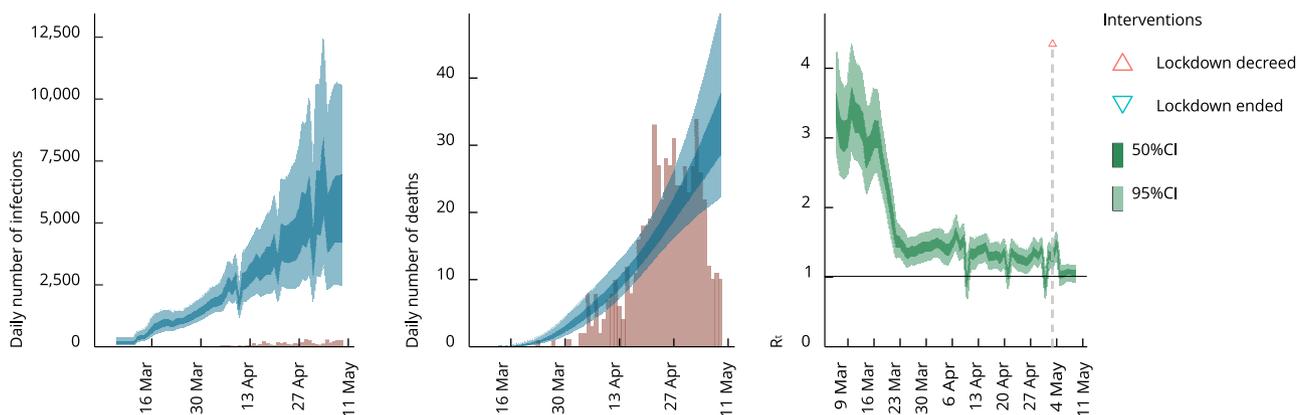
3a) Estimates fitted to deaths by notification date



3b) Estimates fitted to deaths by occurrence date



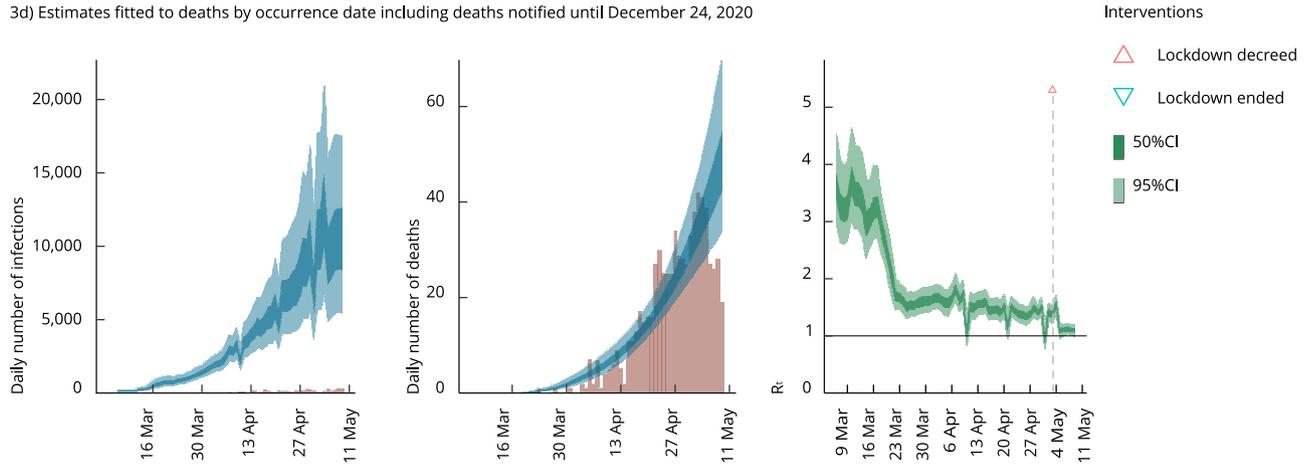
3c) Estimates fitted to deaths by occurrence date correcting notification delay by nowcasting



(continues)

Figure 3 (continued)

3d) Estimates fitted to deaths by occurrence date including deaths notified until December 24, 2020



50%CI: 50% credible interval; 95%CI: 95% credible interval.

Legend: Left: blue bands are estimated infections, dark blue 50%CI, light blue 95%CI and red bars are reported cases. Middle: blue bands are predicted deaths, credible interval as in left plot and red bars are reported deaths. Right: green bands are  $R_t$  estimates, dark green 50%CI and light green 95%CI.

Figure 4 shows the same estimates, but obtained until June 30. Considering deaths by notification date, the daily number of infections increased more steeply in June than in May. From March to May the daily number of deaths rose, then levelled off and finally increased again.  $R_t$  remained around 1 until May, when it increased and stayed above 1, suggesting that the pandemic was still on the rise (Figure 4a). Estimates by occurrence date, however, depicted a different scenario. The number of infections and deaths rose steeply until early May and then decreased sharply, reaching low values by June 30, suggesting that the worst of the pandemic had passed.  $R_t$  values dropped below 1 in early May (Figure 4b). Estimates corrected by nowcasting (Figure 4c) or including deaths until December (Figure 4d) showed little change.

## Discussion

Our results showed a high delay in COVID-19 death reporting in the city of São Luís, which affected the city's ability to monitor the pandemic dynamics over time, first underestimating and more recently overestimating the actual number of deaths. We also observed that combining death data with reporting delay affected the results of the Bayesian model of COVID-19, changing its estimates and depicting a scenario incompatible with the reality seen in São Luís at the time. The use of Bayesian nowcasting technique to minimize the delay in death notification partially improved the quality of mortality data during the peak of the pandemic, presenting estimates that best match the scenario observed in the city, but became less useful almost 2 months after the peak.

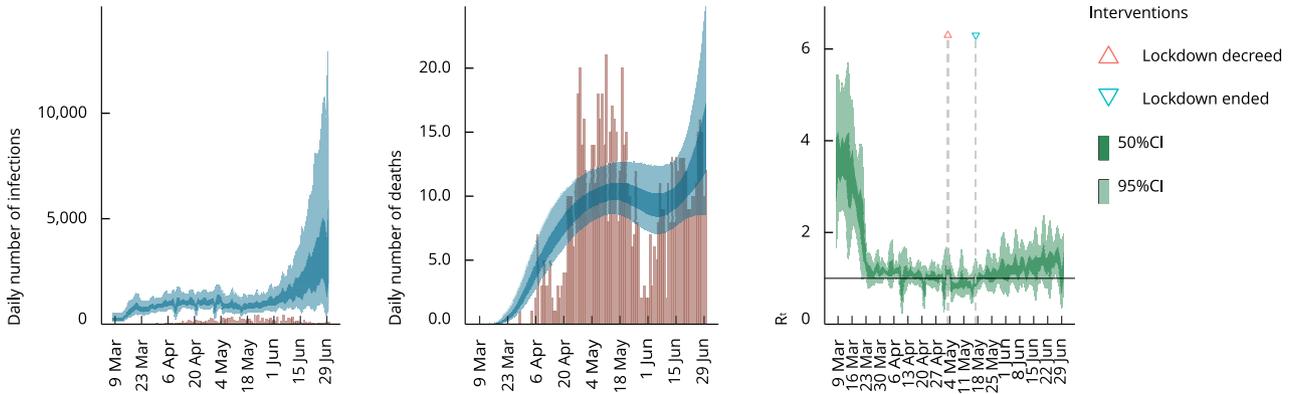
The death notification peak was 102% lower, taking place one week after the actual peak of deaths in the city. Thus, the reporting delay caused the pandemic to be seen retroactively, appearing to be of lesser extent in the official records. Measures taken to combat COVID-19 could therefore have occurred out of step with the current epidemiological dynamics if only death reporting data were considered in decision-making.

From June 7 to July 18, 2020, the Epidemiological Weeks after the peak, deaths by reporting date showed an upward curve incompatible with the scenario observed *in-loco*, characterized by a reduced  $R_t$  to values below 1<sup>19</sup>, successive weeks of reduction in the number of new cases and decreasing occu-

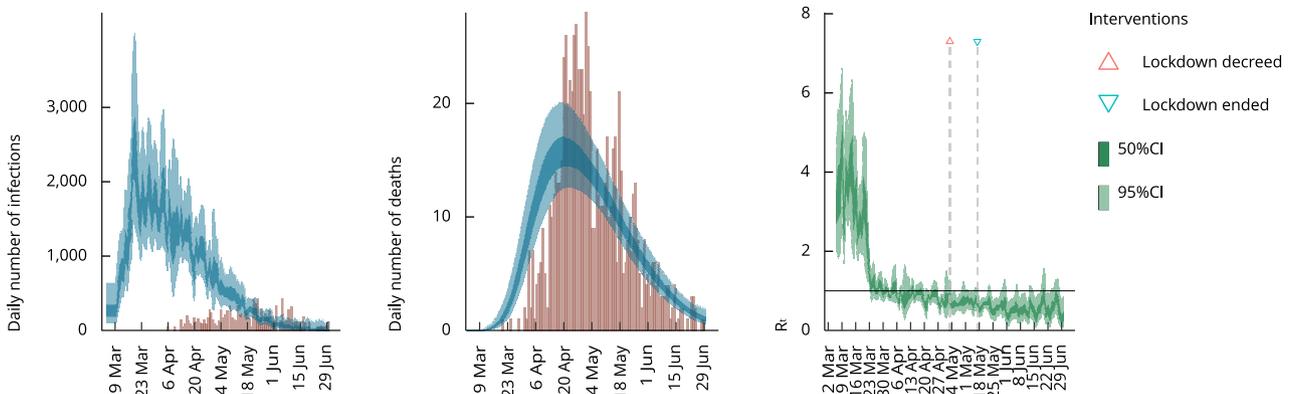
**Figure 4**

Daily number of infections, deaths and time-varying reproduction number ( $R_t$ ) until June 30, 2020 estimated by a semi-mechanistic Bayesian hierarchical model of COVID-19. São Luís, Maranhão State, Brazil.

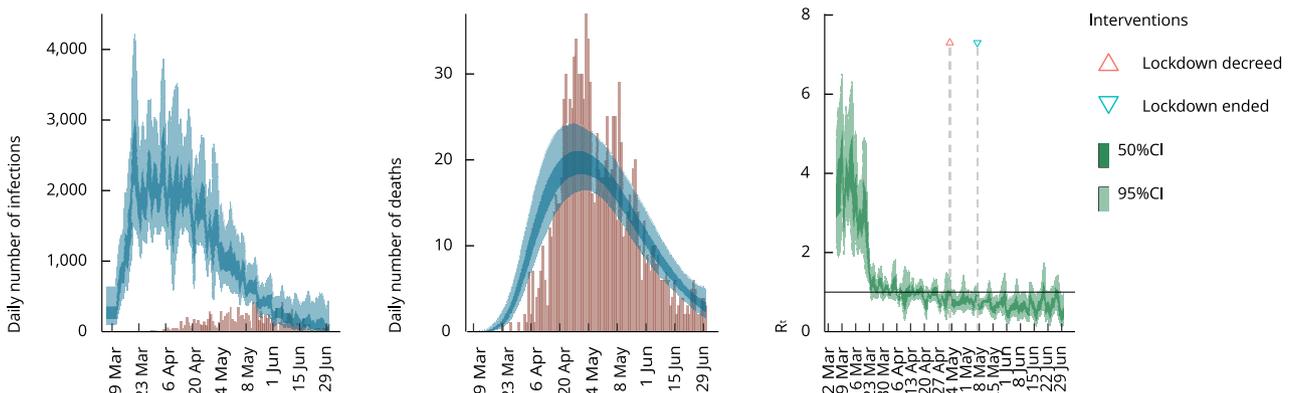
4a) Estimates fitted to deaths by notification date



4b) Estimates fitted to deaths by occurrence date



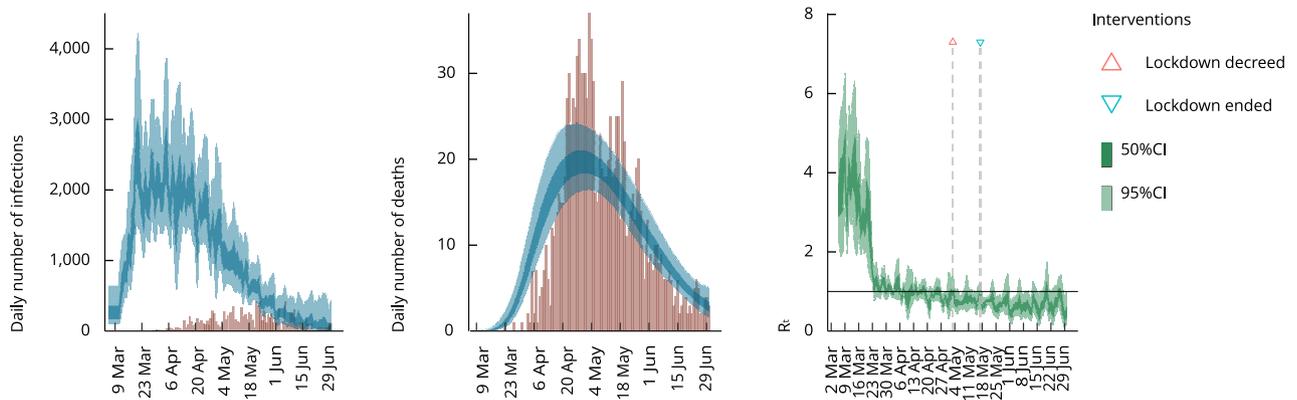
4c) Estimates fitted to deaths by occurrence date correcting notification delay by nowcasting



(continues)

Figure 3 (continued)

4c) Estimates fitted to deaths by occurrence date correcting notification delay by nowcasting



50%CI: 50% credible interval; 95%CI: 95% credible interval.

Legend: Left: blue bands are estimated infections, dark blue 50%CI, light blue 95%CI and red bars are reported cases. Middle: blue bands are predicted deaths, credible interval as in left plot and red bars are reported deaths. Right: green bands are  $R_t$  estimates, dark green 50%CI and light green 95%CI.

pancy of clinical and intensive care unit (ICU) beds<sup>20</sup>. This upward death curve resulted from death reporting delays, which had a median of 57 days between June 14 and August 29, 2020 and 81.6% of deaths notified more than 30 days before their occurrence. Death reporting delays therefore led to an overestimation of the number of deaths as of June 7, 2020.

Difficulty in obtaining properly updated data on mortality can hinder decision making, increasing the likelihood that disease control actions will not be implemented at the most opportune times. After court order, São Luís enforced the lockdown from May 5 to 17, 2020, motivated, among other factors, by the high and growing number of deaths by COVID-19 recorded in the city<sup>21</sup>. The analysis of deaths by occurrence date shows that the start of the lockdown coincided with the peak of deaths, registering a decline in the following week. As the estimated mean timeframe from infection to death is around 23 days<sup>22,23</sup>, one can conclude that the lockdown only had an impact on COVID-19 mortality after its peak, when we observed a decreasing trend. Moreover, the peak number of deaths that prompted the lockdown decree (April 30, 2020) had already been reached at least one week earlier. If the death surveillance system in São Luís had been thoroughly updated, then the decision to implement the lockdown would likely have been made earlier. Timely decision-making is crucial in a pandemic like COVID-19, with its very rapid transmission dynamics and effects that profoundly affect the health system and the economy.

Death reporting delay also negatively impacts the results of statistical models that use death numbers as a starting point to estimate the number of cases and the  $R_t$ . Using the same Bayesian model implemented by the Imperial College of London to analyze the pandemic dynamics in 16 Brazilian states<sup>17</sup>, we estimated the daily number of cases, deaths and  $R_t$  for the city of São Luís on May 9, 2020 during the pandemic peak. Models adjusted for deaths by reporting date correctly showed the upward pandemic trend, but underestimated its magnitude. Models adjusted for deaths by occurrence date, on the other hand, suggested a flatter increase, correctly predicted  $R_t$  below 1 in early May, but grossly underestimated the magnitude of the pandemic. The nowcasting-corrected estimates, in turn, came close to the real situation, correctly predicted  $R_t$  but still underestimated the magnitude of daily infections and deaths.

We also estimated the daily number of cases, deaths and  $R_t$  as of June 30, 2020, after the peak, but when death reporting was still on the rise. Models adjusted for deaths by reporting date proved unus-

able as they erroneously predicted that the pandemic was still on the rise when the worst was over. Almost 2 months after the peak, models adjusted for deaths by occurrence date, nowcasting-corrected or not or including deaths notified until December estimated more accurate daily number of infections and deaths and  $R_t$ .

These results suggests benefits of using nowcasting to correct estimates of the daily number of cases, deaths and  $R_t$ , thus allowing for better monitoring of the pandemic when faced with a huge delay in death reporting. But nowcasting still underestimated the magnitude of the pandemic. After the pandemic peak, nowcasting ceased to be advantageous, as over time most of the previously unreported deaths had already been included. After the peak, analyses based on deaths by occurrence date corresponded more closely to the pandemic dynamics in the city, which was showing a reduction in the number of new cases and a decrease in the occupancy of clinical and ICU hospital beds <sup>20</sup>.

Using death data unadjusted for reporting delay can result in estimates or interpretations that correspond to a reality of days or even weeks past, and is unable to fulfill its purpose of helping to predict pandemic trends over time, as shown in this study. Places with significant death reporting delay, are recommended to avoid using data on deaths by reporting date as a parameter for monitoring the pandemic. In this context, using mortality data by occurrence date or adjusted by nowcasting are better options for timely pandemic monitoring and decision-making. In cities with high quality mortality data and shorter reporting delay, we expect the impact of nowcasting to be smaller.

The median delay of death reporting increased over time, with the most recent Epidemiological Weeks showing a delay then times greater than the previous weeks. Most deaths reported from June 7 to August 29, 2020 corresponded to deaths that occurred in April and May 2020, when the city of São Luís experienced the peak of the disease and the number of suspected cases was very high <sup>24</sup>. Such scenario overwhelmed the public testing laboratories. We believe that during this time testing priority was given to suspected cases rather than deaths, which were mostly tested and reported after the peak. Another factor that may contribute to the observed delay in death reporting is the failure of reporting and logistics management. Increases in the number of deaths should always be analyzed considering the delay in notification, otherwise they may generate “false alarms” of a second wave.

A limitation of the present study is the possibility that underreporting of deaths due to COVID-19 may still be present, leading to underestimation of the Bayesian model results. A second limitation is that the Google mobility data is an estimate based exclusively on individuals who use this technology, so it may not be representative of the population. Moreover, demographic groups most affected by COVID-19, such as older adults, may be underrepresented in this data, as may more vulnerable socio-economic groups <sup>25</sup>. It should be noted, however, that Google mobility data is the only publicly available mobility indicator for São Luís, and the use of a mobility measure is key to improve  $R_t$  estimation <sup>26</sup>. Thus, we believe that the benefits of its use outweigh the bias due to possible underrepresentation.

On the other hand, a strong point of the study is the use of Bayesian nowcasting to minimize reporting delay and to track pandemic dynamics based on death data. Bayesian models are flexible enough to capture relevant data properties. We can incorporate information on underreporting of deaths and define the prior distribution. These data are combined with the information from the prior distribution to predict deaths.

This study observed that the delay in reporting deaths due to COVID-19 in São Luís was high and impacted the timely monitoring of pandemic dynamics. These problems can directly interfere with assertive and timely decision making, particularly in the face of rapidly spreading pandemics like COVID-19, with serious health, economic, and social repercussions. Our findings therefore point to institutional weaknesses in ensuring the quantity and quality in the recording of data necessary to describe the health reality. Moreover, these results alert to the need to review the epidemiological surveillance system in Maranhão and Brazil, as it is likely that a similar data quality issue exists in other state municipalities and in the country. In Brazil, the COVID-19 pandemic revealed flaws that point to years of underfunding and undervaluation of mortality and epidemiologic surveillance data.

We suggest that researchers and managers investigate the quality of mortality data available in their city or country, to check for reporting delays that may affect tracking the COVID-19 pandemic dynamics. In places where the quality of mortality data is low, we recommend that local epidemiological surveillance should seek to reduce or solve this issue by relegating the analysis of suspected deaths to the background. We also suggest disclosing detailed information, such as the number of suspected

deaths and date of death. Finally, researchers and/or managers who intend to use death information in statistical models should check for reporting delay and account for it in their estimates. To this end, the nowcasting technique proved to be somewhat effective in improving the quality of mortality data, allowing the estimation of dynamic transmission parameters that best fit the epidemiological situation at the pandemic peak. Sometime after the peak, however, when the surveillance system was less overloaded, nowcasting hardly improved the monitoring of pandemic dynamics.

In conclusion, improved up-to-date reporting of deaths is mandatory for better monitoring of transmission dynamics. More investments to improve epidemiologic surveillance systems are urgently needed.

## Contributors

All authors participated in the study design and planning, data acquisition, analysis, and interpretation and manuscript development.

## Additional informations

ORCID: Carolina Abreu de Carvalho (0000-0001-7900-4642); Vitória Abreu de Carvalho (0000-0002-7831-5761); Marcos Adriano Garcia Campos (0000-0001-8924-1203); Bruno Luciano Carneiro Alves de Oliveira (0000-0001-8053-7972); Eduardo Moraes Diniz (0000-0002-6201-1771); Alcione Miranda dos Santos (0000-0001-9711-0182); Bruno Feres de Souza (0000-0003-1997-4983); Antônio Augusto Moura da Silva (0000-0003-4968-5138).

## References

1. Antunes LCC. Modelação do atraso na notificação dos óbitos no sistema de vigilância diária da mortalidade [Master's Thesis]. Lisboa: Faculdade de Ciências da Universidade de Lisboa; 2014.
2. França EB, Ishitani LH, Teixeira RA, Abreu DMX, Corrêa PRL, Marinho F, et al. Deaths due to COVID-19 in Brazil: how many are there and which are being identified? *Rev Bras Epidemiol* 2020; 23:e200053.
3. Paixão B, Baroni L, Salles R, Escobar L, Sousa C, Pedroso M, et al. Estimation of COVID-19 under-reporting in Brazilian States through SARI. *ArXiv* 2020; 23 jun. <https://arxiv.org/abs/2006.12759>.
4. Prado MF, Antunes BB, Bastos LSL, Peres IT, Silva AAB, Dantas LF, et al. Analysis of COVID-19 under-reporting in Brazil. *Rev Bras Ter Intensiva* 2020; 32:224-8.
5. Alves THE, Souza TA, Silva SA, Ramos NA, Oliveira SV. Underreporting of death by COVID-19 in Brazil's second most populous state. *Front Public Health* 2020; 8:578645.
6. Silva AAM, Lima-Neto LG, Azevedo CMPES, Costa LMM, Bragança MLBM, Barros Filho AKD, et al. Population-based seroprevalence of SARS-CoV-2 and the herd immunity threshold in Maranhão. *Rev Saúde Pública* 2020; 54:131.
7. Cobre AF, Böger B, Fachi MM, Vilhena RO, Domingos EL, Tonin FS, et al. Risk factors associated with delay in diagnosis and mortality in patients with COVID-19 in the city of Rio de Janeiro, Brazil. *Ciênc Saúde Colet* 2020; 25 Supl 2:4131-40.

8. Moreno AC, Vieira B, Tito F, Manzano F, Marques P MR. Menos da metade dos testes de COVID-19 do Brasil permite saber se vírus está ativo no organismo. Portal 2020; 21 aug. <https://g1.globo.com/bemestar/coronavirus/noticia/2020/08/21/menos-da-metade-dos-testes-de-covid-19-do-brasil-permite-saber-se-virus-esta-ativo-no-organismo.ghtml>.
9. Schumacher FL, Ferreira CS, Prates MO, Lachos A, Lachos VH. A robust nonlinear mixed-effects model for COVID-19 death data. *Statistics and its Interface* 2021; 14:49-57.
10. Lau H, Khosrawipour T, Kocbach P, Ichii H, Bania J, Khosrawipour V. Evaluating the massive underreporting and undertesting of COVID-19 cases in multiple global epicenters. *Pulmonology* 2021; 27:110-5.
11. Observatório COVID-19 BR. Óbitos por COVID-19 no país. <https://covid19br.github.io/analises.html?aba=aba2> (accessed on 01/Oct/2020).
12. McGough SF, Johansson MA, Lipsitch M, Menzies NA. Nowcasting by Bayesian smoothing: a flexible, generalizable model for real-time epidemic tracking. *PLoS Comput Biol* 2020; 16:e1007735.
13. Günther F, Bender A, Katz K, Küchenhoff H, Höhle M. Nowcasting the COVID-19 pandemic in Bavaria. *Biom J* 2021; 63:490-502.
14. Altmejd A, Rocklöv J, Wallin J. Nowcasting COVID-19 statistics reported with delay: a case-study of Sweden. *arXiv* 2020; 11 jun. <http://arxiv.org/abs/2006.06840>.
15. Secretaria de Estado da Saúde do Maranhão. Boletim Epidemiológico COVID-19. Boletim atualizado 25/01/2021. <https://www.saude.ma.gov.br/wp-content/uploads/2021/01/BOLETIM-25-01.pdf> (accessed on 01/Oct/2020).
16. Flaxman S, Mishra S, Gandy A, Unwin HJT, Coupland H, Mellan TA, et al. Report 13 - Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries. <https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/report-13-europe-npi-impact/> (accessed on 01/Oct/2020).
17. Mellan T, Hoeltgebaum H, Mishra S, Whittaker C, Schnekenberg R, Gandy A, et al. Report 21: estimating COVID-19 cases and reproduction number in Brazil. <https://spiral.imperial.ac.uk:8443/handle/10044/1/78872> (accessed on 01/Oct/2020).
18. Wickham H. *Ggplot2: elegant graphics for data analysis*. New York: Springer; 2009.
19. Observatório COVID-19 BR. R efetivo em São Luís. [https://covid19br.github.io/municipios.html?aba=aba3&uf=MA&mun=Sao\\_Luis&q=dia#](https://covid19br.github.io/municipios.html?aba=aba3&uf=MA&mun=Sao_Luis&q=dia#) (accessed on 07/Oct/2020).
20. Secretaria de Estado da Saúde do Maranhão. Boletim Epidemiológico COVID-19. Boletim atualizado 04/07/2020. <http://www.saude.ma.gov.br/wp-content/uploads/2020/07/BOLETIM-04-07.pdf> (accessed on 01/Oct/2020).
21. Alves SMC, Ramos EMB, Delduque MC. Decretação de lockdown pela via judicial: medida (des)necessária? *Cad Saúde Pública* 2020; 36:e00116020.
22. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus – infected pneumonia. *N Engl J Med* 2020; 382:1199-1207.
23. Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. *Lancet Infect Dis* 2020; 20:669-77.
24. Secretaria de Estado da Saúde do Maranhão. Boletim Epidemiológico COVID-19. Boletim atualizado 02/05/2020. <http://www.saude.ma.gov.br/wp-content/uploads/2020/05/BOLETIM-02-05.pdf> (accessed on 01/Oct/2020).
25. Sulyok M, Walker M. Community movement and COVID-19: a global study using Google's Community Mobility Reports. *Epidemiology & Infection* 2020; 148:e284.
26. Kraemer MUG, Yang CH, Gutierrez B, Wu CH, Klein B, Pigott DM, et al. The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* 2021; 368:493-7.

## Resumo

O estudo teve como objetivos, descrever o atraso na notificação de óbitos por COVID-19 na cidade de São Luís, Maranhão, Brasil, e demonstrar o impacto sobre o monitoramento oportuno e modelagem da pandemia. O estudo teve como objetivo secundário determinar a medida em que a nowcasting é capaz de diminuir a defasagem na notificação de óbitos. Analisamos os dados de mortalidade por COVID-19 registrados diariamente no Boletim Epidemiológico da Secretaria de Estado da Saúde do Maranhão e calculamos o atraso na notificação entre 23 de março e 29 de agosto de 2020. Para ilustrar o impacto do atraso na notificação de óbitos e testar a efetividade de uma nowcasting bayesiana para melhorar a qualidade dos dados, ajustamos um modelo hierárquico bayesiano semi-mecanicista. Apenas 17,8% dos óbitos foram notificados sem atraso ou no dia seguinte, enquanto 40,5% foram atrasados em mais de 30 dias. Devido ao atraso na notificação, houve uma subestimação inicial nos óbitos. Entre 7 de junho e 29 de agosto, 644 óbitos foram notificados, mas apenas 116 mortes ocorreram nesse período. O uso da técnica de nowcasting bayesiana melhorou parcialmente a qualidade dos dados de mortalidade no pico da epidemia, apresentando estimativas mais ajustadas ao cenário observado na cidade, mas não se mostrou útil quase dois meses depois do pico. O atraso na notificação de óbitos pode interferir diretamente nas decisões assertivas e oportunas sobre o combate à pandemia da COVID-19. Portanto, o sistema brasileiro de vigilância epidemiológica deve ser revisto urgentemente, e o registro da data do óbito deve ser obrigatório. A técnica de nowcasting mostrou ser parcialmente eficaz na melhoria dos dados de mortalidade no auge da pandemia, mas não depois.

COVID-19; SARS-CoV-2; Registros de Mortalidade; Confiabilidade dos Dados

## Resumen

La propuesta de este estudio es describir la demora en la notificación de muertes por COVID-19, en la ciudad São Luís, Maranhão, Brasil, y demostrar su impacto en el seguimiento puntual, así como en el modelaje de la pandemia de COVID-19. Un objetivo secundario fue confirmar el alcance, donde la previsión inmediata es capaz de mejorar el retraso en la notificación de las muertes. Analizamos los datos de muertes por COVID-19 diariamente en el Boletín Epidemiológico de la Secretaría de Estado de la Salud de Maranhão y calculamos los atrasos notificados desde el 23 de marzo al 29 de agosto, 2020. Con el fin de ilustrar el impacto del retraso en la notificación de muertes, y para probar la efectividad de la predicción inmediata bayesiana en la mejora de los datos de calidad, ajustamos un modelo jerárquico bayesiano semi-mecanicista. Solo un 17.8% de las muertes se notificaron sin atrasos o el día después, mientras que un 40.5% se vieron retrasadas durante más de 30 días. Debido a la demora informada, se produjo una subestimación inicial de muertes. No obstante, desde el 7 de junio al 29 de agosto, se informó de 644 muertes, pero solamente 116 muertes se produjeron durante este periodo. El uso de la técnica de predicción inmediata bayesiana mejoró parcialmente la calidad de la información de mortalidad durante el pico de la epidemia, presentando estimaciones que se ajustan mejor al escenario observado en la ciudad, pero no fue útil casi 2 meses después del pico. El retraso en la notificación de muertes podría interferir directamente en la toma de decisiones asertivas y puntuales, respecto a la pandemia de COVID-19. Por consiguiente, se debe revisar urgentemente el sistema brasileño de vigilancia epidemiológica y la notificación de la fecha de muerte debería ser obligatoria. La técnica de predicción inmediata ha demostrado ser bastante efectiva para mejorar la calidad de los datos de mortalidad solamente en el pico pandémico, pero no después.

COVID-19; SARS-CoV-2; Registros de Mortalidad; Exactitud de los Datos

Submitted on 08/Oct/2020

Final version resubmitted on 05/Apr/2021

Approved on 15/Apr/2021