journal of
global
health

PAPERS

Early detection of respiratory disease outbreaks through primary healthcare data

Thiago Cerqueira-Silva[1,2] ,
Izabel Marcilio[2] , Vinicius de
Araújo Oliveira[2] , Pilar Tavares
Veras Florentino[2] , Gerson
O Penna[3] , Pablo I Pereira
Ramos[2] , Viviane S
Boaventura[1,4] , Manoel Barral-
Netto[1,2]

[1]Laboratório de Medicina e Saúde Pública
de Precisão – Instituto Gonçalo Moniz,
Salvador, Bahia, Brazil
[2]Centro de Integração de Dados e
Conhecimentos para Saúde – Instituto
Gonçalo Moniz, Salvador, Bahia, Brazil
[3]Centro de Medicina Tropical –
Universidade de Brasília, Escola Fiocruz
de Governo, Brasília, Brazil
[4]Faculdade de Medicina da Bahia,
Universidade Federal da Bahia, Salvador,
Brazil

Background The emergence of coronavirus disease 2019 (COVID-19) in 2020 highlighted the relevance of surveillance systems in detecting early signs of potential outbreaks, thus enabling public health authorities to act before the pathogen becomes widespread. Syndromic digital surveillance through web applications has played a crucial role in monitoring the spread of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus. However, this approach requires expensive infrastructure, which is not available in developing countries. Pre-existing sources of information, such as encounters in primary health care (PHC), can provide valuable data for a syndromic surveillance system. Here we evaluated the utility of PHC data to identify early warning signals of the first COVID-19 outbreak in Bahia-Brazil in 2020.

Methods We compared the weekly counts of PHC encounters due to respiratory complaints and the number of COVID-19 cases in 2020 in Bahia State – Brazil. We used the data from December 2016 to December 2019 to predict the expected number of encounters in 2020. We analysed data aggregated by geographic regions (n = 34) and included those where historical PHC data was available for at least 70% of the population.

Results Twenty-one out of 34 regions met the inclusion criteria. We observed that notification of COVID-19 cases was preceded by at least two weeks with an excess of encounters of respiratory complaints in 18/21 (86%) of the regions analysed and four weeks or more in 10/21 (48%) regions.

Conclusions Coronavirus disease 2019 (COVID-19) syndromic surveillance systems based on already established PHC databases may add time to preparedness and response to emerging epidemics.

The coronavirus disease 2019 (COVID-19) pandemic highlighted the importance of surveillance systems capable of detecting respiratory outbreaks at early stages, even before pathogen identification. The recent re-emergence of avian influenza A (H5N1), a potential pandemic virus, reinforces the urgency of implementing syndromic tools for the early detection of respiratory diseases. Digital surveillance systems using syndromic data have played a crucial role in monitoring the spread of the severe acute respiratory syndrome coronavirus 2 virus (SARS-CoV-2). Initiatives such as those proposed by the Zoe COVID-19 symptom group and COvid-19: Operation for Personalized Empowerment to Render smart prevention And care seeking (COOPERA) exemplify the high utility of early warning systems [1,2]. However, they require the population's engagement by accessing the web application and self-reporting information, as well as widespread access to the internet and expensive infrastructures, which limits their application in developing countries.

Correspondence to:
Manoel Barral-Netto
Instituto Gonçalo Moniz, Fundação
Oswaldo Cruz, Salvador
121 Waldemar Falcão Street, Candeal,
Salvador, Bahia, Brazil
Brazil
manoel.barral@fiocruz.br

Here, we present a case study of a syndromic surveillance tool in Brazil that uses digital information routinely collected by the primary healthcare (PHC) system to feed an early warning system.

METHODS

We used data from the PHC system described in another study [3]. We compared the weekly counts of PHC encounters due to respiratory complaints and the number of COVID-19 cases in 2020 in the state of Bahia, Brazil's fourth most populous state with 14.9 million inhabitants in 417 municipalities and a territorial extension comparable to France (Figure S1 in the **Online Supplementary Document**). We used the data from December 2016 to December 2019 to predict the number of encounters expected for 2020. We classified encounters coded with any International Classification of Diseases, 10th Revision (ICD-10) or International Classification of Primary Care, 2nd edition (ICPC-2) related to acute respiratory infection (Table S1 in the **Online Supplementary Document**) as encounters due to respiratory complaints. We used the total number of encounters per week as the population offset in the model.

We assessed the completeness of historical data by the number of weeks with a record of any encounters in previous years. For example, a city was considered to have complete historical data if it had at least 155 weeks with PHC records from 31 October 2016, to 29 December 2019 (total of 165 weeks). The cities were aggregated by immediate geographic region (n = 34), as per the Brazilian Institute of Geography and Statistics [4]. We included the regions with more than 70% of the population with complete historical data. We also conducted a similar analysis using the three largest cities in Bahia: Salvador (population = 2 953 986), Feira de Santana (population = 627 477), and Vitória da Conquista (population = 348 718).

Statistical analysis

We employed the algorithm proposed by Farrington and modified by Noufaliy to estimate the expected number of encounters due to respiratory complaints each week and their respective 95% confidence interval [5]. The upper limit of the 95% confidence interval was defined as the threshold for considering excess cases. The parameters used in the algorithm were three years of historical data using a window half size of three weeks.

We extracted the number of confirmed COVID-19 cases from Bahia's Secretary of Health public dashboard [6] and employed a piecewise linear regression in the weekly COVID-19 cases in each region to define the onset of the rapid SARS-CoV-2 transmission per region (Figure S5 in the **Online Supplementary Document**).

We extracted the number of hospitalisations due to severe acute respiratory syndrome (SARS) from the *Sistema de Informação de Vigilância Epidemiológica da Gripe* (*SIVER-Gripe*) nation-wide surveillance database [7].

We conducted all statistical analyses in R, version 3.6.1 (R Core Team, Vienna, Austria) and its "surveillance", version 1.20.3 [8] and "segmented" version 1.6.2 [9] packages.

Ethics

The Bahia State Health Secretary for research purposes provided weekly aggregated data, as authorised in the Brazilian General Protection law. We did not seek ethical approval, as the data could not be re-identified in any way.

RESULTS

The first COVID-19 case was detected on 25 February 2020 in Brazil and on 6 March 2020 in Bahia. We evaluated a total of 24 882 367 encounters from 2016 to 2019 in 34 immediate geographic regions, 21 (62%) of which met the inclusion criteria (Table S2 in the **Online Supplementary Document**), resulting in 18 036 110 encounters used as baseline data. The validation data were comprised of 4 096 495 encounters that occurred in 2020 (**Figure 1**, Panels A and B)

We observed an excess of encounters due to respiratory complaints in 18 (86%) regions at least two weeks before the increase in COVID-19 cases, with a median of four weeks (interquartile range (IQR) = 3-6) (Figure S2 in the **Online Supplementary Document**). For example, the region of Vitória da Conquista presented six weeks with an excess of encounters in the nine weeks between the first confirmed COVID-19
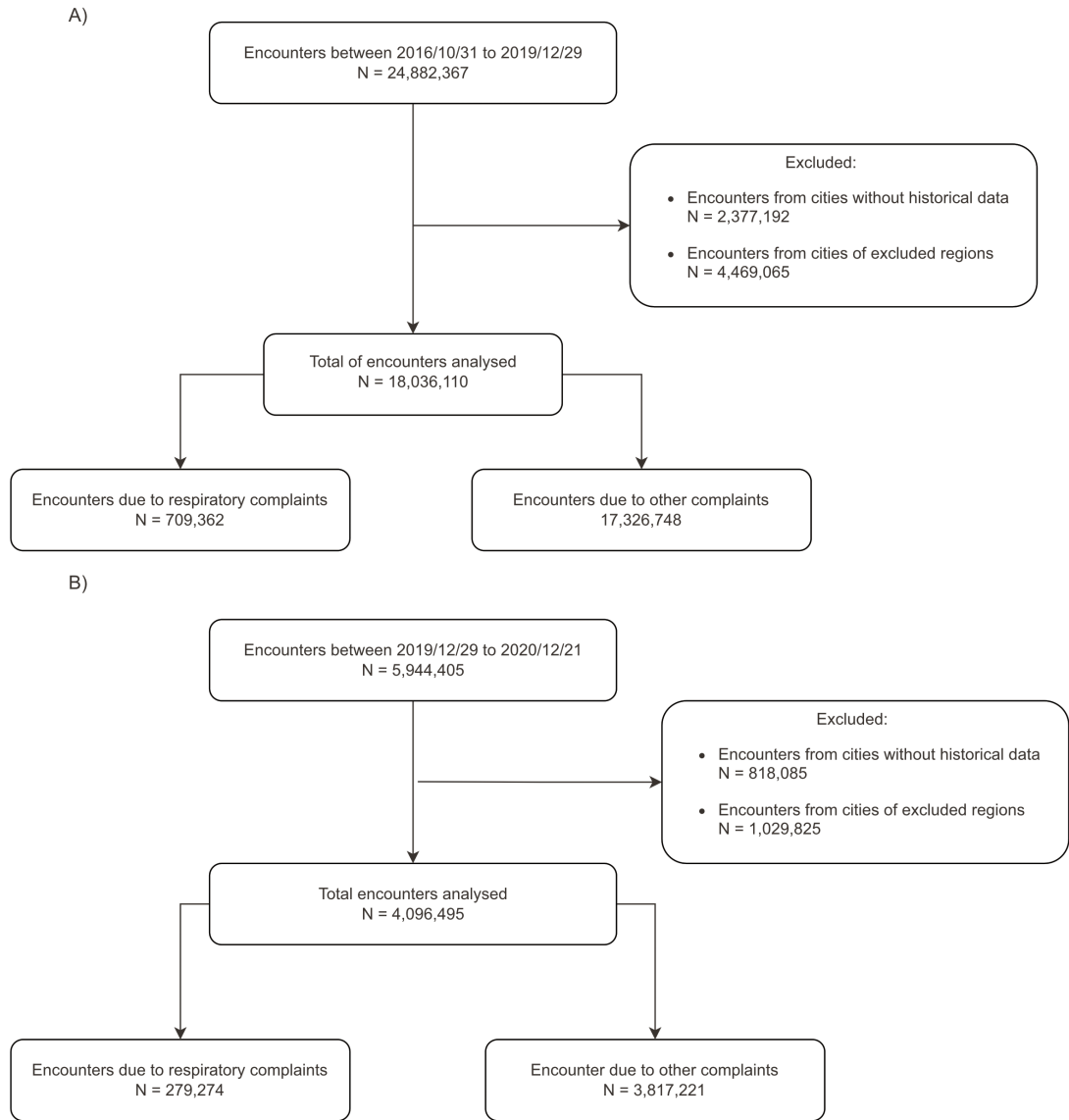
A)

Encounters between 2016/10/31 to 2019/12/29
N = 24,882,367

Excluded:
- Encounters from cities without historical data
  N = 2,377,192
- Encounters from cities of excluded regions
  N = 4,469,065

Total of encounters analysed
N = 18,036,110

Encounters due to respiratory complaints
N = 709,362

Encounters due to other complaints
17,326,748

B)

Encounters between 2019/12/29 to 2020/12/21
N = 5,944,405

Excluded:
- Encounters from cities without historical data
  N = 818,085
- Encounters from cities of excluded regions
  N = 1,029,825

Total encounters analysed
N = 4,096,495

Encounters due to respiratory complaints
N = 279,274

Encounter due to other complaints
N = 3,817,221

Figure 1. Flowchart of the PHC data analysed. Panel A. Baseline. Panel B. Test period.
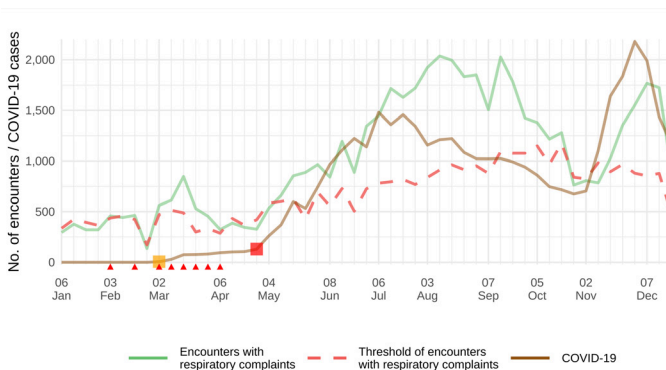

Figure 2. Weekly syndromic surveillance signals from primary health care and weekly COVID-19 cases in the region of Vitoria da Conquista in 2020. The yellow square indicates the week of the first COVID-19 case confirmed in Bahia, the red square indicates the first week of the rapid growth of COVID-19 cases based on piecewise linear regression, and red triangles denote weeks with excess encounters due to respiratory complaints prior to the rapid growth of COVID-19 cases.

case in Bahia and the week of rise in COVID-19 cases in this region (**Figure 2**). Regarding SARS-related hospitalisations, all 21 regions presented COVID-19 hospitalisations after April 2020, coinciding with the pattern of total COVID-19 cases (Figure S4 in the **Online Supplementary Document**).

In the remaining three regions, two presented only one week with excess encounters before the increase in COVID-19 cases, and one did not exhibit any excess encounters due to respiratory complaints before the rise in COVID-19 cases. Analysis of the three largest cities in Bahia showed a minimum of three weeks with an excess of encounters before the rise of COVID-19 (Feira de Santana) and up to seven weeks with an excess of encounters (Vitória da Conquista) (Figure S3 in the **Online Supplementary Document**).

DISCUSSION

We demonstrate that it is possible to anticipate the rise in respiratory infection cases by analysing routinely collected PHC data in most Bahia State regions. The lack of detection seemed to relate to areas with sparse populations or those underreporting PHC data. This suggests a potential utility of digital health data in early-warning systems for surveillance purposes.
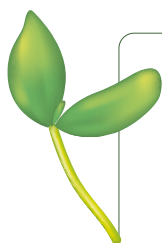
Our findings regarding regions with an excess of PHC encounters due to respiratory complaints prior to the first reported COVID-19 case in Bahia-Brazil are consistent with studies that found evidence of the circulation of SARS-CoV-2 virus two to four weeks in three countries before the first confirmed case by local public health agencies, either using serological surveillance [10,11] or genomic characterisation [12,13]. Only hospitalised patients were reported at the beginning of the COVID-19 pandemic. However, most COVID-19 cases do not require hospitalisation. Since there is a lag of up to two weeks between SARS-CoV-2 infection and progress to severe disease, this situation may also have contributed to the observed excess PHC respiratory-related encounters before the confirmation of the first COVID-19 case. These factors combined hindered the detection of COVID-19 cases in the early period of the pandemic.

In our analysis, the number of encounters in the PHC after the peak of COVID-19 differed greatly by immediate region. One possible reason for this phenomenon is the lack of standardisation in public health recommendations in Brazil during COVID-19 in 2020. The Brazilian Ministry of Health did not develop a national plan in response to the pandemic and did not implement non-pharmacological interventions [14]. Consequently, Brazilian municipalities had to develop individual plans to mitigate the burden of COVID-19, which resulted in heterogeneous and uncoordinated actions [15].

PHC is an integral part of the Brazilian Unified Health System (SUS), which serves as the entrance to healthcare services for all levels of complexity. It is based on family health teams comprising medical doctors, nurse practitioners, nursing assistants, and community healthcare agents. PHC covers at least 74% of the Brazilian population, including the marginalised and those living in remote areas [16]. Despite these characteristics, some factors can interfere with PHC's potential for early detection of future epidemics. One such challenge arises from the lack of data derived from the private healthcare system, which becomes particularly problematic when addressing diseases initially imported from other countries, affecting primarily the more affluent populations who do not use SUS [17] Additionally, concerns persist regarding the potential delays in reporting PHC data, although these delays tend to be shorter than traditional systems relying on laboratory testing [18].

Our study is subject to several limitations. First, we could not link individual PHC encounters and hospitalisation data. This would have allowed us to gain further insights into the number of individuals who progressed to severe cases and their trajectory within the health system. Additionally, our analysis was limited to a single State in Brazil. However, we should note that the State of Bahia is geographically vast and has a highly diverse range of municipalities, with Human Development Index values ranging from 0.5 to 0.75. The early detection in multiple regions of the state provides evidence of the broad applicability and utility of PHC data. This suggests that our approach could be effective across a wide range of situations, further emphasising the potential of PHC data in epidemiological surveillance for early detection of outbreaks.

Digital syndromic surveillance systems based on already established databases, such as the PHC database, may add timeliness to preparedness and response to emerging epidemics. Additionally, implementing such a system could improve the number and quality of PHC data registration, contributing to a structural improvement of the local health system. This strategy is especially advantageous for developing countries that chronically face constrained health resources.

Disclosure of interest: The authors completed the ICMJE Disclosure of Interest Form (available upon request from the corresponding author) and disclose no relevant interests.

Additional material
Online Supplementary Document

REFERENCES

1 Rossman H, Keshet A, Shilo S, Gavrieli A, Bauman T, Cohen O, et al. A framework for identifying regional outbreak and spread of COVID-19 from one-minute population-wide surveys. Nat Med. 2020;26:634-8. Medline:32273611 doi:10.1038/s41591-020-0857-9

2 Desjardins MR. Syndromic surveillance of COVID-19 using crowdsourced data. Lancet Reg Health West Pac. 2020;4:100024. Medline:34013214 doi:10.1016/j.lanwpc.2020.100024

3 de Araújo Oliveira V, Sironi A, Florentino PTV, Marcilio I, Cerqueira-Silva T, Flores-Ortiz R, et al. Syndromic Detection of Upper Respiratory Infections in Primary Healthcare as a Candidate for COVID-19 Early Warning in Brazil: A Retrospective Ecological Study [preprint]. Available from: https://papers.ssrn.com/abstract=4364869. Accessed: 16 March 2023.

4 Instituto Brasileiro de Geografia e Estatística. Geográficas Imediatas E Regiões Geográficas Intermediárias: 2017/IBGE. 2017. Available: https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhes&id=2100600. Accessed: 13 March 2023.

5 Noufaily A, Enki DG, Farrington P, Garthwaite P, Andrews N, Charlett A. An improved algorithm for outbreak detection in multiple surveillance systems. Stat Med. 2013;32:1206-22. Medline:22941770 doi:10.1002/sim.5595

6 Central Integrada de Comando e Controle da Saúde. COVID-19. Available: https://bi.saude.ba.gov.br/transparencia/. Accessed: 13 March 2023.

7 Ranzani OT, Bastos LSL, Gelli JGM, Marchesi JF, Baião F, Hamacher S, et al. Characterisation of the first 250 000 hospital admissions for COVID-19 in Brazil: a retrospective analysis of nationwide data. Lancet Respir Med. 2021;9:407-18. Medline:33460571 doi:10.1016/S2213-2600(20)30560-9

8 Hoehle M, Meyer S, Paul M, Held L, Burkom H, Correa T, et al. surveillance: Temporal and Spatio-Temporal Modeling and Monitoring of Epidemic Phenomena. 2022. Available: https://CRAN.R-project.org/package=surveillance. Accessed: 11 March 2023.

9 Muggeo VMR. segmented: Regression Models with Break-Points / Change-Points (with Possibly Random Effects) Estimation. 2022. Available from: https://CRAN.R-project.org/package=segmented. Accessed: 11 March 2023.

10 Chappell JG, Tsoleridis T, Clark G, Berry L, Holmes N, Moore C, et al. Retrospective screening of routine respiratory samples revealed undetected community transmission and missed intervention opportunities for SARS-CoV-2 in the United Kingdom. J Gen Virol. 2021;102:001595. Medline:34130773 doi:10.1099/jgv.0.001595

11 Cereda D, Manica M, Tirani M, Rovida F, Demicheli V, Ajelli M, et al. The early phase of the COVID-19 epidemic in Lombardy, Italy. Epidemics. 2021;37:100528. Medline:34814093 doi:10.1016/j.epidem.2021.100528

12 Alteri C, Cento V, Piralla A, Costabile V, Tallarita M, Colagrossi L, et al. Genomic epidemiology of SARS-CoV-2 reveals multiple lineages and early spread of SARS-CoV-2 infections in Lombardy, Italy. Nat Commun. 2021;12:434. Medline:33469026 doi:10.1038/s41467-020-20688-x

13 Franco D, Gonzalez C, Abrego LE, Carrera JP, Diaz Y, Caicedo Y, et al. Early Transmission Dynamics, Spread, and Genomic Characterization of SARS-CoV-2 in Panama. Emerg Infect Dis. 2021;27:612-5. Medline:33496228 doi:10.3201/eid2702.203767

14 Ferigato S, Fernandez M, Amorim M, Ambrogi I, Fernandes LMM, Pacheco R. The Brazilian Government's mistakes in responding to the COVID-19 pandemic. Lancet. 2020;396:1636. Medline:33096042 doi:10.1016/S0140-6736(20)32164-4

15 Lui L, Albert CE, dos Santos RM, Vieira L da C. Disparidades e heterogeneidades das medidas adotadas pelos municípios brasileiros no enfrentamento à pandemia de Covid-19. Trab Educ Saúde. 2021;19:e00319151. doi:10.1590/1981-7746-sol00319

16 Massuda A, Hone T, Leles FAG, de Castro MC, Atun R. The Brazilian health system at crossroads: progress, crisis and resilience. BMJ Glob Health. 2018;3:e000829. Medline:29997906 doi:10.1136/bmjgh-2018-000829

17 Macinko J, Lima-Costa MF. Horizontal equity in health care utilization in Brazil, 1998–2008. Int J Equity Health. 2012;11:33. Medline:22720869 doi:10.1186/1475-9276-11-33

18 Heil J, ter Waarbeek HLG, Hoebe CJPA, Jacobs PHA, van Dam DW, Trienekens TAM, et al. Pertussis surveillance and control: exploring variations and delays in testing, laboratory diagnostics and public health service notifications, the Netherlands, 2010 to 2013. Euro Surveill. 2017;22:30571. Medline:28749331 doi:10.2807/1560-7917.ES.2017.22.28.30571

PAPERS