Neonatal Mortality Rates in Brazilian Health Facilities: from 2005 to 2017

Lais Baroni

Federal Center for Technological Education Celso Suckow da Fonseca

Rebecca Salles

Oswaldo Cruz Foundation

Eduardo Ogasawara

Federal Center for Technological Education Celso Suckow da Fonseca

Patricia de Moraes Mello Boccolini

Faculdade de Medicina de Petrópolis

Jefferson Costa Lima

Oswaldo Cruz Foundation

Cristiano Siqueira Boccolini (✉ cristiano.boccolini@iciict.fiocruz.br)

Oswaldo Cruz Foundation

Research Article

Keywords: neonatal mortality rate, time series data, health facilities

Posted Date: July 20th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-726636/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Neonatal Mortality Rates in Brazilian Health Facilities: from 2005 to 2017

Lais Baroni^{1,3†}, Rebecca Salles^{1,2†}, Eduardo Ogasawara^{1†}, Patricia de Moraes Mello Boccolini^{4†}, Jefferson da Costa Lima^{2†} and Cristiano Siqueira Boccolini^{3*†}

*Correspondence:

cristiano.boccolini@icict.fiocruz.br

³ Oswaldo Cruz Foundation,

Fiocruz, Rio de Janeiro, Brazil

Full list of author information is available at the end of the article

†Equal contributor

Abstract

Objectives: Neonatal mortality is a global public health problem. The efforts to reduce child mortality are one of the 2030 Agenda for Sustainable Development goals. The availability of neonatal mortality rates (NMR) data in Brazilian health facilities is crucial to understanding the dynamic, identify gaps, and provide information to allocate resources and actions to reduce NMR at the health facility level. The objective is to offer an integrated dataset containing monthly data in a historical series from 2007 to 2017 with information on all births, neonatal deaths, and NMR (total, early, and late components) enriched with the municipality's information.

Data description: Dataset with historical data with information on the number of births, the neonatal deaths, the neonatal mortality rate (including early and late components), and geographic information for each health facility in Brazil, comprising the years between 2005 and 2017.

Keywords: neonatal mortality rate; time series data; health facilities

Background

Neonatal mortality is a global public health problem [1]. The efforts to reduce child mortality are one of the 2030 Agenda for Sustainable Development goals [2]. The availability of neonatal mortality rates (NMR) data in Brazilian health facilities is crucial to understand the dynamics, identify gaps, and provide information to allocate resources and enable actions to reduce NMR at the health facility level [3]. To tackle these scenarios, in this work, we built the dataset called *NM_Facility*. The objective is to offer an integrated dataset containing monthly data in a historical series from 2007 to 2017 with information on all births, neonatal deaths, and NMR (total, early, and late components) enriched with the municipality's information. This dataset and related files are available for download, as shown in Table 1.

All files listed in Table 1 are stored in the Synapse repository within the research project “Assessing breastfeeding technologies” to which it belongs. The Synapse platform provides the tools for data sharing. It is responsible for the user interface that can view all code, data, and descriptions available. It is a free platform. It is possible to download all data as long as one is registered on the platform.

The authors have previously published a dataset with information on neonatal mortality rates in Brazilian municipalities [5]. The dataset presented here differs from that by having health facilities as its unit of analysis. The availability of a dataset with this unit of analysis is important due to the variability between

Table 1 Overview of data files/datasets.

Label	Name of data file/dataset	File type (file extension)	Data repository and identifier
Dataset 1	NM.Facility	delimited text (.csv)	Synapse: https://doi.org/10.7303/syn23651701 [4]
Data file 1	dict_NM_facility	delimited text (.csv)	Synapse: https://doi.org/10.7303/syn23651701 [4]
Data file 2	data_construction	portable document (.pdf)	Synapse: https://doi.org/10.7303/syn23651701 [4]
Data file 3	data_validation	R code (.r)	Synapse: https://doi.org/10.7303/syn23651701 [4]
Data file 4	municipalities_codes_IBGE	delimited text (.csv)	Synapse: https://doi.org/10.7303/syn23651701 [4]
Data file 5	data_construction	jupyter notebook (.ipynb)	Synapse: https://doi.org/10.7303/syn23651701 [4]
Data file 6	ETLTMCNES	Dataiku DSS project (.zip)	Synapse: https://doi.org/10.7303/syn23651701 [4]
Data file 7	NMR_CNES_translation	jupyter notebook (.ipynb)	Synapse: https://doi.org/10.7303/syn23651701 [4]

different health facilities in the same municipality concerning neonatal deaths [6]. It opens up a range of possibilities for discoveries oriented towards case studies and comparative analyses [7].

Construction and content

The dataset developed is the result of a construction process organized into three main activities: (i) data extraction and aggregation, (ii) computation of neonatal mortality rates, and (iii) data enrichment. Data management activities were based on detailed studies of the dataset and support of specialists in the field ^[1].

i. The developed dataset encompasses data from the Live Birth Certificate System (SINASC) and the Death Certificate System (SIM) from DATASUS (Ministry of Health). Both were made available by the Platform of Data Science applied to Health (PCDaS) [8]. It results in annual datasets representing individual records of live births and deaths, respectively.

Such records were aggregated (counted) by each competency date (combination of month and year), municipality (place of birth in Brazil), and health facility (located in each municipality). The resulting aggregations create absolute values with the total number of births and deaths used as denominators and numerators, respectively.

In particular, we restrict ourselves to neonatal deaths for building numerators from SIM death records, which means deaths occurred in the age group between 0 and 27 days. We further divide this period, aggregating neonatal deaths between 0 and 6 days (early component) and 7 and 27 days (late component). Moreover,

^[1]Description of the construction of the datasets is available at <https://doi.org/10.7303/syn23651701>.

we are only interested in neonatal deaths that occurred in hospitals or other health facilities. These steps are explained in detail and can be reproduced with the Jupyter Notebook file available at <https://doi.org/10.7303/syn23651701>.

Finally, additional temporal aggregations (semester and annual moving average) were constructed from the absolute values of births and neonatal deaths (denominators and numerators) for computing neonatal mortality rates.

ii. The total, early, and late neonatal mortality rates (NMR) were calculated for each instance aggregated in time (month, semiannual moving average, and annual) and space (health facilities) according to Equation 1.

$$\text{NMR} = \frac{\text{number of neonatal deaths}}{\text{number of births}} \cdot 1000 \quad (1)$$

iii. Next, the resulting dataset was enriched with neonatal mortality rates calculated over the different scales of the Brazilian territorial division where the health facilities are located. These rates are analogous to the previously calculated. However, they correspond to different spatial aggregations: municipalities, states, regions, and the entire country. These data were retrieved from the published dataset *NM_County_B* [5].

Finally, the dataset was enriched with information about the municipalities, states, and regions obtained from the DATASUS FTP server^[2]. The attributes resulting from this process are the coordinates, area, regional health codes, and other Boolean attributes, indicating if the municipality is the state's capital or border, for example. The data dictionary^[3] presents a full description.

Utility and discussion

The dataset *NM_Facility* presents 671,489 records and 128 attributes. One attribute is responsible for identifying the maternity by its National Health Facilities Registry (or 'Cadastro Nacional de Estabelecimentos de Saúde', CNES, in Portuguese) registration number. There are forty-five attributes regarding the number of deaths. They are organized according to combinations of time, space, and type of neonatal death component. Time attributes include month, moving average, and year. Space attributes include maternity, city, state, region, country. Finally, the type of neonatal death includes total, late, and early.

Forty-five attributes associated with rates also exist, analogous to those of deaths. There are fifteen attributes regarding births include space and time dimensions. There are nineteen attributes with information about the cities and states to which each maternity belongs, like coordinates, areas, and other information described in the last paragraph of the previous section. There are still the attributes year, month, and year/month in timestamp format^[4].

In a time-series study, it is essential to count on the availability of a continuous historical series. There is a maximum of 149 competency dates for each health facility in the *NM_Facility*, from August 2005 to December 2017. In *NM_Facility*,

^[2] Available at <ftp://ftp.datasus.gov.br/dissemin/publicos/SINASC/NOV/TABELAS/>.

^[3] Available at <https://doi.org/10.7303/syn23651701>.

^[4] The timestamp format represents the seconds since the date of origin, 01/01/1970.

approximately 28% of the health facilities have more than half of the competency dates, and a little more than 17% have more than 90%. Only 1.3% have all 149 competency dates present.

The absence of complete registration can be related to (a) creation of a health facility after August 2005; (b) discontinuation of a health facility before December 2017; (c) temporary interruption in the health facilities services (e.g., for maintenance; by intervention from sanitary agencies; or by judicial decisions); (d) changes in administration hierarchy or funding (e.g., from the federal level to the municipality level, or from private to public funding), which will imply in changes on the CNES registration number; (e) lack of updates despite the mandatory monthly updates; or (f) discontinuation of maternity delivery wing services (with the continuation of other services).

Since neonatal deaths are a relatively rare event, the number of neonatal deaths in maternity is equal to zero for the vast majority of competency dates. It makes the corresponding NMR also equal to zero. In *NM_Facility*, 81.68% of the records have a monthly NMR equal to zero. *NM_Facility* includes 4,369 cities of the 5,570 existing in Brazil. Not all municipalities in Brazil have birth delivery facilities or maternity wings, forcing mothers to have their babies in surrounding cities. All 27 Brazilian states were represented in the data. The number of observations is stable over the months and years. That is, there is no observed pattern of lack of information in time.

Figure 1 shows the distribution of neonatal mortality rate values in 2017 for each region of Brazil^[5]. It considers only the 864 health facilities with at least 1,000 births during the year. They correspond to 60 in the Midwest region, 232 in the Northeast region, 78 in the North region, 364 in the southeast region, and 130 in the south region. It can be seen that the northeast and midwest regions present more significant variations in annual NMR among their health facilities. The northeast region has the lowest median. The south region has the highest median, which does not vary much between regions.

Figure 2.a presents a map of the Brazilian states and their category according to the number of health facilities. The state of São Paulo has the highest number of health facilities among all the states, which is consistent with the fact that it also has a considerably larger population. Three states in the northern region (AP, AC, and RR) and two in the northeast region (SE and AL) have less than 70 health facilities. This finding is consistent with the population and territorial extension of these states.

Figure 2.b shows the map with categories and labels indicating the NMR in 2017 in the states. Amapá and Sergipe are the only states to present NMR higher than 11. The south and southeast regions present, in general, lower rates when compared to other regions of Brazil. An exception is Rio de Janeiro, which is the only state to present an NMR higher than eight in these regions. The north and northeast regions comprise the states with the highest NMR, except for the state of Rondônia. Rondônia, together with Mato Grosso do Sul (central-west region), are the only states outside the south and southeast regions with NMR lower than eight.

^[5]The outliers were smoothed using winsorization of 1% for better characterization of the distribution

Observing Figures 1 and 2.b, there is an apparent disagreement in the values of NMR of the Brazilian regions. Figure 1 shows the south region with the highest NMR median and the northeast with the lowest NMR median. When looking at Figure 2.b, however, it is noted that the states of the northeast region, in general, have higher rates than the states of the south region. It is justified because in Figure 1 we included only health facilities with more than 1,000 births in 2017.

Conversely, in Figure 2.b, all health facilities in the state were considered. It reveals an important feature of these data. The health facilities that perform a higher number of births are, with few exceptions, those with the higher maternity infrastructure and are specialized. Naturally, these health facilities will have higher NMR.

This dataset construction belongs to a sequence of datasets oriented towards explaining the determinants of neonatal mortality. The previous dataset [5] had infor-

mation on NMR at the city, state, region, and country level. The present dataset has the health facility as the unit of analysis and was enriched with the previous one (*NM_County_B* dataset).

Figure 3 is presented to make a comparison between the municipality's NMR and three of its health facilities. The cities Porto Alegre (RS), Rio de Janeiro (RJ), and Belém (PA) were selected since they have, respectively, the minimum, median, and maximum NMR values in 2017. The three health facilities in each municipality were selected because they had the highest number of births in 2017.

It is noticed that the NMR of the health facilities has a very particular NMR pattern, differing considerably from the NMR pattern observed in the municipality. This difference is both in the pattern of behavior, and the magnitude of the NMR (note that the scales for NMR vary from graph to graph) justifies creating this dataset.

Conclusions

The dataset *NM_Facility* was designed and built to make available historical data on neonatal mortality in Brazilian health facilities. Its availability opens up a range of possibilities for research topics, from case studies in a single health facility and period, through analysis of a group of health facilities with characteristics of interest, to a broader study using the entire dataset.

The availability of monthly variables for the total of neonatal deaths and births per month and by health facility enables novel neonatal mortality rates to be computed using any period and region of interest. Furthermore, the fact that there is a line for each health facility/month/year is easier to integrate *NM_Facility* with other datasets.

The dataset analysis presented in the discussion section shows that *NM_Facility* is a rich and detailed dataset for the temporal analysis of neonatal mortality in health facilities Brazilians. In addition to being a potential object of study in the health area, the dataset is also convenient in data science, especially for studies focused on time series.

Limitations

- Some health facilities do not have a complete historical series. They may only have data corresponding to a short period or sparse intervals related. They may be either discontinued or new health services within the dataset time frame or small health facilities, with no deaths or deliveries during long periods.
- There are cases where the NMR exceeds one thousand. They represent the death of a baby in a month without deliveries in the health facility, possibly due to transferring of babies from a nearby city or another health facility. We suggest, during analysis, the addition of a filter to ignore the NMR above 200/1,000 births.
- There are cases (mostly between 2005 and 2016) in which the same health facility was registered as belonging to different municipalities. It resulted in distinct aggregated values for the number of births or neonatal deaths. In these cases, we suggest summing the values regarding the same health facility in the same competence.

Figure 3 Time series plots. NMR by Region.

- There is a health facility registered with the number CNES 8888888. It occurred in only three years (from 2012 to 2014) and did not correspond to any registered health facility. It is assumed to be a typo or system error.
- The CNES was implemented in August 2005 and went through gradual implementation so that the data may have inconsistencies in the first few months.
- The municipalities of Figueirão and Itanhangá were created on January 1st, 2005. Therefore, it was not possible to find information on coordinates and areas for them. These are the only empty fields (NAs) in the datasets.

Ethics approval and consent to participate

The datasets used in this study were obtained from the DATASUS open-data. The Ministry of Health of Brazil is committed to respect the ethical precepts and to guarantee the privacy and reliability of the data.

Availability of data and materials

The dataset generated during the current study and additional documentation is freely and openly available on the Synapse repository at <https://doi.org/10.7303/syn23651701> [4].

Competing interests

The authors declare that they have no competing interests.

Author's contributions

All authors contributed equally to the study. LB; RS; CB performs the study design, manuscript writing, data curation, and data analysis. JF; PB; EO were responsible for data engineering and revised it critically for intellectual content. All authors read and approved the final manuscript.

Funding

This work was supported, in whole or in part, by the Bill & Melinda Gates Foundation [OPP1202042], National Council for Scientific and Technological Development (CNPq), Coordination of Superior Level Staff Improvement (CAPES) finance code 001, and Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) partially financed. Under the grant conditions of the Foundation, a Creative Commons Attribution 4.0 Generic License has already been assigned to the Author Accepted Manuscript version that might arise from this submission.

Acknowledgements

The authors thank PCDaS team members Sérgio Cruz (team management), Igor Morais, Lucas Carraro and Carlos Cardoso (computational infrastructure), Raphael de Freitas Saldanha, Carlos Sousa and Balthazar Paixão (data engineering and analysis) for their support enabling the development of this project.

We thank Odd.Studio for their support in the conceptual and strategic design of the development of this data project.

Author details

¹ Federal Center for Technological Education of Rio de Janeiro, CEFET/RJ, Rio de Janeiro, Brazil. ² Platform of Data Science applied to Health (PCDaS), ICICT, Oswaldo Cruz Foundation, Fiocruz, Rio de Janeiro, Brazil. ³ Oswaldo Cruz Foundation, Fiocruz, Rio de Janeiro, Brazil. ⁴ Petropolis Medical School, UNIFASE, Petrópolis, Brazil.

References

1. Lehtonen, L., Gimeno, A., Parra-Llorca, A., Vento, M.: Early neonatal death: a challenge worldwide. In: *Seminars in Fetal and Neonatal Medicine*, vol. 22, pp. 153–160 (2017). Elsevier
2. Hug, L., Alexander, M., You, D., Alkema, L., for Child, U.I.-a.G.: National, regional, and global levels and trends in neonatal mortality between 1990 and 2017, with scenario-based projections to 2030: a systematic analysis. *The Lancet Global Health* **7**(6), 710–720 (2019)
3. Hill, K., Choi, Y.: Neonatal mortality in the developing world. *Demographic research* **14**, 429–452 (2006)
4. Baroni, L., Salles, R., Boccolini, P., Lima, J., Sousa, C., Ogasawara, E., Boccolini, C.: Neonatal mortality in brazilian health facilities. Technical report, <https://doi.org/10.7303/syn23651701> (January 2021)
5. Baroni, L., Salles, R., Salles, S., Pedroso, M., Lima, J., Morais, I., Carraro, L., de Freitas Saldanha, R., Sousa, C., Cardoso, C., *et al.*: Neonatal mortality rates in brazilian municipalities: from 1996 to 2017. *BMC Research Notes* **14**(1), 1–3 (2021)
6. Moster, D., Lie, R.T., Markestad, T.: Neonatal mortality rates in communities with small maternity units compared with those having larger maternity units. *British Journal of Obstetrics and Gynaecology* **108**(9), 904–909 (2001)
7. Miranda, E.C.S., Rodrigues, C.B., Machado, L.G., Gomes, M.A.d.S.M., Augusto, L.C.R., Simões, V.M.F., Magluta, C., Lamy-Filho, F.: Situação dos leitos neonatais em maternidades brasileiras: uma análise exploratória. *Ciência & Saúde Coletiva* **26**, 909–918 (2021)
8. PCDaS, Plataforma de Ciência de Dados aplicada à Saúde. Laboratório de Informação em Saúde (Lis). Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (Icict). Fundação Oswaldo Cruz (Fiocruz). doi:10.7303/syn25882127. <https://bigdata.icict.fiocruz.br/>