CHAPTER
50

Estimation of excess mortality during COVID-19 pandemic based on death registry quality analysis.

https://doi.org/10.56238/emerrelcovid19-050

Assel Muratovna Shigayeva Ferreira
PhD student in Decision Models and Health, Department of Statistics, Federal University of Paraiba - UFPB.
E-mail: shigyassel@gmail.com;

Lucas de Almeida Moura
PhD student in Public Health Nutrition, Department of Nutrition, University of São Paulo - USP.
E-mail: lucasdemoura@usp.br;

Neir Antunes Paes
Full Professor in the Department of Statistics, Federal University of Paraiba - UFPB.
E-mail: antunes@de.ufpb.br.

ABSTRACT
Background: There was a debate whether an accurate measurement of deaths was exacerbated during COVID-19 pandemic 2020 in Brazil.
Purpose: This study aimed to analyze the magnitude and temporal course of excess mortality from the COVID-19 pandemic in 2020 with evaluation of mortality data quality in one state of Northeast region of Brazil.
Methods: Excess mortality estimation and evaluation of the quality of COVID-19 death records during 2020 pandemic was applied for the state of Paraiba. The assessment of deaths records quality included analysis of its coverage and completeness (redistribution of garbage codes (GCs) with redistribution of ill-defined causes).
Results: The total number of deaths due COVID-19 in Paraiba for 2020 increased by 37.4% after data quality correction analysis. The study proved statistically significant excess of deaths in the state for all-cause mortality and natural causes, with approximately 2-fold higher number for respiratory diseases. The quality data correction added significant fractions on top of initially observed deaths which makes the scope of exceeding deaths in Paraiba even larger.
Conclusions. There is strong evidence that the full impact of the 2020 pandemic in Paraiba has been much greater than what is indicated by reported deaths due to COVID-19 alone. It is noteworthy that the methodology applied to the state of Paraiba can be replicated for any Brazilian state.

Keywords: Covid-19, mortality, quality of death registry.

1 INTRODUCTION

Reliable and valid mortality data are essential for monitoring population health and developing target policies [1]. There was a debate whether an accurate measurement of deaths was exacerbated during COVID-19 pandemic [2].

Brazil experienced significant challenges during 2020 pandemic with registration of COVID-19 cases and associated deaths [3]. Among the possible causes of reporting bias were low testing capacity, delays in information delivery and errors in identifying causes of death. The quality of mortality data in Brazil has been gradually improving over time. Nonetheless, someissues remain, as regional differences in routine death registration [4-6]. Some problems relatedto mortality statistics quality in the country possibly have been aggravated during pandemic.

Estimating excess mortality due to the COVID-19 pandemic was globally recognized as an important tool for measuring the real toll of deaths [2,7,8]. It can draw attention to the magnitude of the crisis by providing a comprehensive comparison of additional deaths amongstthe countries and

allowing for further analysis of its causes. Empirical evidence demonstrates that COVID-19 pandemic is associated with the excess mortality due its direct (infectious disease) and indirect effects (shortage of health care resources, altered access to health care, patient's comorbidity status and others) However, there are some limitations in regional/location scope and evaluation of the quality of reported deaths. Many studies were attributed to all-cause excess mortality without cause specification [2, 9-11].

This study aimed to analyze the magnitude and temporal course of excess mortality from the COVID-19 pandemic in 2020 with evaluation of mortality data quality in one state of Northeastregion of Brazil.

2 METHODS

Data sources and study setting. Excess mortality estimation and evaluation of the quality of COVID-19 death records during 2020 pandemic was applied for the state of Paraiba, the third most densely populated federative unit in Northeast region of Brazil [12]. This state was the geographic focus part of the master's thesis of one of the authors of this work [13], which makes Paraíba a representation of what can be investigated for other Brazilian states.

The data was retrieved from the Mortality Information System (Sistema de Informação sobre Mortalidade -SIM)- a nationally consolidated official Brazilian mortality data source, which is the one among all, concisely presenting information from the death certificate including causes of deaths [14]. Weekly (1st-52nd epidemiological weeks) all -causes, natural causes and respiratory system diseases registered deaths in Paraiba for 2015-2020 were compiled [15]. The Mortality Information System protocol for recording deaths due COVID-19 was based on nationally adopted COVID-19 death definitions, certifications, and classification [16]. The data was adjusted by sex and age groups. An age breakdown of 0-59 years and 60+ years was applied.

The study followed 2 main methodological steps:

i) Assess of death registries quality and

ii) Estimation of excess mortality.

Death registries data quality analysis. To assess the data quality of deaths due COVID-19 the following approaches were determined: a) analysis of the coverage of deaths, b) analysis of deaths registration completeness by redistribution of garbage codes (GCs) and c) redistribution of ill-defined causes (IDC). The methodological approach and results of mortality data quality analysis for 2020 COVID-19 pandemic in Paraiba were published previously [17].

The coverage of all registered deaths was estimated using the General Growth Balance (GGB) method originally proposed by Brass [18]. Due to the errors inherent to any estimates and to facilitate

the interpretation of death coverage estimates, was applied the classification adapted by Paes [19], which established four categories: excellent (>90%); good (81%–90%); regular (71%–80%) and deficient (≤70%).

For redistribution of Garbage Codes (GCs) attributed to COVID-19 deaths was applied the conceptual approach outlined by Naghavi et al. (2010) [20]. The study followed four key steps:

i) Identification of GCS;

ii) Identification of target causes where the deaths assigned to a GCS should in principle be reassigned (based on pathophysiology);

iii) And calculation of the fraction of deaths assigned to a GCS that should be reallocated to target cause. In addition to identify specific GCS list and mapping it across major groups of causes was determined 4 levels typology: Level 1- codes with serious policy implications; Level 2- codes with substantial policy implications; Level 3-codes with important policy implications; Level 4- codes with limited implications. Each GC was revised in terms of pathophysiology connection to COVID- 19, and those that are likely belong to another ICD chapter or rubrics were excluded. It is important to highlight, that there is no specified GCS for COVID-19 since this disease is relatively new, with newly opened ICD-10 rubrics, and yet under observation/investigations in terms of pathogenesis, systems involved to pathology and specific reasons of deaths. The cgs for investigation were chosen based on evidence about primarily respiratory system damages due COVID-19 and its generalized sepsis complications [21]. All GCS previously extracted from major groups were redistributed accordingly to calculated fraction for COVID-19 considering their typology, sex and age-adjusted distribution.

Redistribution of ill-defined causes (IDC) was estimated by application of the Ledermann's method based on simple linear regression between the proportion of a specific cause of death and the proportion of ill-defined causes of death [22].

2.1 EXCESS MORTALITY ESTIMATION

Before estimation of expected deaths, each year of data for any mortality "shocks", such as other epidemics or natural disasters, which might not represent long-term mortality trends was observed and removed from the data. In a study were considered all-cause mortality, due to natural causes (after subtraction of external causes) and respiratory system diseases. The trend line of the deaths (weekly historical data 2015-2019) to calculate excess deaths in the period of interest was constructed. Two sets of data were used to excess deaths estimation and their comparison: observed deaths and expected deaths count. The observed deaths are referred to the deaths as registered by SIM

database. Expected deaths are the number of deaths after corrections based on the mortality quality study.

By calculating standard error (SE) and confidence interval (CI) was obtained baseline value, which was the weekly upper limit for expected deaths under normal conditions. Next, was obtained the number of excess deaths for each week by separately subtracting weekly baseline from observed/ expected number of deaths corresponding to the same week of 2020. Negative values, where the observed count fell below baseline, were set to zero. The total number of excess deaths for Paraiba was calculated by summing the excess deaths in each epidemiological week, from January 1, 2020, to December 31, 2020.

The Triple Exponential Smoothing (Holt-Winter Method) function was used to project deaths in 2020 based on historical time series 2015-2019. Two measures were used for analysis– absolute number of excess deaths, and relative measure, P-score. The absolute number of excess deaths is the difference between number of observed and number of projected deaths (baseline):

$$Excess\ deaths_{\ week\ (x)2020} = Observed\ Deaths_{\ week\ (x)2020} - Projected\ Deaths_{week(x)2020}$$

The absolute number of excess deaths provides the measure of scale. However, it's less comparable across the territories due to large differences in populations. If the result was a positive number, the week was marked as having experienced excess mortality. A measure that is more comparable across countries is the P-score, which calculates excess mortality as the percentage difference between the number of deaths in 2020 and the average number of deaths in the same period — week or month — over the years 2015–2019.

$$P-score = \frac{Deaths_{Period\ \#2020} - Average\ Deaths_{Period\ \#2015-2019}}{Average\ Deaths_{Period\#2015-2019}} * 100$$

The study has limitations: the data source considered for the study had presented certain limitations in 2020, such as: preliminary mortality data, registration of suspected cases along with confirmed, limited variables for analysis and disagreements in deaths counts. The proportions of ill-defined causes (ICD) and Garbage Codes were limited by target causes that possibly could be attributed to COVID-19, and not included all ICD and GCs. In this case, the number of misclassified deaths could be even higher. The statistical models for registered deaths correction and calculation of excess mortality could present some acceptable errors.

3 RESULTS

Table 1 summarizes the results of COVID-19 expected deaths estimation for Paraiba after investigation of mortality sources data quality (coverage by application of the General Growth Balance method and completeness by redistribution of ill-defined causes and Garbage Codes).

Table 1 - Total and sex-specific COVID-19 deaths considering observed and expected counts, Paraiba, 2020.

Deaths	**Males**	**Females**	**Total**
Observed	1929	1497	3426
Expected	2530	2177	4707
Difference	601	680	1281
(%)	31.1	45.4	37.4

Sources: Mortality Information System (SIM), 2020

The total number of deaths due COVID-19 in Paraiba for 2020 increased by 37.4% after data quality correction analysis. The percentage was higher for females comparing to males (45.4% vs 31.1%).

Table 2 shows the results of excess deaths (observed and expected) for Paraiba in 2020 with statistical outcomes based on application of historical series average and exponential smoothing function for all-cause deaths, natural causes and deaths due respiratory system diseases.

All-cause mortality. The analysis of all-cause mortality during 2020 pandemic identified 4461 of excess deaths (95%CI, 4174-4816). The percentage difference between deaths in 2020 and the average number of deaths over the years 2015–2019 was 16.5%, P <.0001. After quality data correction, an estimated number of excess deaths was 6702 (95% CI, 6467 – 7009), p- score 25%, P <.0001.

Table 2 - Excess of deaths due COVID-19 for 2020 based on historical 2015-2019 trend for all-cause mortality, natural causes, and respiratory diseases, Paraiba 2020.

	2015-2019 (Baseline)	**2020**	**Excess of deaths**		
	N (95% CI)	N	N (95% CI)	P-score (%)	p-value
Observed deaths					
All-causes	27055 [24732, 29378]	31517	4461 [4174, 4816]	16.5	<.0001
Natural -causes	24177 [23872, 24482]	28576	4399 [4094, 4704]	18.2	<.0001
Resp	3221 [3110, 3332]	6236	3015 [2904, 3126]	93.6	<.0001
Expected deaths[a]					
All-causes	27055 [24732, 29378]	33757	6702 [6467,7009]	24.7	<.0001
Natural-causes	24177 [23872, 24482]	30626	6449 [6144, 6754]	26.7	<.0001
Resp	3221 [3110, 3332]	6678	3457 [3346, 3568]	107.3	<.0001

[a]- deaths after quality data correction by application of General Growth Balance (GGB) method
Legend: Resp- respiratory system diseases, Source: SIM-Mortality Information System, 2020.

Considering the number of observed deaths due COVID-19 for Paraiba in 2020 (Table 1), the excess of deaths probably was in 75% directly impacted by virus (3426 vs 4532).

Natural causes mortality. Analysis of observed deaths due natural causes showed difference in 4399 (95% CI, 4094 -4704, p-score 18.2%, P <.0001). Estimation based on expected counts were higher (p-score 26.7%) than estimates based on observed counts with an increase of excess deaths of approximately 8.5%.

Mortality due to respiratory illness. The exploratory analysis of excess deaths due respiratory diseases in Paraiba during pandemic 2020 identified values that were two-fold higher than baseline for both observed and expected deaths. The difference with 5-year weekly average baseline showed excess of deaths in 93.6% and 107.3%. The statistically significant (P <.0001) excess of deaths in Paraiba due respiratory diseases proved direct influence of COVID-19 on excess mortality in 2020. Figure 1 showed graphic visualization of the magnitude of excess mortality due to all-causes, natural causes and respiratory diseases.

Sex and age-group distribution. Analysis of sex adjusted all-causes, natural causes and respiratory causes distribution showed the same mortality pattern for Paraiba in 2020 as was

seen in previous 5 years (2015-2019): prevalence of males compared with females. Statistically significant excess of deaths was observed both among males and females and was higher for deaths due respiratory diseases (Table 3).

For all-causes, p-score for males was 18.5% observed deaths and 24.1% expected deaths (2723 vs 3551, P<0.001). For females, the difference compared with baseline was 13.6% for observed deaths and 24.7% for expected deaths.

Table 3 - Sex-adjusted excess of observed and expected deaths due all-causes, natural causes and respiratory diseases, Paraiba, 2020

	2015-2019 (Baseline)	**2020**	**Excess of deaths**		
	N (95% CI)	N	N	P-score (%)	p-value
Observed deaths					
All-causes					
Males	14711 [12998, 16423]	17434	2723	18.5	<.0001
Females	12395 [10823, 13966]	14083	1688	13.6	<.0001
Natural causes					
Males	12503 [10924, 14081]	14880	2377	19.0	<.0001
Females	11943 [10401, 13486]	13696	1753	15.4	<.0001
Resp					
Males	1512 [963, 2061]	3360	1848	122.2	<.0001
Females	1691 [1110, 2271]	2876	1185	70.1	<.0001
Expected deaths[a]					

All-causes					
Males	14711 [12998, 16423]	18262	3551	24.1	<.0001
Females	12395 [10823, 13966]	15454	3059	24.7	<.0001
Natural causes					
Males	12503 [10924, 14081]	15587	3084	24.6	<.0001
Females	11943 [10401, 13486]	15024	3081	25.7	<.0001
Resp					
Males	1512 [963, 2061]	3519	2007	132.7	<.0001
Females	1691 [1110, 2271]	3158	1467	86.7	<.0001

[a]- deaths after quality data correction by application of General Growth Balance (GGB) method.
Legend: Resp- respiratory system diseases, Source: SIM-Mortality Information System, 2020.

Analysis of excess mortality for natural causes in terms of the magnitude showed the same tendency as for all-causes. The gap between males and females was less prominent comparing with all-causes, because of external causes subtraction. For males, the difference was in 1174 deaths (19% vs 25%, P<0.001). The highest proportion of exceed deaths was observed for respiratory diseases considering both males and females. For males was notified statistically significant (<0.0001) P-score 122.2% based on observed deaths, which increased in 10% for expected deaths.

Age-adjusted scope of deaths in 2020 as well as for weekly historical baseline showed significant prevalence of individuals 60-years-old and higher for all-causes, natural causes and respiratory diseases (Table 4). However, despite the prevalence of older age group (≥ 60-years- old) in total count of deaths for both baseline and 2020, the higher P-score was notified for group under 60-years-old.

Table 4 - Age-adjusted excess of observed and expected deaths due all-causes, natural causes and respiratory diseases, Paraiba 2020.

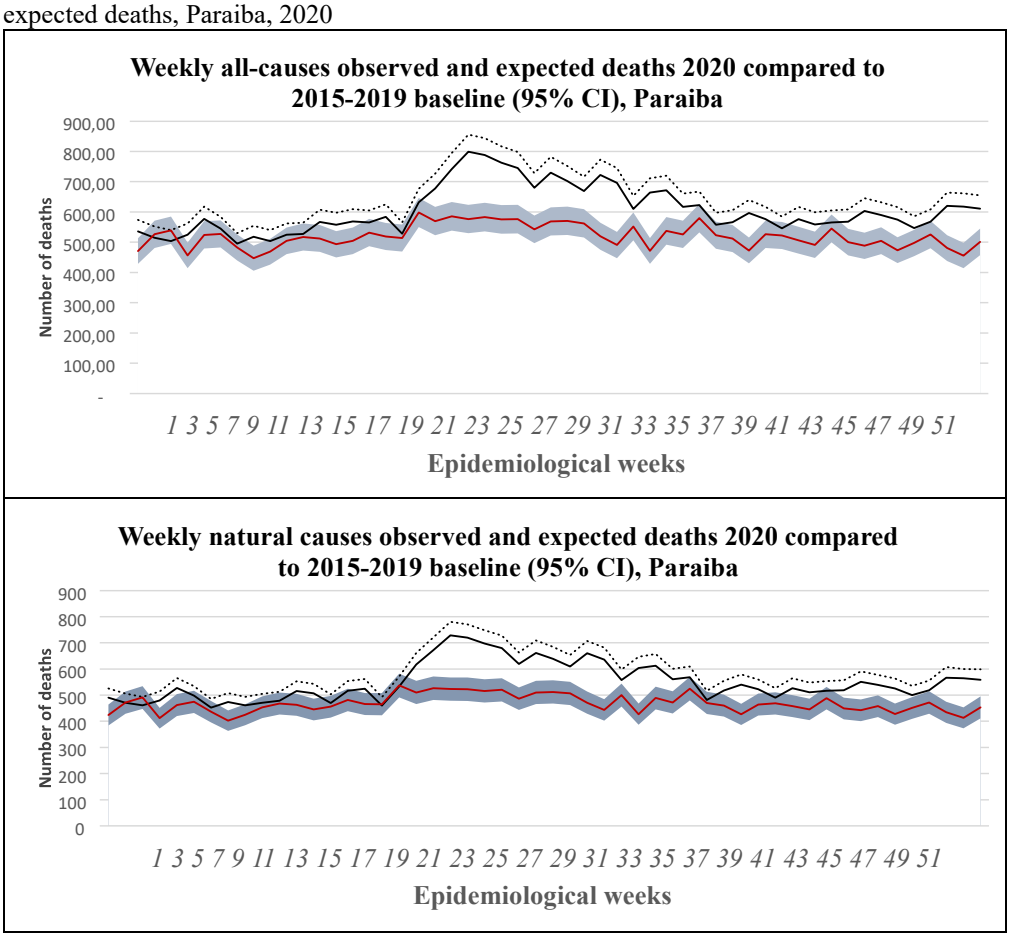
	2015-2019 (Baseline)	**2020**	**Excess of deaths**		
	N (95% CI)	N	N	P-score (%)	p-value
Observed deaths					
All-causes					
<60-years-old	8226 [6946, 9506]	9768	1545	18.7	<.0001
≥60-years-old	18827 [16890, 20764]	21749	2992	15.9	<.0001
Natural causes					
<60-years-old	6069 [4970, 7168]	7678	1609	26.5	<.0001
≥60-years-old	18311 [16401, 20221]	20898	2587	14.1	<.0001
Resp					
<60-years-old	483 [175, 790]	1404	921	190.6	<.0001
≥60-years-old	2717 [1982, 3454]	4832	2115	77.8	<.0001

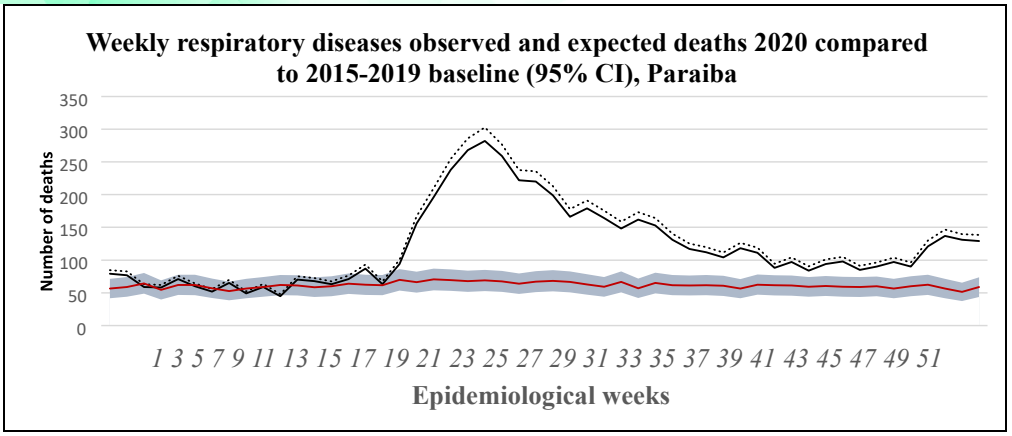
Expected deaths[a]					
All-causes					
<60-years-old	8226 [6946, 9506]	10479	2253	27.3	<.0001
≥60-years-old	18827 [16890, 20764]	23332	4505	23.9	<.0001
Natural causes					
<60-years-old	6069 [4970, 7168]	8240	2171	35.7	<.0001
≥60-years-old	18311 [16401, 20221]	22419	4108	22.4	<.0001
Resp					
<60-years-old	483 [175, 790]	1506	1023	211.8	<.0001
≥60-years-old	2717 [1982, 3454]	5184	2467	90.7	<.0001

[a]- deaths after quality data correction by application of General Growth Balance (GGB) method
Legend: Resp- respiratory system diseases, Source: SIM-Mortality Information System, 2020

For instance, looking at the historical baseline for respiratory diseases, the magnitude of excess deaths for <60-years-old group was at least 3 times higher than was seen in previous 5 years. The number of expected deaths increased for this particular group in 27.3%, 35.7% and 212% (all-causes, natural causes and respiratory diseases deaths, respectively).

Figure 1 - Excess deaths due all-causes, natural causes and respiratory system diseases compared to baseline considering observed and expected deaths, Paraiba, 2020





4 DISCUSSION AND FINAL CONSIDERATIONS

The incompleteness of COVID-19 deaths reporting, untested cases, time lags, overwhelmed health care capacities, among other issues, may create variations in death counts, so they do not represent real magnitude of mortality. Estimation of excess mortality proved to be a more accurate measure of direct and indirect impact of pandemic to population's mortality [8].

The full impact of the 2020 pandemic in Paraiba has been much greater than what is indicated by reported deaths due to COVID-19 alone. The study proved statistically significant excess of deaths in the state for all-cause mortality and natural causes, with approximately 2-fold higher number for respiratory diseases. The quality data correction added significant fractions on top of initially observed deaths which makes the scope of exceeding deaths in Paraiba even larger.

Even though confirmed COVID-19 deaths present a higher difference in excess deaths compared to historical baseline, about one-fourth of the excess deaths have been attributed to indirect impact of pandemic. It is likely that some of these deaths were COVID-19 cases that were miscoded or misdiagnosed. Many studies ultimately show that around one-third of the excess deaths that occurred at different times throughout the COVID-19 pandemic are not assigned to virus [23].

Analysis of age and sex-adjusted deaths in Paraiba during 2020 pandemic, proved statistically significant excess of deaths in all observed demographic characteristics: males and females, group under 60-years-old and 60-years-old and higher for all-cause mortality, natural causes and respiratory diseases. The historical trend of the male's prevalence among all deaths had a larger magnitude during 2020 pandemic.

Despite the significant prevalence of total number of deaths in group ≥60-years-old, the proportion of excess deaths was higher for group under 60-years-old. It proves that not only elderly people were affected by COVID-19 pandemic, but younger people also died and probably were underreported or misdiagnosed.

Northeast region of Brazil, including Paraiba, had shown recent improvements in a quality of vital events registration. However, the studies of the mortality information quality in region and

Paraiba are still severely limited [4,5]. This study has demonstrated that any analysis of mortality and excess mortality requires understanding of data quality issues. The analysis shown significant number of underreported deaths due COVID-19 after the General Growth Balance method application. One of the main contributions of this study is that for excess mortality estimation was considered mortality data quality issues which proved that deaths count released by Brazilian official mortality system do not fully represent the real burden of pandemic.

5 CONCLUSIONS

In conclusion, there is strong evidence that the full impact of the 2020 pandemic in Paraiba has been much greater than what is indicated by reported deaths due to COVID-19 alone. It is noteworthy that the methodology applied to the state of Paraiba can be replicated for any Brazilian state.

REFERENCES

United Nations. United Nations Statistics Division. Principles and Recommendations for a Vital Statistics System. Revision 3. Department of Economic and Social Affairs Statistics Division Statistical Papers, Series M No. 19/Rev.3. New York: United Nations; 2014.

COVID-19 Excess Mortality Collaborators. Estimating excess mortality due to the COVID-19 pandemic: a systematic analysis of COVID-19-related mortality, 2020-21. Lancet. 2022 Apr 16;399(10334):1513-1536. doi: 10.1016/S0140-6736(21)02796-3. Erratum in: Lancet. 2022 Apr 16;399(10334):1468. PMID: 35279232; PMCID: PMC8912932.

Cimerman S, Chebabo A, Cunha CAD, Rodriquez-Morales AJ. Deep impact of COVID-19 in the healthcare of Latin America: the case of Brazil. Braz J Infect Dis. 2020;24(2):93-95.

Paes NA. Quality of death statistics by unknown causes in Brazilian states. Rev Saúde Pública. 2007;41(3):436-445.

Paes NA, Santos CSA, Coutinho TDF. Quality of children's death records for regionalized spaces: a methodological route. Rev. bras. Epidemiol. 2021;24. Access in: 20 Jun. 2021.

Teixeira RA, et al. Quality of cause-of-death data in Brazil: Garbage codes among registered deaths in 2000 and 2015. Revista Brasileira de Epidemiologia. 2019;22(3). Available from: https://doi.org/10.1590/1980-549720190002.supl.3. Access in: 14 Jun. 2021.

Banerjee A, Pasea L, Harris S, Gonzalez-Izquierdo et al. Estimating excess 1-year mortality associated with the COVID-19 pandemic according to underlying conditions and age: a population-based cohort study. Lancet. 2020;395(10238):1715–1725.

CDC. Centers for Diseases Control and Prevention. Excess Deaths Associated with COVID-19.2020. Available from: https://www.cdc.gov/nchs/nvss/vsrr/covid19/excess_deaths.htm#techNotes. Access in: 14 May 2021.

WHO. The true death toll of COVID-19: estimating global excess mortality. May, 2021. Available from: https://www.who.int/data/stories/the-true-death-toll-of-covid-19- estimating-global-excess-mortality

Stokes AC, Lundberg DJ, Elo IT, Hempstead K, Bor J, Preston SH. COVID-19 and excess mortality in the United States: a county-level analysis. PLoS Med. 2021;18.

Kontis V, Bennett JE, Rashid T, et al. Magnitude, demographics and dynamics of the effect of the first wave of the COVID-19 pandemic on all-cause mortality in 21 industrialized countries. Nat Med. 2020; 26:1919–1928.

IBGE. Instituto Brasileiro de Geografia e Estatística. IBGE cidades. 2020. Available from: https://cidades.ibge.gov.br/

BRASIL. Ministério da Saúde. Sistema de informação sobre mortalidade - SIM. Apresentação. Secretaria de Vigilância em Saúde, Brasília. 2018. Available from: http://svs.aids. gov.br/dantps/cgiae/sim/apresentacao/

Paraíba. Governo do Estado da Paraíba. Paraíba tem primeiro óbito confirmado por Covid-19. Notícias. 2020. Available from: https://Paraíba.pb.gov.br/noticias/Paraíba- tem-primeiro-obito-confirmado-por-covid-19.

BRASIL. Ministério da Saúde. Orientações para o preenchimento da Declaração de Óbito no contexto da COVID-19. Secretaria de Vigilância em Saúde, Departamento de Análise em Saúde e Vigilância de Doenças não Transmissíveis, Coordenação Geral de Informações e Análise Epidemiológica, Brasília. 2020. Available from: https://saude.rs.gov.br/upload/arquivos/202005/06141402-nt-med-covid-04-05-2020- final.pdf

Paes NA, Ferreira AMS, Moura L de A. Proposta metodológica para avaliação de registros de óbitos por COVID-19. Cad Saúde Pública. 2023;39(1): e00096722. Available from: https://doi.org/10.1590/0102-311XPT096722

Brass W. Methods for estimating fertility and mortality from limited and defective data. International Program of Laboratories for Population Statistics. Chapel Hill: International Program of Laboratories for Population Statistics/University of North Carolina; 1975.

Paes NA. Demografia Estatística dos Eventos Vitais, 2th edition. João Pessoa: CCTA; 2018.

Naghavi M. et al. Algorithms for enhancing public health utility of national causes-of- death data. Popul Health Metr. 2010; 8:9. Available from: https://doi.org/10.1186/1478- 7954-8-9

Zhou F. et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. The Lancet (London, England). 2020;395(10229):1054-1062. Available from: https://doi.org/10.1016/S0140- 6736(20)30566-3

Leddermann S. La répartition de decés de causa indeterminée. Rev. Inst. Int. Stat. 1955;23(1):47-57. Available from: https://doi.org/10.2307/1402010.

Woolf SH, Chapman DA, Sabo RT et al. Excess Deaths From COVID-19 and Other Causes in the US, March 1, 2020, to January 2, 2021. JAMA. 2021;325(17):178