Assessment of Underreporting of Breast Cancer Mortality in Northeastern Brazil Over 40 Years

https://doi.org/10.32635/2176-9745.RBC.2024v70n4.4792

Avaliação dos Sub-registros da Mortalidade por Câncer de Mama no Nordeste do Brasil ao Longo de 40 Anos

Evaluación de los Sub-registros de Mortalidad por Cáncer de Mama en el Nordeste de Brasil a lo Largo de 40 Años

Juliana Dantas de Araújo Santos Camargo[1]; Sávio Ferreira Camargo[2]; Ana Karla Monteiro Santana de Oliveira Freitas[3]; Taynãna Cesar Simões[4]; Flávio Henrique Miranda de Araújo Freire[5]; Marcos Roberto Gonzaga[6]; Karina Cardoso Meira[7]

ABSTRACT

Introduction: Health planning and evaluation are compromised by poor-quality mortality data. **Objective:** To assess the stages of correction for breast cancer death records in the Northeast Region from 1980 to 2019. **Method:** Ecological study of breast cancer deaths among women aged 20 and over, residing in the States of the Northeast Region between 1980 and 2019. Data from the Mortality Information System of the Department of Informatics of the National Health System (SIM/DATASUS) were used. Corrections were made for unknown age, ill-defined causes, and incomplete cancer diagnoses. Proportional redistribution was carried out according to year, age group, and State. Underreporting was corrected using the extinct generations method. Mortality rates were calculated by age group and standardized using the direct method. Friedman tests and multiple comparisons with Bonferroni correction were used to assess differences in mortality rates across correction stages. **Results:** The average rate was 11.91/100 thousand women. A 61% increase (19.19/100 thousand) was observed after the correction stages. The greatest increase after corrections was in Maranhão (97%), and the smallest was in Pernambuco (26%). The highest average rates were in Pernambuco (19.99/100 thousand) and Ceará (19.33/100 thousand), while the lowest were in Maranhão (11.99/100 thousand) and Piauí (14.03/100 thousand). Significant differences were found in all localities between uncorrected and corrected breast cancer mortality rates ($p < 0.01$). **Conclusion:** After corrections, significant changes in breast cancer mortality rates were observed in all States of the Northeast. The greatest increases occurred in States with the poorest socioeconomic conditions, highlighting the importance of data correction.

Key words: Mortality Registries; Breast Neoplasms/mortality; Underreporting; Data Accuracy.

RESUMO

Introdução: Planejamento e avaliação em saúde são prejudicados por informações de mortalidade de má qualidade. **Objetivo:** Avaliar as etapas de correção dos registros de óbitos por câncer de mama na Região Nordeste no período de 1980 a 2019. **Método:** Estudo ecológico dos óbitos por câncer de mama em mulheres residentes nos Estados da Região Nordeste, a partir dos 20 anos, entre 1980 e 2019. Foram utilizados dados do Sistema de Informação de Mortalidade do Departamento de Informática do Sistema Único de Saúde (SIM/DATASUS). Correções foram feitas para idade ignorada, causa mal definida e diagnóstico incompleto de câncer. Realizada redistribuição proporcional segundo ano, faixa etária e Estado. Sub-registros corrigidos com gerações extintas ajustadas. Taxas de mortalidade calculadas segundo faixa etária e padronizadas pelo método direto. Testes de Friedman e comparações múltiplas com correção de Bonferroni para diferenças nas taxas de mortalidade entre etapas de correção. **Resultados:** Taxa média de 11,91/100 mil mulheres. Aumento de 61% (19,19/100 mil) após as etapas de correção. Maior incremento após as correções: Maranhão (97%). Menor: Pernambuco (26%). Maiores taxas médias: Pernambuco (19,99/100 mil) e Ceará (19,33/100 mil). Menores: Maranhão (11,99/100 mil) e Piauí (14,03/100 mil). Houve diferenças significativas em todas as localidades entre as taxas de mortalidade por câncer de mama sem correção e após a aplicação das etapas de correção para qualidade da informação e inclusão da cobertura ($p < 0,01$). **Conclusão:** Após correções, houve alterações significativas nas taxas de mortalidade por câncer de mama em todos os Estados do Nordeste. Maiores incrementos ocorreram nos Estados com piores condições socioeconômicas. Assim, ficou evidenciada a relevância da correção.

Palavras-chave: Registros de Mortalidade; Neoplasias da Mama/mortalidade; Sub-Registro; Confiabilidade dos Dados.

RESUMEN

Introducción: La planificación y evaluación en salud se ven perjudicadas por información de mortalidad de mala calidad. **Objetivo:** Evaluar las etapas de corrección de los registros de defunciones por cáncer de mama en la región Nordeste en el período de 1980 a 2019. **Método:** Estudio ecológico de las defunciones por cáncer de mama en mujeres residentes en los estados de la región Nordeste, a partir de los 20 años, entre 1980 y 2019. Los datos fueron tomados del Sistema de Información de Mortalidad del Departamento de Informática del Sistema Único de Salud (SIM/DATASUS). Se realizó corrección para edad ignorada, causa mal definida y diagnóstico incompleto de cáncer. Se hizo redistribución proporcional según año, grupo de edad y estado. Sub-registros fueron corregidos con el método de generaciones extintas ajustado. Tasas de mortalidad fueron calculadas según grupo de edad y estandarizadas por el método directo. Pruebas de Friedman y comparaciones múltiples con corrección de Bonferroni fueron usadas para evaluar diferencias en las tasas de mortalidad entre etapas de corrección. **Resultados:** Tasa media de 11,91/100 000 mujeres. Aumento del 61% (19,19/100 000) tras las etapas de corrección. Mayor incremento tras las correcciones: Maranhão (97%). Menor: Pernambuco (26%). Mayores tasas medias: Pernambuco (19,99/100 000) y Ceará (19,33/100 000). Menores: Maranhão (11,99/100 000) y Piauí (14,03/100 000). Hubo diferencias significativas en todas las localidades entre las tasas de mortalidad por cáncer de mama sin corrección y tras la aplicación de las etapas de corrección para la calidad de la información e inclusión de la cobertura ($p < 0,01$). **Conclusión:** Tras las correcciones, hubo cambios significativos en las tasas de mortalidad por cáncer de mama en todos los estados del Nordeste. Los mayores incrementos se dieron en los estados con peores condiciones socioeconómicas. Así, se evidenció la relevancia de la corrección.

Palabras clave: Registros de Mortalidad; Neoplasias de la Mama/mortalidad; Omisiones de Registro; Exactitud de los Datos.

[1,2]Universidade Federal do Rio Grande do Norte (UFRN), Maternidade Escola Januário Cicco, Empresa Brasileira de Serviços Hospitalares (Ebserh). Natal (RN), Brasil. E-mails: juliana_ily@hotmail.com; saviocamargo@gmail.com. Orcid iD: https://orcid.org/0000-0001-8692-5706; Orcid iD: https://orcid.org/0000-0002-5165-1976
[3]UFRN, Departamento de Tocoginecologia. Natal (RN), Brasil. E-mail: akfreitas7@gmail.com. Orcid iD: https://orcid.org/0000-0003-3907-9601
[4]Instituto René Rachou (Fiocruz Minas). Belo Horizonte (MG), Brasil. E-mail: taynanasimoes@ensp.fiocruz.br. Orcid iD: https://orcid.org/0000-0002-5849-343X
[5,6]UFRN, Programa de Pós-graduação em Demografia (PPGDEM). Natal (RN), Brasil. E-mails: flaviohfreire@gmail.com; mrcs.roberto@gmail.com. Orcid iD: https://orcid.org/0000-0002-7416-9947; Orcid iD: https://orcid.org/0000-0002-6088-3453
[7]UFRN, PPGDEM. Natal (RN), Brasil. Universidade Federal de São Paulo (Unifesp), Escola Paulista de Enfermagem. São Paulo (SP), Brasil. E-mail: karina.meira@unifesp.br. Orcid iD: https://orcid.org/0000-0002-1722-5703

This article is part of a master's thesis titled "Temporal evolution of breast cancer mortality in the States of the Northeast Region from the perspective of age, period and cohort effects".

Corresponding author: Juliana Dantas de Araújo Santos Camargo. Av. Nilo Peçanha, 259 – Petrópolis. Natal (RN), Brasil. CEP 59012-310. E-mail: juliana_ily@hotmail.com

INTRODUCTION

The Mortality Information System (SIM), managed by the Ministry of Health[1], provides the basic cause of death according to the tenth review of the International Classification of Diseases and Related Health Problems (ICD-10)[2]. It started being used in 1975 and, over the last decades, gradually evolved regarding the quality of the records. This quality assessment is the object of several studies that aim at measuring both the coverage levels and the accuracy of data regarding the basic cause of death[3-5].

The high percentages of under reports and reports with unknown basic cause of deaths may compromise the quality of mortality information, impairing the planning and management of health services. Those high percentages are believed to be negatively related to health, social and economic conditions and are observed exactly in the least developed regions of the country, the North and Northeast regions[6,7].

In addition to issues related to underreporting, the gaps in mortality assessment related to basic cause of death are fairly common, representing an issue in the distribution of deaths due to a specific cause. Two classifications have been broadly assessed: poorly defined causes and incomplete diagnoses. However, the last classification allows the allocation of death within a group of similar causes in a chapter of ICD-10, reflecting a minor loss of information[8].

There is still no consensus in the literature on which is the best methodology for assessing the quality of mortality data, due to specific limitations and advantages of each method. Moreover, using different techniques according to the data in research makes it harder to compare methods[5,9]. Despite the controversies found in several studies on correction methods, it's a known fact that one of the techniques should be applied to improve the quality of the report in studies in which SIM was used as data source. Many studies have demonstrated the importance of correction, mainly for the 1980 and 1990 decades and in the North and Northeast Regions[5,6].

Since breast cancer is a disease in which high mortality rates predominate in old age[10,11], the application of methods to assess the quality of information may improve those estimates. Thus, the objective of this study is to assess the stages of correction for breast cancer death records in the States of the Northeast Region from 1980 to 2019.

METHOD

The studied population is comprised of women living in the Northeast States, aged 20 and over, that died from 1980 to 2019 and had breast cancer as their basic cause of death. The 20 and onwards age group was chosen due to the low incidence of breast cancer before that age.

The death records were obtained from SIM, managed by the Ministry of Health[1], and the records of interest were those classified with the codes 174, referring to female breast neoplasm in the ninth review of the ICD (ICD-9) and C50, referring to breast neoplasm in the tenth review of the ICD (ICD-10) [2].

The population data were obtained from the Department of Informatics of the National Health System (DATASUS)[12], that come from the *Instituto Brasileiro de Geografia e Estatística* (IBGE). Values referring to the years 1980, 1991, 2000 and 2010 were obtained through the respective census, and the intercensal periods, in addition to the 2011-2019 period, were calculated by IBGE through population projections for July 1st[12].

The data correction process was composed of the following steps: i) proportional redistribution of the records classified as ignored age; ii) proportional redistribution of 50% of ill-defined death causes among all the death causes, except for external causes; iii) proportional redistribution of deaths by neoplasm with incomplete diagnosis among the cancer types constituted by the codes: malignant neoplasms of ill-defined, secondary and unspecified sites: 195 (CID-9) and C76 (CID-10); secondary and unspecified malignant neoplasm of lymph nodes: 196 (CID-9) and C77 (CID-10); secondary malignant neoplasm of respiratory and digestive organs: 197 (CID-9) and C78 (CID-10); secondary malignant neoplasm of other and unspecified sites: 198 (CID-9) and C79 (CID-10); malignant neoplasm, without specification of site: 199 (CID-9) and C80 (CID-10); malignant neoplasms of independent (primary) multiple sites: C97 (CID-10); these deaths were redistributed among all the cancers according to the proportion related to the Chapter II — Neoplasms; and iv) coverage assessment of the mortality records of each State and correction of under-report.

The fourth step in the process was performed using the correction factors of the 1980-2010 period found through the method of extinct generations adjusted by Queiroz et al.[5]. The three main methods for assessing coverage in the death records were discussed in this work: General Growth Balance (GGB), proposed by Hill[13], the Synthetic Extinct Generation (SEG), proposed by Bennett and Horiuchi[14], and the Adjusted Synthetic Extinct Generations (SEG+GGB), proposed by Hill et al.[13]. The GGB method is based on the fundamental demographic balancing equation, which states the growth

rate as the difference between the increase and decrease rates of a population. This relationship also applies to any age group with an open interval from x years (people with x years or more)[13,15,16].

The Synthetic Extinct Generation[14] (SEG) method is also detailed in the study by Queiroz et al.[5]. This method uses the specific growth rates by age group to change a death distribution by age to an age distribution of the population. Although the authors have demonstrated a good performance of both techniques, the robustness of the combined methods (SEG+GGB) is the main motivator for using it[5].

The specific mortality rates by age group and according to five-year periods were calculated dividing the number of breast cancer deaths by the female population on July 1st and multiplying the result by 100 thousand. The direct method was applied to standardize the rates, using the world population as the standard, proposed by Sergi and modified by Doll and Payne[17]. The Friedman test was executed to determine if there were differences in the mortality rates obtained between the correction stages. Pairs comparison was performed[16] with a Bonferroni correction for multiple comparisons. The 5% level of significance was adopted in the analyses. The software used for analyses was the Statistical Package for the Social Sciences (SPSS)[18], version 28.0.

According to Resolution CNS 510/2016[19], studies with publicly available secondary data are exempt from ethical analysis.

RESULTS

In the analyzed period, all the locations showed increase in the death records and mortality rates after the correction steps. In the Northeast Region, during the 1980-1984, 2,541 breast cancer deaths were recorded, and this number increased to 18,299 records in the 2015-2019 period. After the correction of records, these numbers correspond to 5,171 and 24,334 deaths, indicating percentage increases of 104% and 33%, in the periods of 1980-1984 and 2015-2019, respectively. The results exhibited in Tables 1 and 2 refer to the mortality rates by breast cancer standardized and obtained in each record correction step for each location (Northeast Region and States).

After correcting death records, it is possible to see increases in breast cancer mortality rates throughout the period, 1980 to 2019, and for all locations. Comparing the 1980-1984 and 2015-2019 periods, it is possible to identify that most percentage variations happened in the first period, with the States of Maranhão and Piauí holding the greatest variations, 173% and 155%, respectively. Pernambuco, Rio Grande do Norte and Bahia have the smaller increase percentages, 47%, 51% and 53%, respectively (Tables 1 and 2).

It is possible to identify the gradual increase in the mortality rates from each correction step. Those differences are representative for some States when the coverage factor is applied. Increases of 60% and 29% were verified in Maranhão and Piauí, respectively, in average mortality rates, after correction for underreporting of death records. The State of Pernambuco obtained the smallest percentage gain after death records correction, with 26% of increase; its mortality rate increased from 15.92/100 thousand to 19.99/100 thousand. Maranhão presented the greatest percentage variation after applying the record correction steps (97%) (Figure 1).

Aiming to compare correction steps to identify the existence of differences in mortality rates after applying death records correction steps, the Friedman test was performed. Test results showed statistical significance for every location ($p < 0.05$), thus, multiple comparisons *post hoc* tests were applied to identify the steps that actually caused significant changes in the standardized mortality rates and are shown in Table 3.

Significant differences ($p < 0.01$) in breast cancer mortality rates were observed in all locations when comparing the non-corrected rates and the two steps called: breast cancer correction for information quality (BCCIA + ID + IDC) and total correction of breast cancer including coverage (BCCIA + ID + IDC +DC) (Table 3).

DISCUSSION

The breast cancer mortality rate record correction in the Northeast Region from 1980 to 2019 recorded a 61% increase (11.91/100 thousand women *vs.* 19.91/100 thousand women). The greatest increases were verified in the States of Maranhão (97%) and Piauí (50%) and the smallest in Pernambuco (26%). Furthermore, these differences were significant between breast cancer mortality rates without correction and after applying the correction steps for information quality and inclusion of coverage in all locations.

The substantial increase in deaths shown in the present study, after the correction steps, reaffirms the need for applying indirect death correction techniques in research that uses SIM records[1] to compare distinct locations in a long period of time, above all when studying less developed Regions using the death records from the 1980s and 1990s[5,6].

The greatest correction percentages observed in the less developed States in the Region, Maranhão and Piauí,

Table 1. Standardized breast cancer mortality rates before and after the death record correction steps in the Northeast Region and States of Alagoas, Bahia, Ceará and Maranhão, in five-year periods, from 1980 to 2019 (Natal, Brazil, 2021)

Northeast						
Periods	**NCBC**	**BCCIA**	**BCCIA + ID**	**BCCIA + IDC**	**BCCIA + ID + IDC**	**BCCIA + ID + IDC +DC**
1980-1984	7.16	7.21	7.67	10.02	10.58	14.70
1985-1989	7.86	7.92	8.45	10.83	11.47	15.95
1990-1994	8.69	8.79	9.38	11.46	12.25	15.92
1995-1999	9.68	9.95	10.49	11.96	13.05	16.97
2000-2004	11.33	11.36	12.53	13.34	14.55	18.63
2005-2009	14.95	14.96	16.35	15.74	17.16	21.96
2010-2014	16.87	16.88	18.26	17.51	18.91	24.20
2015-2019	18.76	18.76	19.37	19.37	19.70	25.22
Alagoas						
Periods	**NCBC**	**BCCIA**	**BCCIA + ID**	**BCCIA + IDC**	**BCCIA + ID + IDC**	**BCCIA + ID + IDC +DC**
1980-1984	7.95	7.99	8.42	11.57	12.07	12.31
1985-1989	7.23	7.26	7.77	10.23	10.79	11.03
1990-1994	7.71	7.77	8.26	10.53	11.15	11.56
1995-1999	7.15	7.23	7.98	9.38	10.28	10.67
2000-2004	8.99	8.99	10.13	10.84	11.98	13.70
2005-2009	13.71	13.71	15.24	14.44	15.96	18.20
2010-2014	15.88	15.88	17.41	16.47	18.00	20.53
2015-2019	16.10	16.10	16.43	17.17	17.50	19.95
Bahia						
Periods	**NCBC**	**BCCIA**	**BCCIA + ID**	**BCCIA + IDC**	**BCCIA + ID + IDC**	**BCCIA + ID + IDC +DC**
1980-1984	7.77	7.84	8.39	10.08	10.76	11.85
1985-1989	8.76	8.83	9.33	11.06	11.7	12.87
1990-1994	8.97	9.02	9.59	11.09	11.75	13.52
1995-1999	9.41	9.46	10.24	11.36	12.23	14.05
2000-2004	10.5	10.51	11.53	12.33	13.36	15.5
2005-2009	13.12	13.13	14.29	14.39	15.57	18.06
2010-2014	16.08	16.09	17.48	17.23	18.63	21.62
2015-2019	17.60	17.60	18.83	20.16	20.16	23.39
Ceará						
Periods	**NCBC**	**BCCIA**	**BCCIA + ID**	**BCCIA + IDC**	**BCCIA + ID + IDC**	**BCCIA + ID + IDC +DC**
1980-1984	7.41	7.46	7.89	9.91	10.45	13.60
1985-1989	7.92	7.96	8.57	10.89	11.58	15.06
1990-1994	8.55	8.59	9.50	11.50	12.50	14.86
1995-1999	11.51	11.63	12.65	13.85	15.11	17.98
2000-2004	13.99	14.07	15.87	16.19	18.14	20.68
2005-2009	17.05	17.05	18.93	17.75	19.62	22.38
2010-2014	18.30	18.30	19.82	18.81	20.33	23.17
2015-2019	21.82	21.82	22.26	23.18	23.61	26.92
Maranhão						
Periods	**NCBC**	**BCCIA**	**BCCIA + ID**	**BCCIA + IDC**	**BCCIA + ID + IDC**	**BCCIA + ID + IDC +DC**
1980-1984	2.34	2.36	2.54	3.18	3.39	6.39
1985-1989	3.14	3.18	3.53	4.48	4.91	9.28
1990-1994	3.53	3.56	3.81	5.21	5.51	8.39
1995-1999	3.91	3.95	4.10	5.43	5.66	8.60
2000-2004	4.33	4.34	4.64	5.77	6.09	9.50
2005-2009	8.38	8.39	8.89	8.84	9.35	14.59
2010-2014	10.89	10.9	11.55	11.23	11.89	18.55
2015-2019	12.11	12.11	12.38	12.92	13.20	20.59

Captions: NCBC = non-corrected breast cancer; BCCIA = breast cancer correction for ignored age; BCCIA + ID = breast cancer correction for ignored age and incomplete diagnosis; BCCIA + IDC = breast cancer correction for ignored age and ill-defined causes; BCCIA + ID + IDC = breast cancer correction for information quality; and BCCIA + ID + IDC +DC = total correction of breast cancer including coverage.

Table 2. Standardized breast cancer mortality rates before and after the death record correction steps in the States of Paraíba, Pernambuco, Piauí, Rio Grande do Norte and Sergipe, in five-year periods, from 1980 to 2019 (Natal, Brazil, 2021)

Paraíba						
Periods	**NCBC**	**BCCIA**	**BCCIA + ID**	**BCCIA + IDC**	**BCCIA + ID + IDC**	**BCCIA + ID + IDC +DC**
1980-1984	5.40	5.45	6.09	10.05	10.79	11.22
1985-1989	5.25	5.26	5.84	9.10	9.69	10.08
1990-1994	6.59	6.61	7.44	10.38	11.26	11.98
1995-1999	6.07	6.13	6.70	9.34	10.04	10.69
2000-2004	8.51	8.52	9.78	10.97	12.25	13.70
2005-2009	15.16	15.17	16.83	15.91	17.58	19.70
2010-2014	15.66	15.66	17.15	16.17	17.65	19.74
2015-2019	16.84	16.84	17.27	18.25	18.25	20.92
Pernambuco						
Periods	**NCBC**	**BCCIA**	**BCCIA + ID**	**BCCIA + IDC**	**BCCIA + ID + IDC**	**BCCIA + ID + IDC +DC**
1980-1984	10.01	10.05	10.59	14.11	14.73	14.76
1985-1989	10.94	11.00	11.69	15.07	15.87	15.88
1990-1994	12.77	13.02	13.60	16.45	17.54	17.90
1995-1999	15.16	15.22	16.20	17.65	18.75	19.13
2000-2004	16.40	16.40	17.93	18.19	19.72	20.70
2005-2009	19.68	19.69	21.42	20.26	22.01	23.09
2010-2014	20.19	20.19	21.80	20.62	22.24	23.36
2015-2019	22.19	22.19	22.54	23.61	23.96	25.16
Piauí						
Periods	**NCBC**	**BCCIA**	**BCCIA + ID**	**BCCIA + IDC**	**BCCIA + ID + IDC**	**BCCIA + ID + IDC +DC**
1980-1984	5.42	5.48	5.72	7.39	7.75	13.84
1985-1989	3.78	3.84	4.05	5.53	5.86	10.38
1990-1994	5.09	5.15	5.39	6.82	7.18	8.53
1995-1999	4.84	4.85	5.15	6.01	6.34	7.51
2000-2004	8.62	8.65	9.29	10.23	10.94	12.71
2005-2009	13.32	13.32	14.24	13.81	14.73	17.09
2010-2014	15.93	15.93	16.92	16.25	17.25	19.99
2015-2019	17.80	17.80	18.14	18.75	16.94	22.15
Rio Grande do Norte						
Periods	**NCBC**	**BCCIA**	**BCCIA + ID**	**BCCIA + IDC**	**BCCIA + ID + IDC**	**BCCIA + ID + IDC +DC**
1980-1984	5.74	5.77	6.06	7.69	8.03	8.68
1985-1989	9.52	9.56	10.11	12.55	13.18	14.41
1990-1994	10.63	10.67	11.50	13.45	14.35	16.11
1995-1999	10.97	11.17	12.01	13.29	14.52	16.23
2000-2004	11.51	11.56	12.76	13.28	14.57	16.89
2005-2009	15.05	15.05	16.94	15.35	17.24	20.00
2010-2014	17.53	17.53	19.49	17.79	19.75	22.88
2015-2019	20.07	20.07	20.37	21.79	22.09	25.63
Sergipe						
Periods	**NCBC**	**BCCIA**	**BCCIA + ID**	**BCCIA + IDC**	**BCCIA + ID + IDC**	**BCCIA + ID + IDC +DC**
1980-1984	7.80	7.84	8.08	11.92	12.24	12.58
1985-1989	8.55	8.76	8.91	13.04	13.61	13.84
1990-1994	9.48	10.00	10.08	13.05	14.16	14.94
1995-1999	9.64	10.28	10.29	12.35	13.63	14.5
2000-2004	13.49	13.50	14.46	15.64	16.62	18.25
2005-2009	17.44	17.44	18.47	18.13	19.17	21.08
2010-2014	20.94	20.94	21.96	21.56	22.58	24.87
2015-2019	22.25	22.25	23.08	23.15	23.99	26.38

Captions: NCBC = non-corrected breast cancer; BCCIA = breast cancer correction for ignored age; BCCIA + ID = breast cancer correction for ignored age and incomplete diagnosis; BCCIA + IDC = breast cancer correction for ignored age and ill-defined causes; BCCIA + ID + IDC = breast cancer correction for information quality; and BCCIA + ID + IDC +DC = total correction of breast cancer including coverage.

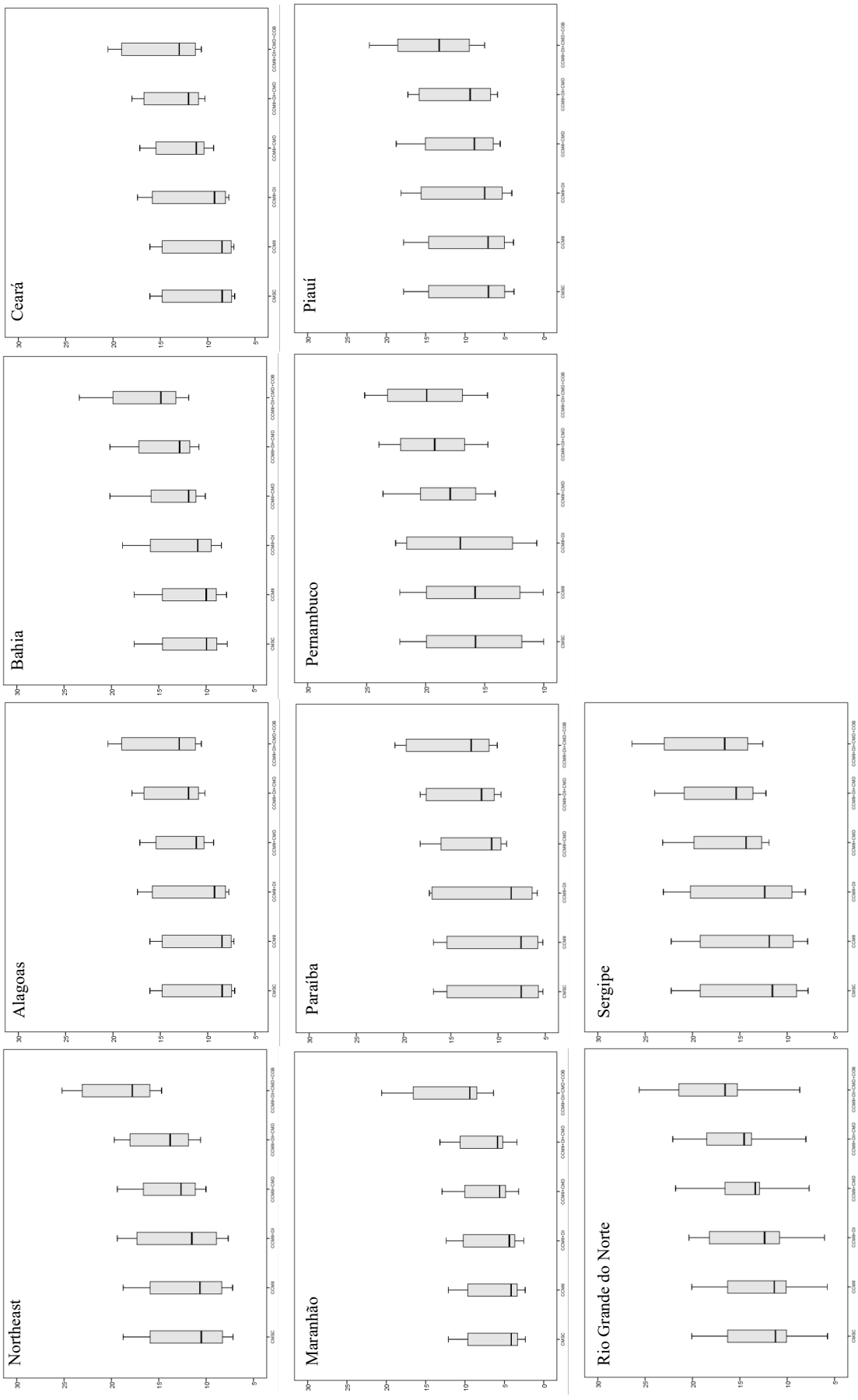

Figure 1. Distribution of mortality rates by breast cancer in the Northeast Region and respective States, period of 1980-2019, according to record correction steps
Captions: NCBC = non-corrected breast cancer; BCCIA = breast cancer correction for ignored age; BCCIA + ID = breast cancer correction for ignored age and incomplete diagnosis; BCCIA + IDC = breast cancer correction for ignored age and ill-defined causes; BCCIA + ID + IDC = breast cancer correction for information quality; and BCCIA + ID + IDC +DC = total correction of breast cancer including coverage.

Table 3. Multiple comparison tests to assess the death records correction steps in the Northeast Region and respective States, in five-year periods, from 1980 to 2019 (Natal, Brazil, 2021)

Multiple comparisons*	Location									
	Northeast	Alagoas	Bahia	Ceará	Maranhão	Paraíba	Pernambuco	Piauí	Rio Grande do Norte	Sergipe
NCBC *vs.* BCCIA	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
NCBC *vs.* BCCIA + ID	0.242	0.488	0.412	0.412	0.290	0.347	0.412	0.412	0.412	0.412
NCBC *vs.* BCCIA + IDC	0.075	0.113	**0.032**	0.092	0.061	0.061	0.092	0.092	0.092	0.092
NCBC *vs.* BCCIA + ID + IDC	**$p < 0.01$**	**0.001**	**0.001**	**0.001**	**$p < 0.01$**	**0.001**	**0.001**	**0.010**	**0.001**	**0.001**
NCBC *vs.* BCCIA + ID + IDC +DC	**$p < 0.01$**	**$p < 0.01$**	**$p < 0.01$**	**$p < 0.01$**	**$p < 0.01$**	**$p < 0.01$**	**$p < 0.01$**	**$p < 0.01$**	**$p < 0.01$**	**$p < 0.01$**
BCCIA *vs.* BCCIA + ID	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
BCCIA *vs.* BCCIA + IDC	0.921	0.488	0.488	0.575	0.790	0.575	0.575	0.575	0.575	0.575
BCCIA *vs.* BCCIA + ID + IDC	**0.016**	**0.008**	**0.020**	**0.010**	**0.016**	**0.016**	**0.010**	0.092	**0.010**	**0.010**
BCCIA *vs.* BCCIA + ID + IDC +DC	**$p < 0.01$**	**$p < 0.01$**	**$p < 0.01$**	**$p < 0.01$**	**$p < 0.01$**	**$p < 0.01$**	**$p < 0.01$**	**$p < 0.01$**	**$p < 0.01$**	**$p < 0.01$**
BCCIA + ID *vs.* BCCIA + IDC	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
BCCIA + ID *vs.* BCCIA + ID + IDC	1.000	0.921	0.790	0.921	0.921	1.000	0.921	1.000	0.921	0.921
BCCIA + ID *vs.* BCCIA + ID + IDC +DC	0.061	**0.049**	**0.033**	**0.049**	**0.049**	**0.049**	**0.049**	0.075	**0.049**	**0.049**
BCCIA + IDC *vs.* BCCIA + ID + IDC	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
BCCIA + IDC *vs.* BCCIA + ID + IDC +DC	0.201	0.242	0.412	0.242	0.242	0.290	0.242	0.347	0.242	0.242

Captions: (*) *p* values were obtained through multiple comparison tests after Bonferroni corrections. The values in bold were significant to 5%; NCBC = non-corrected breast cancer; BCCIA = breast cancer correction for ignored age; BCCIA + ID = breast cancer correction for ignored age and incomplete diagnosis; BCCIA + IDC = breast cancer correction for ignored age and ill-defined causes; BCCIA + ID + IDC = breast cancer correction for information quality; and BCCIA + ID + IDC +DC = total correction of breast cancer including coverage.

corroborate the findings of this study that also assessed the quality of the death records in Brazil and its regions, indicating superior magnitude in the Regions that showed greater gaps in social, economic and health issues[6,7].

In Brazil, there is heterogeneity in the country's Regions regarding the quality of information in death records in the proportions classified as ill-defined causes. These percentages decreased from 27.1% to 8.5% from 1991 to 2010; however, the North and Northeast Regions still hold the greatest proportions of inaccurately classified death records. This explains why, in this study, after proportional redistribution of 50% of ill-defined causes, the greatest increase percentage in average mortality rates were verified in the State of Paraíba (26%) and the smallest in the States of Piauí (13%) and Rio Grande do Norte (14%). This is related to the difficulties in accessing health services and the bad quality of the care offered and contributes to the existence of low-quality mortality indicators[4]. Thus, more attention to training and awareness of medical professionals that work in the Northeast Region is needed regarding the importance of correctly fulfilling the basic cause of death.

Furthermore, on the need to rectify the SIM/DATASUS[1] records regarding quality and underreporting, studies developed by Gamarra et al. (2010)[3] and Santos-Silva[20] with deaths by cervix cancer in the Northeast Region States – the first, from 1996 to 2005, the second, 1980 to 2014 – showed that using the records with no proper correction may mislead conclusions regarding the time trend and the period effect due to information bias.

Thus, the indirect techniques for correcting breast cancer death records applied in this research have proven to be essential, considering that improving quality of SIM generates period effect on mortality and, thus, rise in its coefficients. Such rise, it should be stressed, may occur due to a better certification of deaths, and not due to increase in incidence and mortality by the neoplasm.

It is important to highlight some inherent limitations to this research. Studies with ecological designs do not allow the affirmation of analytical relationships, just the formation of hypotheses in the aggregated set of a population. However, they are important in epidemiological studies where great samples are used, favoring the creation of hypotheses related to risk factors and protection for illness development[21].

Other limitations refer to the quality and underreporting of the death records obtained from SIM/DATASUS[1] regarding three aspects: 1) coverage of adult mortality rate records; 2) redistribution of ill-defined causes and 3) age declaration. There are different demographic methods to assess the coverage of adult mortality records, though each has its limitations. In some methods – Brass (1975)[19] balancing equation and Preston et al.[23] – the main limitation is related to the population stability assumption. The latest methods GGB[13] and SEG[14] have the advantage of eliminating such assumption.

Moreover, the combination of these two methods, SEG + GGB[13], may be more robust than applying each method individually[2]. However, these methods are used for general mortality correction assuming that the distribution of mortality by a specific cause would present a similar and constant behavior. Studies aiming to solve this gap are still scarce, reason why general mortality correction factors estimated by Queiroz et al.[5] were used here.

Proportional redistribution of 50% of ill-defined causes among all known natural causes, excluding external causes[24], as proposed by the World Health Organization (WHO), is also an object of discussion. This methodology does not consider the fact that the probability of incomplete cause of death record may not present a constancy for each specific cause[25].

On the quality assessment of age declaration, there are strong evidence of incorrect notification of old age declarations in developing countries, affecting demographic estimates based, mainly, on the age distribution analysis or the mortality of populations. Such inaccuracies may generate lower mortality rate estimates in older ages[26]. Several studies highlight the importance of this investigation, since the occurrence of systematic memory errors is common, however, the studies in this area are not fully consolidated yet and there are many proposed methods to assess this declaration[27].

The theme addressed here is extremely broad and there are innumerable possibilities for future research. The development of cause-specific mortality correction factors still represents a challenge for demographers – this would be the first possibility. Another proposal would be the application of statistical imputation techniques for the distribution of ill-defined causes, such as multiple imputation and Approximate Bayesian Bootstrap (ABB), used by Silva Júnior[28], or even the statistical procedure proposed by Ledermann and applied by Paes and Gouveia[29].

Despite the referred limitations, this study represents a step forward with the application and comparison of indirect correction techniques both for quality and underreporting of death records, presenting more trustworthy mortality rates for an important public health issue of the Brazilian Northeast, given that breast cancer is the most frequent among women, excluding non-melanoma skin cancer.

CONCLUSION

The main results indicate significant changes in mortality rates by breast cancer in every Northeast State after correcting the death records, with the greatest increases being observed in the States with worse socioeconomic conditions. Thus, the relevancy of the correction due to the substantial increase verified after applying the procedures became proven.

The 1980 and 1990 decades presented high indexes of underreporting and ill-defined causes, mainly in the least developed States. From the corrections, mortality estimates approached a more reliable reality, allowing to identify existing disparities among the States in the Region, that is, the more developed States showed higher mortality rates. Thus, it is not recommended to use data from the Northeast Region extracted from SIM to evaluate mortality trends, especially from the 1980 and 1990 decades, without applying the adequate record correction techniques.

Considering the findings of this research, it is essential that governmental bodies and health professionals recognize the importance of carefully recording data on mortality by breast cancer, since these are indispensable tools for monitoring actions regarding primary and secondary prevention. Furthermore, mortality systems allow to assess the efficacy and distribution of tertiary prevention of breast cancer in a Region of Brazil with differentiated conditions of health services and treatments.

ACKNOWLEDGMENTS

To the *Empresa Brasileira de Serviços Hospitalares (Ebserh)* and the *Programa de Pós-Graduação em Demografia (PPGDEM)* of *Universidade Federal do Rio Grande do Norte (UFRN)*.

CONTRIBUTIONS

All the authors have substantially contributed to the study design, acquisition, analysis and interpretation of the data, wording, and critical review. They approved the final version for publication.

DECLARATION OF CONFLICT OF INTERESTS

There is no conflict of interest to declare.

FUNDING SOURCES

Coordination for the Improvement of High Education Personnel (Capes) – Brazil. Funding code 001.

REFERENCES

1. SIM: Sistema de Informação sobre Mortalidade [Internet]. Versão 3.2.1.2. Brasília (DF): DATASUS. [data desconhecida] - [acesso 2024 ago 3]. Disponível em: http://sim.saude.gov.br/default.asp
2. Organização Mundial da Saúde. CID-10: Classificação Estatística Internacional de Doenças e problemas relacionados à saúde. São Paulo: Edusp; 2008.
3. Gamarra CJ, Valente JG, Azevedo e Silva G. Correction for reported cervical cancer mortality data in Brazil, 1996-2005. Rev Saude Publica. 2010;44(4):629-38. doi: https://doi.org/10.1590/S0034-89102010000400006
4. Lima EEC, Queiroz BL. Evolution of the deaths registry system in Brazil: Associations with changes in the mortality profile, under-registration of death counts, and ill-defined causes of death. Cad Saude Publica. 2014;30(8):1721-30. doi: https://doi.org/10.1590/0102-311X00131113
5. Queiroz BL, De Araujo FreireI FHM, Gonzaga MR, De Lima EEC. Estimativas do grau de cobertura e da mortalidade adulta (45q15) para as unidades da federação no Brasil entre 1980 e 2010. Rev Bras Epidemiol. 2017;20(Supl 1):21-33. doi: https://doi.org/10.1590/1980-5497201700050003
6. Paes NA. Quality of death statistics by unknown causes in Brazilian States. Rev Saude Publica. 2007;41(3):436-45. doi: https://doi.org/10.1590/S0034-89102007000300016
7. Silva GA, Gamarra CJ, Girianelli VR, et al. Cancer mortality trends in Brazilian state capitals and other municipalities between 1980 and 2006. Rev Saude Publica. 2011;45(6):1009-18. doi: https://doi.org/10.1590/S0034-89102011005000076
8. Jorge MHPM, Gotlieb SLD, Laurenti R. O sistema de informações sobre mortalidade: problemas e propostas para o seu enfrentamento I - Mortes por causas naturais. Rev Bras Epidemiol. 2002;5(1):3-17. doi: http://dx.doi.org/10.1590/S1415-790X2002000200007
9. Lanza Queiroz B, Simões Agostinho C. Estimativas da mortalidade adulta para o brasil no período 1980/2000: uma abordagem metodológica comparativa [Internet]. 16 Encontro Nacional de Estudos Populacionais; 2008 Set 29-Out 3; Caxambu, MG, Brasil. Belo Horizonte: UFMG; 2008. [acesso 2024 ago 15]. Disponível em: https://www.researchgate.net/publication/239932426
10. Meira KC, Guimarães RM, Santos J, et al. Análise de efeito idade-período-coorte na mortalidade por câncer de mama no Brasil e regiões. Rev Panam Salud Publica. 2015;37(2):83-9.
11. Rocha-Brischiliari SC, Bulgarelli AF, Street ME, et al. The rise in mortality from breast cancer in young women: trend analysis in Brazil. LoS One. 2017;12(1). doi: https://doi.org/10.1371/journal.pone.0168950
12. TABNET [Internet]. Brasília (DF): DATASUS. c2008 – [acesso 2024 maio 28]. Disponível em: https://datasus.saude.gov.br/informacoes-de-saude-tabnet/
13. Hill K. Estimating census and death registration completeness. Asian Pac Popul Forum. 1987;1(3):8-13.
14. Bennett NG, Horiuchi S. Estimating the Completeness of Death Registration in a Closed Population. Popul Index. 1981;47(2):207-21. doi: http://dx.doi.org/10.2307/2736447
15. Hill K, You D, Choi Y. Death distribution methods for estimating adult mortality: Sensitivity analysis with simulated data errors. Demogr Res. 2009;21(9):235-54. doi: https://doi.org/10.4054/DemRes.2009.21.9
16. Hill K, Choi Y, Timæus IM. Unconventional approaches to mortality estimation. Demogr Res. 2005;13(12):281-99. doi: https://doi.org/10.4054/DemRes.2005.13.12
17. Doll R, Payne PM, Waterhouse JAH. Introduction. In: Agência Internacional de Pesquisa em Câncer. Cancer Incidence in Five Continents. Lyon: IARC;1966.
18. SPSS®: Statistical Package for Social Science (SPSS) [Internet]. Versão 28.0. [Chicago]. International Business Machines Corporation. [acesso 2023 mar 9]. Disponível em: https://www.ibm.com/br-pt/spss?utm_content=SRCWW&p1=Search&p4=43700077515785492&p5=p&gclid=CjwKCAjwgZCoBhBnEiwAz35Rwiltb7s14pOSLocnooMOQh9qAL59IHVc9WP4ixhNTVMjenRp3-aEgxoCubsQAvD_BwE&gclsrc=aw.ds

19. Conselho Nacional de Saúde (BR). Resolução nº 510, de 7 de abril de 2016. Dispõe sobre as normas aplicáveis a pesquisas em Ciências Humanas e Sociais cujos procedimentos metodológicos envolvam a utilização de dados diretamente obtidos com os participantes ou de informações identificáveis ou que possam acarretar riscos maiores do que os existentes na vida cotidiana, na forma definida nesta Resolução [Internet]. Diário Oficial da União, Brasília, DF. 2016 maio 24 [acesso 2024 ago 3]; Seção I:44. Disponível em: http://bvsms.saude.gov.br/bvs/saudelegis/cns/2016/res0510_07_04_2016.html
20. Santos-Silva I. Políticas de controle do câncer de mama no Brasil: quais são os próximos passos? Cad Saude Publica. 2018;34(6):e00097018. doi: https://doi.org/10.1590/0102-311X00097018
21. Medronho RA, Bloch KV, Luiz RR, et al. Epidemiologia. São Paulo: Atheneu; 2009.
22. Brass W. Methods for estimating fertility and mortality from limited and defective data. Chapel Hill: Occasional publication of the Centre for Population Studies; 1975.
23. Preston S, Coale AJ, Trussell J, et al. Estimating the completeness of reporting of adult deaths in populations that are approximately stable Popul Index. 1980;46(2):179-202. doi: https://doi.org/10.2307/2736122
24. Mathers C, Bernard C, Iburg K. Global burden of disease in 2002: data sources, methods and results. Geneva: WHO; 2003. (Global Programme on Evidence for Health Policy Discussion 54).
25. Cavalini LT, Ponce de Leon ACM. Correção de subregistros de óbitos e proporção de internações por causas mal definidas. Rev Saude Publica. 2007;41(1):85-93. doi: https://doi.org/10.1590/S0034-89102007000100012
26. Preston SH, Elo IT, Stewart Q. Effects of age misreporting on mortality estimates at older ages. Popul Stud (NY). 1999;53(2):165-77.
27. Gomes Andrade P, Gomes C, Veras RP, et al. Evolução da qualidade da declaração da idade na América Latina e Caribe: uma proposta de escolha de métodos a partir da estrutura etária [Internet]. 7 Congresso da Associação Latino-Americana de População (ALAP) e o 20 Encontro Nacional de Estudos Populacionais (ABEP); 2016 Out 17-22.; Foz do Iguaçu, Brasil. Caxambu: ALAP; 2016. [acesso 2024 ago 15]. Disponível em: https://files.alapop.org/congreso7/files/pdf/313-202.pdf
28. Silva Júnior, WP. Diferenciais regionais na mortalidade adulta por escolaridade no Brasil em 2010 [dissertação]. Natal: Universidade Federal do Rio Grande do Norte; 2018.
29. Paes NA, Gouveia JF. Recovery of the main causes of death in the Northeast of Brazil: impact on life expectancy. Rev Saude Publica. 2010;44(2):301-9. doi: https://doi.org/10.1590/s0034-89102010000200010

Recebido em 3/7/2024
Aprovado em 10/9/2024

Associate-editor: Jeane Glaucia Tomazelli. Orcid iD: https://orcid.org/0000-0002-2472-3444
Scientific-editor: Anke Bergmann. Orcid iD: https://orcid.org/0000-0002-1972-8777