DOI: 10.1590/1980-549720190002.supl.3

ORIGINAL ARTICLE/*ARTIGO ORIGINAL*

Quality of cause-of-death data in Brazil: Garbage codes among registered deaths in 2000 and 2015

Qualidade dos dados sobre causas de morte no Brasil: códigos garbage nos

Renato Azeredo Teixeira[I], Mohsen Naghavi[II], Mark Drew Crosland Guimarães[I], Lenice Harumi Ishitani[III], Elizabeth Barboza França[I]

ABSTRACT: ***Introduction:*** reliability of mortality data is essential for health assessment and planning. In Brazil, a high proportion of deaths is attributed to causes that should not be considered as underlying causes of deaths, named garbage codes (GC). To tackle this issue, in 2005, the Brazilian Ministry of Health (MoH) implements the investigation of GC-R codes (codes from chapter 18 "Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified, ICD-10") to improve the quality of cause-of-death data. This study analyzes the GC cause of death, considered as the indicator of data quality, in Brazil, regions, states and municipalities in 2000 and 2015. ***Methods:*** death records from the Brazilian Mortality Information System (SIM) were used. Analysis was performed for two GC groups: R codes and non-R codes, such as J18.0-J18.9 (Pneumonia unspecified). Crude and age-standardized rates, number of deaths and proportions were considered. ***Results:*** an overall improvement in the quality of mortality data in 2015 was detected, with variations among regions, age groups and size of municipalities. The improvement in the quality of mortality data in the Northeastern and Northern regions for GC-R codes is emphasized. Higher GC rates were observed among the older adults (60+ years old). The differences among the areas observed in 2015 were smaller. ***Conclusion:*** the efforts of the MoH in implementing the investigation of GC-R codes have contributed to the progress of data quality. Investment is still necessary to improve the quality of cause-of-death statistics.

Keywords: Vital statistics. Cause of death. Data accuracy. Information systems. Brazil.

[I]Public Health Graduate Program, School of Medicine, Universidade Federal de Minas Gerais – Belo Horizonte (MG), Brazil.
[II]Institute for Health Metrics and Evaluation – Seattle (WA), United States.
[III]Epidemiology and Health Assessment Research Group, Universidade Federal de Minas Gerais – Belo Horizonte (MG), Brazil.
Corresponding author: Elisabeth B. França. Faculty of Medicine. Universidade Federal de Minas Gerais. Avenida Alfredo Balena, 190, room 731, Santa Efigênia, CEP: 30130-100, Belo Horizonte, MG, Brazil. Email: efranca.med@gmail.com

Conflict of interests: nothing to declare – **Financial support:** Funding from Vital Strategies as part of the Bloomberg Philanthropies Data for Health Initiative (Project 23998 Fundep/UFMG).

RESUMO: ***Introdução*:** a confiabilidade dos dados de mortalidade é essencial para avaliação e planejamento da saúde. No Brasil, uma alta proporção de óbitos é atribuída a causas que não devem ser consideradas como causa básica (CBO) de óbitos, códigos *garbage* (CG). Para enfrentar essa questão, em 2005 o Ministério da Saúde (MS) implementou a investigação de CG-códigos R (capítulo 18 "Sintomas, sinais e achados anormais de exames clínicos e de laboratório, não classificados em outra parte, CID-10") para reduzir o impacto do erro de classificação das CBO. O estudo analisa os óbitos classificados como CG, considerados indicadores de qualidade dos dados, para o Brasil, regiões, estados e municípios, em 2000 e 2015. ***Métodos*:** utilizaram-se os registros de óbitos do Sistema de Informações sobre Mortalidade (SIM). A análise foi realizada para dois grupos de CG: códigos R e não códigos R, como J18.0-J18.9 Pneumonias não especificadas). Consideraram-se as taxas, brutas e ajustadas por idade, número absoluto e proporção dos óbitos. ***Resultados*:** foi observada melhoria global na qualidade dos dados de mortalidade em 2015, com variações entre regiões, idade e porte dos municípios. Destaca-se melhoria na qualidade dos dados de mortalidade das regiões Nordeste e Norte para o CG-códigos R. Maiores taxas de CG foram observadas entre idosos. Diferenças regionais observadas em 2015 foram menores. ***Conclusão*:** os esforços do MS na implementação da investigação de CG-códigos R contribuíram para progresso na qualidade dos dados. Ainda é necessário investir em melhorias adicionais na qualidade das estatísticas das causas de mortes.

***Palavras-chaves*:** Estatísticas vitais. Causas de morte. Confiabilidade dos dados. Sistemas de informação. Brasil.

INTRODUCTION

Analysis of cause of death is essential for assessing the population's health situation, thus defining priorities for public health. Despite it representing the most important aspect of vital statistics for supporting actions of public health authorities, useful cause of death (COD) information is not available in many countries[1]. Inequalities in the registration of vital events are substantial, and some of the poorest countries have no official birth or death certificates[2]. Most of developing countries do not have a complete vital registration system, in which deaths receive a death certificate filled in by a physician[1].

Many methods have been developed to assess the quality of a country's health information system. A traditional criterion to analyze the quality of vital statistics is the proportion of deaths from chapter 18 of the International Classification of Diseases – 10th revision (ICD-10), traditionally named Ill-defined causes of death[3-5]. The Global Burden of Disease Studies (GBD), which is responsible for generating health estimates for several countries using standardized analytical methodology for improving data quality, has the redistribution of underlying causes of death considered as garbage codes (GC) as a step in the treatment of the mortality data to obtain more accurate information. GCs are the ill-defined CODs or those that should not have been classified as an underlying cause. They include not only the codes of chapter 18 Symptoms, signs and abnormal clinical and laboratory findings not elsewhere classified, but also codes from other chapters not considered important for planning in public health such as J18.0-J18.9 (Pneumonia unspecified), A40.0-A41-9 (Sepsis) and C80.0-C80.9 (Malignant neoplasm without specification of site), since they do not provide satisfactory definition of actions in public health for the prevention and control of diseases and health problems[6]. A much more detailed and extended analysis of the concept was conducted by Mohsen Naghavi and colleagues in 2010, who classified the GCs in four categories with comparability across ICD revisions[7]. New revisions were published in the GBD 2010 study and following GBD studies[8,9]. Analysis of GCs says a lot about the quality of vital statistics in populations[7,10].

In Brazil, the quality of mortality data was very poor before the 1980's, with high underreporting of deaths. COD used to be collected only in some cities using different forms of death certificates (DC)[11,12]. The Brazilian Mortality Information System (*Sistema de Informações sobre Mortalidade*; SIM) was created in 1976 by the Ministry of Health (MoH), with the responsibility of collecting COD based on a standard World Health Organization (WHO) DC issued by physicians[11].

Since the beginning of the SIM, the MoH has promoted measures to strengthen the health system, including the quality of mortality data in Brazil[11]. After 2005, the Program "Reduction of the percentage of deaths due to ill – defined causes of death" focused especially on the country's Northern and Northeastern regions and initiatives have been developed by the MoH. Since then, several strategies have been implemented and expanded in Brazil such as the investigation of deaths due to ill-defined causes (R-codes) in hospitals and at home by municipal health workers, the implementation of the National Network of Death Verification Service, and suspension of federal funding to municipalities if the entrance of data into the SIM was not in accordance with the expected targets and within the defined deadlines, among other initiatives[13-17]. Considering the importance given to the quality of vital statistics, this study analyzes underlying causes of death assigned as GC, which are considered an indicator of data quality in Brazil, regions, states and municipalities, in 2000 and 2015.

METHODS

The source of data was the public database from SIM in 2000 and 2015, which included more than 2 million deaths with information classified according to the sex, age, city of residence and cause of death. We selected the years 2000 and 2015, before and after the implementation of the MoH initiative of investigation of ill-defined causes of death, which had population estimates corrected at the municipal level.

To analyze geographic inequalities in Brazil, the data was analyzed according to the following groupings:

a) Geographical regions (North, Northeast, Central West, Southeast and South);
b) 26 states and one Federal District;
c) 4 groups of municipalities according to size as follows: 1. 20,000 inhabitants or less (n = 3,879); 2. 20,001 to 100,000 inhabitants (n = 1,406); 3. 100,001 to 500,000 (n = 247); and, 4. more than 500,000 inhabitants (n = 38).

Besides the five traditional Brazilian regions and states, analysis was also presented according to municipalities, as they are also responsible for population health care[18]. Population size was justified due to the complexity of the health system and the criteria for federal and state funding – larger population size municipalities receive larger funding[19]. Although the number of small-sized municipalities is disproportionately greater than medium and large in all Brazilian states, human and financial resources for surveillance functioning are limited[20].

Deaths registered in the SIM in 2000 and 2015 were classified as GC according to the GBD 2015 study list[9] and divided into two groups for analysis in this study: a. GC ill defined (R codes), which are from the chapter 18 of the ICD-10 and include all R codes except R95; and b. GC other chapters (Non-R codes), from all other chapters of the ICD-10.

Number of deaths of GC for mortality rates were estimated after missing data and underreporting were corrected. Unknown sex, age and local of residence were proportionally reallocated using "pro rata" redistribution by year. Underreporting of deaths was corrected using factors generated by the ratio between the estimated number of all-cause mortality from the GBD 2015 study[9] and all number of registered deaths from SIM, according to age, sex and local of residence. The GBD study does not conduct analyses at the municipal level, so we could not correct underreporting at this level.

Population data for age, sex and municipalities were obtained from the Brazilian Institute of Geography and Statistics (*Instituto Brasileiro de Geografia e Estatística*; IBGE), Interagency Network for Health Information (*Rede Interagencial de Informações para a Saúde; Ripsa*)[21] and The Ministry of Health estimates[(a)]. Geographic areas were aggregated according to the size of municipalities observed in 2010, year of the last population census in Brazil.

Age-standardized mortality rates, per 100,000, were estimated by the direct method using the world's population from GBD 2015[9] as reference. For individual municipalities, we used crude mortality rates, i.e., number of deaths divided by each municipality population, per 100,000. The approximate 95% confidence interval (95% C.I.) for age-adjusted death rates were estimated for Brazil, regions and states[22].

To visualize the geographical distribution of changes in GC groups in states and municipalities between 2000 and 2015, the percent changes (PC) of mortality rates were estimated as follows:

$$PC = \left(\left({}^{\text{rate mortability in 2015}} / {}_{\text{rate mortability 2000}}\right) - 1\right) \text{x } 100$$

This study was approved by the Research Ethics Committee of the Universidade Federal de Minas Gerais CAAE 62803316.7.0000.5149 and was technically and financially supported by the Brazilian Ministry of Health – Ministry of Health Surveillance.

[(a)]http://tabnet.datasus.gov.br/cgi/deftohtm.exe?novapop/cnv/popbr.def

RESULTS

In 2000 and 2015, 946,960 and 1,264,176 Brazilians' deaths were registered in SIM, respectively. Among these deaths, 41.3% were reported as GC in 2000 (27.0% by Non-R codes and 14.3% by R codes) and 33.3% in 2015 (27.6% by Non-R codes and 5.7% by R codes). Differences in the proportion of GC-R codes among regions in 2000 were larger, ranging from 6.3%, in the South, to 28.2%, in the Northeast, while in 2015 the proportions were more homogeneous: 2.9% in the Central West and 7.9% in the North, as shown in Table 1. The top five ICD-10 codes of COD in the GC-non-R codes group for both years were: I64 – Stroke, not specified as hemorrhage or infarction (n = 84,353 deaths, 3.8% of total deaths); J189 – Pneumonia, unspecified organism (n = 65,516 deaths, 3.0% of total deaths); I10 – Essential (primary) hypertension (n = 33,138 deaths, 1.5% of total deaths); I500 – Congestive heart failure (n = 28,426 deaths, 1.3% of total deaths); and, A419 – Sepsis, unspecified organism (n = 28,004 deaths, 1.3% of total deaths).

Table 1. Proportions of deaths due to garbage codes, R-codes and Non-R codes, and total deaths by regions, sexes, age groups and size of municipalities. Brazil, 2000 and 2015.

Variables		2000			2015		
		Proportion		Total deaths	Proportion		Total deaths
		R[1]	Non-R[2]		R[1]	Non-R[2]	
Regions	Central West	8.5	25.6	54,292	2.9	24.2	83,381
	Northeast	28.2	24.3	230,848	7.2	28.6	338,940
	North	23.9	23.2	47,561	7.9	25.0	77,944
	Southeast	9.7	29.1	461,512	5.6	28.7	572,738
	South	6.3	26.3	152,477	3.4	25.2	191,173
Sex	Men	13.7	25.6	552,130	5.7	25.3	709,118
	Women	15.1	28.8	393,607	5.6	30.5	554,383
Age group	Under 5	12.5	18.2	79,470	2.3	14.3	43,096
	5 to 14	9.8	28.0	11,659	4.9	24.6	8,159
	15 to 49	10.1	23.9	218,812	5.1	18.0	230,876
	50 to 69	13.2	24.7	264,109	5.3	23.6	361,951
	70 plus	18.1	32.2	368,271	6.2	34.5	616,814
Size municipalities (1,000)	0--20	25.0	25.9	146,300	7.6	28.4	205,928
	20|--100	21.1	26.6	247,995	7.5	28.7	352,918
	100|--500	11.2	27.5	238,457	5.1	27.5	324,660
	500+	6.2	27.1	307,313	3.2	26.2	377,269
Brazil		14.3	27.0	946,690	5.7	27.6	1,264,176

Except for R code group in 2015, men showed smaller proportions of GC than women. Deaths among older people (70+ years old) showed the highest number of total deaths and the highest proportions of GC causes of death among women and men. We should note the lowest proportions in larger municipalities (500,000 and more inhabitants) for GC-R codes.

The Table 2 shows the overall decrease in age-standardized mortality rates due to garbage codes. Mortality due to GC-R codes decreased from 150.2 to 47.3 deaths per 100,000 inhabitants between 2000 and 2015, representing a 68.5% reduction, while the mortality due to GC-non-R codes decreased from 255.2 to 225.7 deaths per 100,000 inhabitants between 2000 and 2015, representing a 11.5% reduction. All the regions showed a decrease in the R code rates in 2015 when compared with 2000 – the Northeastern region had the highest reduction (75.4%) and the South the smallest (52.8%). On the other hand, GC-non-R codes showed higher proportions in 2015 than in 2000 in two regions, Northeast and North, while the other regions showed a less pronounced decrease in GC-non-R codes than in GC-R codes.

Women showed smaller age-standardized mortality rates due to garbage codes than men in both years. The age-standardized rate of GC-R codes in women was 122.8 deaths per 100,000 inhabitants in 2000 and 36.0 deaths per 100,000 inhabitants in 2015, representing a 71% reduction. Men showed 183.9 and 60.9 deaths per 100,000 inhabitants in 2000 and 2015, respectively. For GC-non-R codes, women had 214.4 deaths per 100,000 inhabitants in 2000, decreasing to 188.0 deaths per 100,000 inhabitants in 2015, representing a 12% reduction. Men had 301.6 in 2000 to 272.8 deaths per 100,000 inhabitants in 2015, representing a 10% reduction (rates not shown by sex in Table 2).

Analysis by age grouping showed that older people presented the highest rates of mortality due to garbage codes, and the 5 to 14-year-old-age group presented the lowest. The greatest reductions were observed in the age group of under 5 years old: −89% and −56% for GC-R codes and GC-non-R codes, respectively. Only GC-non-R codes in the group over 70 years old showed an increase of 7%, from 2,067.6 deaths per 100,000 inhabitants in 2000 to 2,219.3 in 2015).

Regarding the size of municipalities, the age-standardized mortality rates due to GC in 2000 presented higher values than in 2015, except for the GC-non-R codes group for 0-20,000 and 20,001-100,000 inhabitants, as shown in Table 2. In 2000, the rate of non-R codes in the 0-20,000 population group increased from 200.4, in 2000, to 212.2 deaths per 100,000 inhabitants, in 2015. For municipalities with 20,001-100,000 inhabitants, GC-non-R code rates increased from 237.8 to 239.4 deaths per 100,000, from 2000 to 2015 (overlapping confidence intervals). Big cities with 500,000 inhabitants or more showed the lowest values for GC-R code rates.

At the State level, high age-standardized mortality rates were observed in 2000 for both groups of GCs, as shown in Table 3. From the 16 states of the Northeast and North, 11 showed higher rates of GC-R codes than GC-non-R codes in 2000 (Acre, Alagoas, Amazonas, Bahia, Maranhão, Pará, Paraíba, Pernambuco, Piauí, Rio Grande do Norte, Sergipe). In 2015, all R code rates were lower than the non-R code rates. Moreover, R code rates decreased in all states, which was not observed in the non-R code group. The percent change of GC-non-R code increased in most of the North and Northeastern states.

Table 2. Age-standardized mortality rates due to garbage codes, R and Non-R, and Percent changes by regions, age groups and size of municipalities. Brazil, 2000 and 2015.

Variables	2000		2015		PC (%)	
	R	Non-R	R	Non-R	R	Non-R
Region	81.9	247.3	21.4	208.1	-73.9	-15.8
Central West	(79.4-84.3)	(243-251.6)	(20.5-22.3)	(205.2-211)		
Northeast	276.8	223	68.1	251.5	-75.4	12.7
	(274.9-278.7)	(221.4-224.7)	(67.3-68.9)	(249.9-253)		
North	240.5	218.3	76.6	237.7	-68.2	8.9
	(236.1-244.9)	(214.2-222.3)	(74.6-78.6)	(234.3-241.2)		
Southeast	94.2	281.5	42.9	220.0	-54.4	-21.9
	(93.3-95)	(280-283)	(42.5-43.4)	(218.9-221)		
South	62.2	257.7	29.4	203.9	-52.8	-20.9
	(60.9-63.5)	(255.1-260.3)	(28.7-30)	(202.2-205.7)		
Age	92.3	120.9	9.9	53.5	-89.3	-55.8
Under 5	(90.9-93.8)	(119.3-122.5)	(9.4-10.4)	(52.3-54.7)		
5 to 14	5.8	14.3	2.0	9.0	-65.9	-36.8
	(5.6-6.1)	(13.9-14.7)	(1.8-2.1)	(8.7-9.4)		
15 to 49	27.3	61.6	12.1	41.3	-55.7	-33
	(27-27.7)	(61.1-62.1)	(11.9-12.3)	(40.9-41.7)		
50 to 69	195.3	344.0	60.8	262.7	-68.9	-23.6
	(193.3-197.2)	(341.5-346.5)	(60-61.6)	(261-264.4)		
70+	1264.4	2067.6	427.1	2219.3	-66.2	7.3
	(1255.6-1273.2)	(2056.3-2078.8)	(423-431.1)	(2210-2228.5)		
Size Mun	213.0	200.4	58.4	212.1	-72.6	5.8
0-20	(211.0-215.0)	(198.5-202.4)	(57.5-59.3)	(210.4-213.7)		
20-100	208.7	237.8	65.9	239.4	-68.4	0.6
	(207-210.3)	(236.1-239.6)	(65.2-66.7)	(237.9-240.8)		
100|-500	128.9	289.7	45.2	243.6	-64.9	-15.9
	(127.3-130.4)	(287.5-292.0)	(44.5-45.9)	(242-245.2)		
500+	66.1	283.6	25.7	208.1	-61.1	-26.6
	(65.2-67.1)	(281.7-285.5)	(25.3-26.2)	(206.9-209.4)		

Size Mun: size of municipalities in thousands of inhabitants.

Table 3. Age-standardized mortality rates (per 100,000) due to garbage code, R-codes and Non-R codes, and Percent changes for states. Brazil, 2000 and 2015.

State		2000		2015		PC (%)	
		R	Non-R	R	Non-R	R	Non-R
North	RO	148.8 (136.8-160.8)	297.7 (281.7-313.7)	52.6 (47.5-57.6)	241.7 (230.7-252.7)	-64.7	-18.8
	AC	320.8 (298-343.7)	191.2 (173.8-208.7)	59.3 (51.4-67.1)	254.3 (237.7-271)	-81.5	33.0
	AM	258.9 (248.5-269.3)	175.6 (167.3-184)	120.9 (115.3-126.5)	207.8 (200.5-215.1)	-53.3	18.3
	RR	93.5 (73.3-113.8)	302.5 (265.6-339.3)	54.3 (41.3-67.2)	246.7 (220.8-272.6)	-42.0	-18.4
	PA	258.3 (252-264.6)	213.1 (207.4-218.7)	71.3 (68.7-74)	249.8 (244.9-254.8)	-72.4	17.3
	AP	158.6 (138.2-179)	265.3 (238.3-292.4)	96.9 (85.1-108.7)	226.3 (207.9-244.6)	-38.9	-14.7
	TO	215.6 (202.9-228.2)	233.8 (221.1-246.6)	26.2 (22.8-29.5)	222.2 (212.2-232.2)	-87.9	-5.0
Central West	MS	83.0 (77.6-88.4)	275.6 (265.9-285.4)	14.0 (12.3-15.6)	195.8 (189.4-202.2)	-83.2	-29.0
	MT	89.8 (83.7-96)	292.4 (281.5-303.4)	58.1 (54.6-61.6)	231.6 (224.5-238.7)	-35.3	-20.8
	GO	93.1 (89.1-97.0)	234.4 (228.1-240.7)	21.2 (19.8-22.5)	224.6 (220.1-229.1)	-77.2	-4.2
	DF	44.9 (40.0-49.7)	220.6 (209.4-231.8)	4.2 (3.3-5.1)	194.1 (187.4-200.7)	-90.5	-12.0
South	PR	55.3 (53.3-57.4)	269.4 (264.9-274)	24.7 (23.7-25.7)	222.6 (219.6-225.7)	-55.4	-17.4
	SC	126.0 (121.7-130.4)	242.4 (236.5-248.2)	18.2 (17.1-19.3)	192.5 (188.7-196.2)	-85.5	-20.6
	RS	41.7 (40.2-43.3)	254.0 (250.2-257.7)	32.1 (31.1-33.1)	188 (185.6-190.5)	-23.1	-26.0

Continue...

Table 3. Continuation.

State		2000		2015		PC (%)	
		R	Non-R	R	Non-R	R	Non-R
Northeast	MA	372.8 (365.9-379.8)	226.6 (221.4-231.8)	49.3 (47.2-51.5)	286.8 (281.6-292.0)	-86.8	26.6
	PI	272.4 (264.4-280.4)	215.2 (208.3-222)	33.6 (31.2-36.0)	280.6 (273.5-287.8)	-87.7	30.4
	CE	184.8 (181.0-188.7)	253.2 (248.7-257.7)	50.2 (48.5-51.9)	284.0 (280.0-288.1)	-72.8	12.2
	RN	225.7 (218.8-232.5)	171.2 (165.4-177.0)	26.6 (24.7-28.5)	200.8 (195.5-206.2)	-88.2	17.3
	PB	397.6 (389.9-405.3)	175.8 (170.7-180.9)	51.9 (49.4-54.4)	258.2 (252.6-263.7)	-86.9	46.8
	PE	241.8 (237.5-246.2)	210.8 (206.8-214.8)	38.5 (37.0-40.0)	259.2 (255.4-263.0)	-84.1	23
	AL	302.7 (293.9-311.5)	241.1 (233.4-248.9)	45.6 (42.7-48.4)	277.8 (270.7-285.0)	-85	15.2
	SE	303.1 (292.1-314.2)	196 (187.4-204.6)	65.9 (61.6-70.2)	252.6 (244.1-261.1)	-78.3	28.9
	BA	283.5 (279.8-287.2)	239.6 (236.2-242.9)	123.4 (121.4-125.5)	246.3 (243.4-249.1)	-56.5	2.8
Southeast	MG	121.0 (119.1-123.0)	264.5 (261.5-267.4)	60.3 (59.2-61.4)	232.4 (230.2-234.5)	-50.2	-12.1
	ES	174.5 (168.2-180.9)	242.3 (234.9-249.6)	7.4 (6.5-8.3)	174.0 (169.5-178.6)	-95.8	-28.2
	RJ	114.1 (112-116.1)	295.5 (292.2-298.8)	46.6 (45.5-47.6)	264.3 (261.9-266.8)	-59.2	-10.6
	SP	65.4 (64.3-66.4)	287.3 (285.0-289.5)	34.9 (34.3-35.5)	204.6 (203.1-206.0)	-46.6	-28.8
BR		150.2 (149.5-151)	255.2 (254.2-256.2)	47.3 (46.9-47.6)	225.7 (224.9-226.4)	-68.5	-11.6

At the municipality level, good results in percent changes of mortality rates of R codes can be seen in 2015 when compared with 2000, as shown in Figure 1A. Almost all municipalities presented green color, representing a decrease in the rates in 2015. On the other hand, Figure 1B showed that the non-R-code group of GC rates did not show as good results as the R-code – red colors stand out more than the green, indicating an increase in the rates for most municipalities in the country.

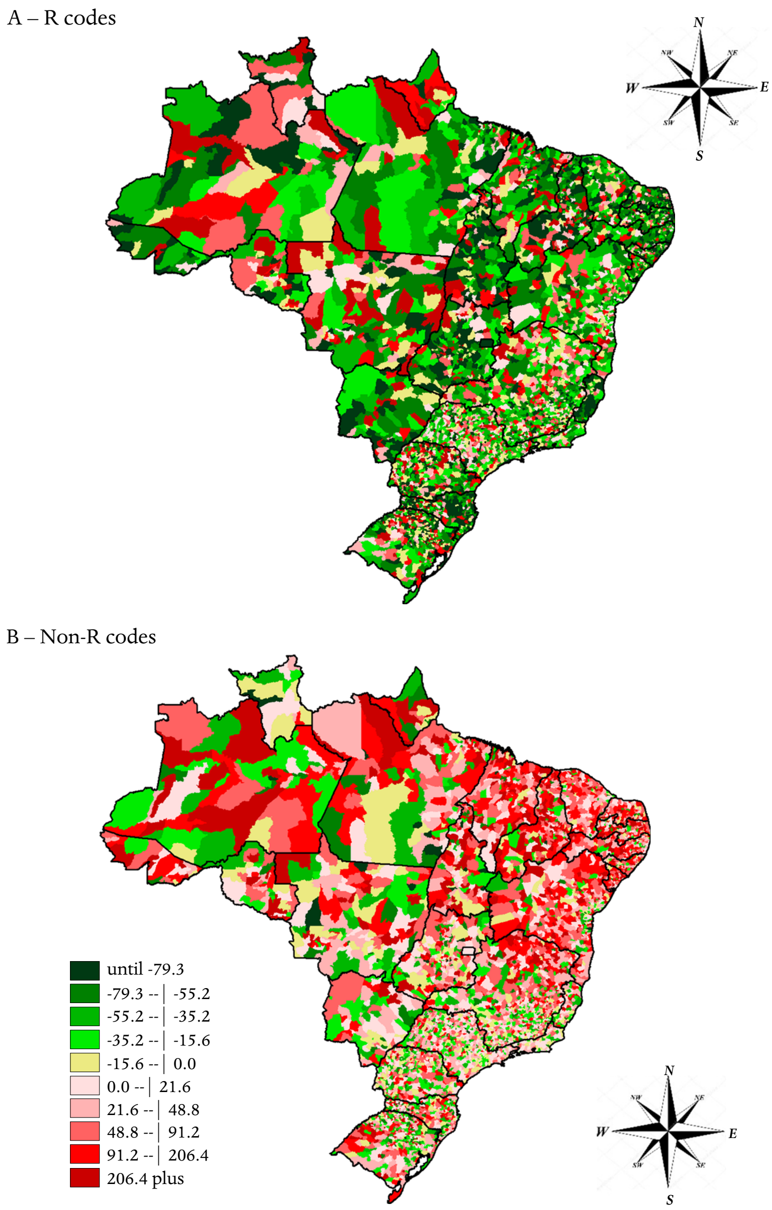

Figure 1. Percent changes of mortality rates due to garbage codes between 2000 and 2015 in municipalities of Brazil.

DISCUSSION

To the best of our knowledge, this is the first study that analyzed the quality of mortality data due to underlying causes of death assigned as GC in Brazil. Generally, substantial progress has been seen in the quality of mortality data, which indicates improvements due to MoH initiatives. However, our results have also revealed important heterogeneity throughout the country while the analysis of GC groups considered the regional level in 2000 and 2015. The decrease in fractions and in age-standardized rates of mortality due to GC-R codes in 2015, corroborates the results concerning the decrease of ill-defined causes of death in Brazil[23], especially in the Northeastern and Northern regions.

Investigation methods of GC-R codes in Brazil have been performed since 2005 by municipal surveillance teams using information on the deceased's terminal disease obtained from hospitals, from the Family Health Strategy – FHS records, and other health services[13,14]. Moreover, data from routine national information systems such as the Disease Notification Information System and the Hospital Information System, autopsy data from the Death Verification Service for natural causes, and from forensic institutes were also used. Household interviews were conducted with the deceased's family by a qualified health professional when these sources were not sufficient to identify the cause of death. Since 2009, a standardized verbal autopsy (VA) form from WHO was implemented in the whole country, and the cause of death of the VAs were certified by a physician[24]. The Northern and Northeastern regions were prioritized in this project, and our findings indicate positive results obtained by the MoH[18].

Percent changes of GC-R code rates between 2000 and 2015 have shown a reduction in all states, also indicating successful results in the investigation of ill-defined COD. Despite it showing less variability in 2015, GC-R code rates were lower in smaller municipalities in both years. However, caution in inferences is necessary, as corrections of the underreporting of deaths in the SIM data for this geographic level were not applied. For state and national level, we applied the correction of underreporting considering all causes of death, not only for GC causes of death, assuming same proportions of causes among reported and non-reported deaths, although some causes such as injuries could have different distribution[25].

The quality of vital statistics presents geographic and socioeconomic differences. These statistics are of a higher standard in more affluent regions[26]. In addition, more is invested in health to larger and more developed areas, which provides a better quality in the definition of COD[27,28]. The decentralized primary health care program from the MoH, the FHS, may be responsible for reducing the number of unattended deaths and registered ill-defined COD in small municipalities, recently. FHS coverage has increased in the country and has included more rural areas, with a higher coverage than in urban areas[29].

In addition to the possibility of verifying the situation in quality of mortality data at a more specific level, the analysis by municipalities is interesting since we could better visualize the real situation of GCs throughout the country, even in smaller areas. The highest percent change of rates between 2000 and 2015 occurred in the smallest municipalities, which reinforces the importance of the decentralized interventions undertaken by the MoH[3].

Concerning age groups, the proportion of deaths assigned as GC R-codes and GC-non-R codes vary in groups and years, always demonstrating higher values within the older adult groups. Likewise, mortality rates have shown higher values, except for under 5-year-old group. Naghavi *et al.* and other researchers also found higher values of GCs for infants of 1 to 4 years old and for older adults over 60 years old[7,30,31]. Older people have multiple morbidities, therefore the underlying diagnosis is more difficult, which may explain the increase of GC in these age groups[32].

Except for GC-R codes in 2015, higher values of GC groups are observed for women than men when we analyze the proportions according to sex. On the other hand, higher age-standardized mortality rates have been observed more in men than in women. Causes of death should be better defined for men due to the higher occurrence of external causes. External causes are assigned by coroners after performing autopsies, or after investigating the circumstances of injuries in police reports[33].

In this study, other analysis concerning the GC-non-R codes are also essential to the quality of mortality data. Those codes have not decreased as GC-R codes. The Northern and Northeastern regions have not presented a negative percent change of GC-non-R codes, which were not routinely investigated by health services[14]. One hypothesis that should be studied is that occasionally the investigated GC-R code cases had their underlying cause of death changed to codes belonging to the GC-non-R code group.

Studies using methods for correcting underreported deaths and redistributing GCs have recently increased, since robust methodologies were developed by groups of researchers[5,7-9]. These studies are essential for minimizing errors in mortality data analyses, but statistical methods and even the GC investigation cannot properly solve the problem concerning correct assignment of the individual underlying cause of death. The best way to guarantee more accuracy with estimates is a correct definition of COD by a physician. Thus, a physician's training in medical certification of cause of death should be implemented throughout the country to enable a sustainable replacement of the GC investigation.

In summary, our findings demonstrate the magnitude of GCs of mortality data in Brazil and indicate successful results of strategies to qualify the information on COD. Despite the effectiveness in the efforts of MoH to improve the quality of vital statistics, focusing on health professionals awareness is still necessary, especially physicians. More in-depth knowledge on the situation of GCs, and the correct filling out of death certificates at local levels, in hospitals and other institutions should be prioritized.

CONCLUSION

The findings of this studies showed the magnitude of GC in Brazil and indicate that investigation of deaths assigned as GC-R codes have been effective and would impact the GC amount. Further studies should be encouraged to improve the knowledge of COD quality in Brazil, especially for the GC-non-R-code group, focusing on the Northern and Northeastern regions. Brazil should continue working on public health initiatives to obtain more accurate data to reach the same degree of accuracy with data quality as developed countries.

REFERENCES

1. Mathers CD, Fat DM, Inoue M, Rao C, Lopez AD. Counting the dead and what they died from: an assessment of the global status of cause of death data. Bull World Health Organ. 2005;83(3):171-77.
2. Setel PW, Macfarlane SB, Seater S, Mikkelsen L, Jha P, Stout S, et al. Monitoring of Vital Events A scandal of invisibility: making everyone count by counting everyone. Lancet. 2007;370(9598):1569-77. https://doi.org/10.1016/S0140-6736(07)61307-5
3. Rasella D, Aquino R, Barreto ML. Impact of the Family Health Program on the quality of vital information and reduction of child unattended deaths in Brazil: an ecological longitudinal study. BMC Public Health. 2010;10:380. https://doi.org/10.1186/1471-2458-10-380
4. Mahapatra P, Shibuya K, Lopez AD, Coullare F, Notzon FC, Rao C, et al. Civil registration systems and vital statistics: successes and missed opportunities. Lancet. 2007;370(9599):1653-63. https://doi.org/10.1016/S0140-6736(07)61308-7
5. França E, Teixeira R, Ishitani L, Duncan BB, Cortez Escalante JJ, Morais Neto OL, et al. Ill-defined causes of death in Brazil: a redistribution method based on the investigation of such causes. Rev Saúde Pública. 2014;48(4):671-81. http://dx.doi.org/10.1590/S0034-8910.2014048005146
6. Murray CJL, Lopez AD. Estimating causes of death: new methods and global and regional applications for 1990. Boston: Harvard School of Public Health; 1996. p. 118-200.
7. Naghavi M, Makela S, Foreman K, O'Brien J, Pourmalek F, Lozano R. Algorithms for enhancing public health utility of national causes-of-death data. Population Health Metrics. 2010;8(1):9. https://doi.org/10.1186/1478-7954-8-9
8. Naghavi M, Wang H, Lozano R, et al. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet. 2015;385(9963):117-171. https://doi.org/10.1016/S0140-6736(14)61682-2
9. Murray JL, Lopez AD, Naghavi M, Wang H, et al. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet. 2016;388(10053):1459-544. https://doi.org/10.1016/S0140-6736(16)31012-1
10. Phillips DE, Lozano R, Naghavi M, et al. A composite metric for assessing data on mortality and causes of death: the vital statistics performance index. Popul Health Metr 2014;12:14. https://doi.org/10.1186/1478-7954-12-14
11. Jorge MH, Laurenti R, Gotlieb SL. Análise da qualidade das estatísticas vitais brasileiras: a experiencia de implantacao do SIM e do SINASC. Ciên Saúde Coletiva. 2007;12(3):643-54. http://dx.doi.org/10.1590/S1413-81232007000300014
12. Ministério da Saúde. Secretaria de Gestão Participativa. Departamento de Informática do SUS. Rede Interagencial de Informações para a Saúde [Internet]. [cited 2019 Apr 20]. Available from: http://tabnet.datasus.gov.br/cgi/idb2012/folder.htm
13. França EB, Cunha CC, Vasconcelos AMN, Escalante JJC, Abreu DX, Lima RB, et al. Investigation of ill-defined causes of death: assessment of a program's performance in a State from the Northeastern region of Brazil. Rev Bras Epidemiol. 2014;17(1):119-134. https://doi.org/10.1590/1415-790X201400010010ENG
14. Ministério da Saúde. Manual para investigação do óbito com causa mal definida [Internet]. Brasília, DF: 2009 [cited 2019 Apr 20]. Available from: http://bit.ly/2HrdeCw
15. Souza MFM, Barea V, Williams D. Improving the mortality information in poor areas: the Brazilian experience. Proceedings of the WHO Family of International Classifications – WHO-FIC; 2007; Trieste, Italy.
16. Ministério da Saúde (BR). Portaria nº 3.252, de 11 de fevereiro de 2009. Aprova as diretrizes para execução e financiamento das ações de Vigilância em Saúde pela União, Estados, Distrito Federal e Municípios e dá outras providências. Diário Oficial da União. Brasília, DF; 2009 Dec.;65.
17. SVO: Portaria nº 1405, de 29 de junho de 2006. Institui a Rede Nacional de Serviços de Verificação de Óbito e Esclarecimento da Causa Mortis (SVO). Diário Oficial da União, Brasília, DF; 2006 June.
18. Brasil. Constituição da República Federativa do Brasil. Brasília, DF: Senado; 1988.

19. Cohn A, Westphal MF, Elias PE. Data and the process of formulating health policies. Rev Saúde Pública 2005;39(1):1-7.
20. Recktenwaldt M, Junges JR. Organization and practice of the health surveillance in small municipalities. Saúde e Sociedade. 2017;26(2):367-381. http://dx.doi.org/10.1590/s0104-12902017171916
21. Rosseti ES. Documento técnico contendo descrição da metodologia aplicada na determinação das projeções e das estimativas das populações dos municípios, por sexo e idade – em papel e meio magnético. Rio de Janeiro: Ministério da Saúde Secretaria de Gestão Estratégica e Participativa Departamento de Informática do SUS Coordenação Geral de Disseminação de Informações em Saúde; 2014.
22. Minino AM. A guide to State implementation of ICD-10 for mortality, part II: applying comparability ratios. Hyattsville, Maryland: National Center for Health Statistics; 2000.
23. Martins Junior DF, Costa TM, Lordelo MS, Felzemburg RDM. Tendência dos óbitos por causas mal definidas na região Nordeste do Brasil, 1979-2009. Rev Assoc Med Bras. 2011;57(3):338-46. http://dx.doi.org/10.1590/S0104-42302011000300019
24. Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Análise de Situação de Saúde. Manual de instruções para o preenchimento da Declaração de Óbito. Brasília, DF; 2011.
25. Campos D, França E, Loschi RH, Souza MFM. Uso da autópsia verbal na investigação de óbitos com causa mal definida em Minas Gerais, Brasil. Cad Saúde Pública. 2010 June;26(6):1221-33. http://dx.doi.org/10.1590/S0102-311X2010000600015
26. Kulhánová I, Menvielle G, Bopp M, Borrell C, Deboosere P, Eikemo TA, et al. Socioeconomic differences in the use of ill-defined causes of death in 16 European countries. BMC Public Health. 2014;14:1295. https://doi.org/10.1186/1471-2458-14-1295
27. Abreu DMX, Sakurai E, Campos LN. A evolução da mortalidade por causas mal definidas na população idosa em quatro capitais brasileiras, 1996-2007. R Bras Est Pop. 2010;27(1):75-88. http://dx.doi.org/10.1590/S0102-30982010000100006
28. Kanso S, Romero DE, Leite IC, Moraes EN. Diferenciais geográficos, socioeconômicos e demográficos da qualidade da informação da causa básica de morte dos idosos no Brasil. Cad. Saúde Pública [Internet]. 2011 July [cited 2017 Apr 28];27(7):1323-1339. http://dx.doi.org/10.1590/S0102-311X2011000700008
29. Malta DC, Sanros MAS, Stopa SR, Vieira JEB, Melo EA, Reis AAC. A cobertura da Estratégia de Saúde da Família (ESF) no Brasil, segundo a Pesquisa Nacional de Saúde, 2013. Cien Saude Colet. 2016; 21(2):327-338. http://dx.doi.org/10.1590/1413-81232015212.23602015
30. Ngo AD, Rao C, Hoa NP, Adair T, Chuc NTK. Mortality patterns in Vietnam, 2006: findings from a national verbal autopsy survey. BMC Research Notes. 2010;3(1):78. https://doi.org/10.1186/1756-0500-3-78
31. Costa MR, Marcopito LF. Mortalidade por causas mal definidas, Brasil, 1979-2002, e um modelo preditivo para idade. Cad Saúde Pública. 2008;24(5):1001-12.
32. Mello Jorge MHP, Laurenti R, Lima-Costa MF, Gotlieb SLD, Chiavegatto Filho ADP. A mortalidade de idosos no Brasil: a questão das causas mal definidas. Epidemiol Serv Saúde. 2008;17(4):271-81. http://dx.doi.org/10.5123/S1679-49742008000400004
33. Matos SG, Proietti FA, Barata RCB. Confiabilidade da informação sobre mortalidade por violência em Belo Horizonte, MG. Rev Saúde Pública. 2007;41(1):76-84. http://dx.doi.org/10.1590/S0034-89102007000100011

Received: 09/24/2018
Final version presented on: 04/16/2019
Accepted: 05/07/2019

Acknowledgments: We would like to thank the *Institute for Health Metrics and Evaluation* (IHME), University of Washington, for the development and methods of database of the study GBD 2015, and also the Brazilian Ministry of Health for providing the access to data sources and technical support.

Author's contribution: Teixeira RA, França EB, Ishitani LH and Naghavi M idealized and projected the study. Teixeira R and França EB developed the management, data exploration and analysis, elaboration and interpretation of results, and discussion. França EB, Ishitani LH, Naghavi M and Guimarães MDC critically reviewed the manuscript and collaborated in the whole process. All the authors have read, contributed to and approved the final manuscript.