






Improving the usefulness of mortality data: reclassification of ill-defined causes based on medical records and home interviews in Brazil

Aprimorando dados de mortalidade: reclassificação de causas mal definidas baseada em registros clínicos/entrevistas domiciliares no Brasil

Elisabeth Barboza França^I , Lenice Harumi Ishitani^{II} , Renato Azeredo Teixeira^{II} , Carolina Cândida da Cunha^{II} , Maria Fatima Marinho^{III} 

ABSTRACT: *Introduction:* Brazil presented a high proportion of ill-defined causes of death (IDCD) in 2000, compromising accurate cause-of-death analysis. *Objective:* To analyze specific underlying causes for deaths originally assigned as IDCD in the Mortality Information System (SIM – *Sistema de Informação sobre Mortalidade*), after investigation activities implemented in country between 2006 and 2017. *Method:* For all IDCD identified in the SIM, municipal health professionals collected information about the final disease obtained from hospital records, autopsies, forms of family health teams, and home investigation. Specific causes among reclassified IDCD after investigation were evaluated according to age groups and four calendar periods. *Results:* Proportions of IDCD reassigned to other causes after review increased over time, reaching 30.1% in 2017. From a total of 257,367 IDCD reclassified in 2006-2017, neonatal-related conditions, injury, ischemic heart disease and stroke were the leading causes detected in the age groups 0-9 years, 10-29 years, 30-69 years, 70 years and over, respectively. *Discussion:* The similarity and plausibility of cause-specific proportions derived from the reclassification of IDCD by age group over time indicate the accuracy of the investigation data. *Conclusion:* High proportions of IDCD reassigned to more informative causes after review indicate the success of this approach to correct misclassification in the SIM, an initiative that should be maintained. Training physicians on death certification along with better quality of medical care and access to health services would lead to further improvement.

Keywords: Information systems. Vital statistics. Mortality. Cause of death. Under registration.

^IPublic Health Graduate Program, Universidade Federal de Minas Gerais, Belo Horizonte (MG) – Brazil.

^{II}Epidemiology and Health Assessment Research Group, Universidade Federal de Minas Gerais, Belo Horizonte (MG) – Brazil.

^{III}Institute of Advanced Studies, Universidade de São Paulo, São Paulo (SP) – Brazil.

Corresponding author: Elisabeth Barboza França. Av. Alfredo Balena, 190, room 731, Santa Efigênia, CEP: 30130-100, Belo Horizonte, MG, Brazil. E-mail: 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:* As proporções de causas de morte mal definidas (CMD) foram elevadas no Brasil em 2000. *Objetivo:* Analisar causas específicas para óbitos originalmente como CMD no Sistema de Informação sobre Mortalidade (SIM), após investigação implementada no país entre 2006 e 2017. *Metodologia:* Para as CMD identificadas no SIM, profissionais das secretarias de saúde coletaram informações sobre a doença final obtida nos registros hospitalares, autópsias, formulários de equipes de saúde da família e investigação domiciliar. Causas específicas entre as CMD reclassificadas após investigação foram avaliadas segundo idade em quatro períodos. *Resultados:* Proporções de CMD reclassificadas para outras causas após revisão aumentaram ao longo do tempo, atingindo 30,1% em 2017. De um total de 257.367 CMD reclassificadas entre 2006 e 2017, causas perinatais, causas externas, cardiopatia isquêmica e doença cerebrovascular foram os principais grupamentos de causas nos grupos etários de 0-9 anos, 10-29 anos, 30-69 anos, 70 anos ou mais, respectivamente. *Discussão:* A similaridade e a plausibilidade das proporções de causas específicas detectadas entre as CMD nos grupos etários ao longo do tempo indicam acurácia dos dados da investigação. *Conclusão:* Causas mais informativas detectadas após a revisão das CMD indicam o sucesso dessa abordagem para corrigir erros de classificação no SIM, que deve ser mantida. Maior capacitação dos médicos no preenchimento do atestado de óbito e melhor acesso e qualificação dos serviços de saúde são importantes para melhoria futura.

Palavras-chave: Sistemas de informação. Estatísticas vitais. Mortalidade. Causa da morte. Sub-registro.

INTRODUCTION

Analyses of reliable levels and cause-of-death (COD) trends are important to guide priority health interventions¹. In Brazil, the Mortality Information System (SIM – *Sistema de Informação sobre Mortalidade*) was created by the Ministry of Health (MoH) in 1975 through the collection of COD data using the World Health Organization (WHO) standard form for death certificate (DC), which is filled in by physicians in accordance with the WHO guidelines. Although the high completeness, the number of unusable causes of death – in particular garbage codes (GC) from chapter 18-ICD-10, or R codes, herein named ill-defined causes of death (IDCD) – has been a significant problem among registered causes at SIM in Brazil. In 2003, the proportion of IDCD was 13.3%, representing 133,434 deaths, with notable differences throughout the country, varying from 6.7% in the South region to 21.2% and 25.9% in the North and the Northeast regions, respectively².

To tackle this issue, the Government introduced a new strategy in 2005 with the aim to identify the underlying causes of IDCD using intentional search in health institutions like hospitals. If necessary, home interviews using verbal autopsy were also conducted to reclassify the cause of death³. After all, the causes of death assigned after investigation activities were registered in the SIM at the end of each year, and became part of final COD statistics in Brazil.

Investigation of IDCD as part of routine data collection on a large scale as had occurred in Brazil in 2006-2017 is an innovative approach to strengthen population-level mortality statistics and is also a significant new endeavor. Causes of death derived from this initiative should be evaluated to assess its plausibility and to enable further research on the use of

empirical algorithms for redistributing deaths due to the remaining IDCDC. Therefore, this study aims to describe the target underlying causes of death identified among IDCDC investigated in health services and by verbal autopsy. Proportions of reclassified deaths by selected variables and cause-specific mortality fractions (CSMF) derived from the reclassified IDCDC by age were analyzed over time to assess the plausibility of our results.

METHODS

This study used data from SIM, available at the website www.datasus.gov.br of the Department of the Brazilian National Health System (Datasus). All deaths registered in the SIM from 2006 to 2017 for both sexes were considered. All four digit ICD-10 R codes from chapter 18 were selected, except for R50.2 (Drug-induced fever), R78.0 (Finding of alcohol in blood), R78.1 (Finding of opiate drug in blood), R78.2 (Finding of cocaine in blood), R78.3 (Finding of hallucinogen in blood), R78.4 (Finding of other drugs of addictive potential in blood), R78.5 (Finding of psychotropic drug in blood), R95 (Sudden infant death syndrome); these items were not classified as ill-defined causes according to the Global Burden of Disease Study (GBD) 2017⁴. Garbage codes in the GBD study were classified into four levels, based on their redistribution across different levels of the GBD cause hierarchy. A garbage code (GC) Level 1 results in redistribution of the cause into a different category at the top level of the GBD cause hierarchy: communicable, maternal or perinatal diseases; non-communicable diseases; and injuries. GC Level 2 includes all GCs that are assigned to specific GBD Level 1 causes but to different Level 2 causes. Level 3 are GCs from Level 2 which are redistributed to a different Level 3 cause. Level 4 includes GCs such as “unspecified stroke” that can be redistributed within the Level 3 specific cause “stroke”. Although the GBD 2017 list divides R codes into three levels of GCs⁴, Level 1 R codes in Brazil represent 92% of the total; therefore, this study presents an analysis based on the total of R codes.

In the SIM database, the DC that originally showed IDCDC is called “original underlying cause (*causa básica original*)”, referring to the underlying cause registered at the time of death in the original DC. This study categorized the DCs that originally showed IDCDC into two groups: (1) investigated IDCDC (variable reported in a specific field in the SIM); (2) reclassified IDCDC. In the latter, the IDCDC is replaced by an underlying cause belonging to any chapter of ICD-10, except for chapter 18. When the cause has no investigation or does not enable the clarification of the underlying cause, the original underlying cause will be the same as the final cause. In the case of substitution of IDCDC for an underlying cause not included in chapter 18, the death will be presented as reclassified.

To investigate IDCDC, municipal health professionals should use the data sent by the MoH, such as records from the SIM linked to the Hospital Information System. However, this procedure was held at Ministry level for only a few years. In general, to clarify the COD, data was first collected from hospitals and other institutions such as civil registries, autopsy services, and municipal health departments including the family health strategy, using a

standardized MoH form. For deaths without a defined cause after those procedures, home interviews using the verbal autopsy (VA) method should be conducted. The WHO 2007 VA questionnaire was introduced in 2007 for home interviews, as a pilot project launched in states from the Northeast, North and Southeast regions. From 2009 on, a VA-adapted questionnaire was implemented throughout the country. A physician should fill out the part of cause of death conclusion into the VA form, herein presented in the same way as the DC form. After that, a coder using an automated program selects the underlying cause of death, which is included in the SIM as a reclassified cause^{3,5,6}.

This study describes the distribution of IDCD investigated and reclassified over time and also proportions of cases according to age, region of residence and place of death. The contribution of each data source to the investigation process was also analyzed: data from other information systems, medical records from hospitals, autopsy reports from the Death Investigation Service (SVO – *Serviço de Verificação de Óbito*) and forensic institutes, verbal autopsy records from home interviews, and other sources such as municipal health departments and physician's interviews. When a death was investigated from more than one of the aforementioned sources, information regarding data sources should be presented as "several sources".

To analyze CODs among reclassified IDCD, the GBD 2017 cause list was used⁴. This list places all causes in a hierarchical level. The first level divides the causes into three broad groups: 1. communicable, maternal, neonatal conditions, and nutritional deficiencies; 2. non-communicable diseases; and, 3. injuries. The second level consists of 20 major CODs, and the third level subdivides Level 2 into types such as cerebrovascular diseases and road traffic injuries. In this paper, the results are reported mainly at the GBD third level. Causes that account for 0.5% of deaths or more were included; all other causes were classified into broad-cause residual categories. For some causes, the following garbage codes were aggregated to specific causes, differently from the GBD 2017 cause list which does not include any garbage codes among the listed causes: (1) Lower respiratory infections: J15.9, J18, J22; (2) Oral, nasopharynx cancer: C14; (3) Diabetes: E12-E14 (including all 4 digit codes); (4) Stroke: I64, I67.4, I67.8, I67.9, I69.4, I69.8.

The CSMF is calculated as the number of deaths due to a specific cause divided by all deaths. CSMFs of reclassified IDCDs were evaluated according to age groups (0-9 years old, 10-29 years, 30-69 years and 70 and over), in four calendar periods: 2006-2008, 2009-2011, 2012-2014, 2015-2017.

This study was approved by the Research Ethics Committee of the Universidade Federal de Minas Gerais (CAEE 7555317.0.0000.5149) and developed according to the ethical rules established in the ordinance no. 466/2012 of the Brazilian National Health Council.

RESULTS

In Brazil, between 2006 and 2017, more than one million deaths occurred annually and around 100,000 deaths per year were originally attributed to IDCD, although the proportions have decreased over time. We observed an increase in proportions of investigated and reclassified

COD over time, more pronounced in 2017. In this year, 40,728 deaths were investigated (40.2%) and 30,595 deaths had the underlying cause reclassified after investigation (Table 1).

Table 1. Proportions of deaths due to ill-defined causes (IDCD) according to investigation and reclassification. Brazil, 2006-2017.

Year death	Deaths all causes	IDCD before investigation*		Investigated IDCD**		Reclassified IDCD**		IDCD after investigation*	
	n	n	%	n	%	n	%	n	%
2006	1,031,691	95,133	9.2	24,897	26.2	15,204	16.0	79,929	7.7
2007	1,047,824	95,917	9.2	24,597	25.6	15,988	16.7	79,929	7.6
2008	1,077,007	95,027	8.8	26,090	27.5	17,079	18.0	77,948	7.2
2009	1,103,088	94,699	8.6	26,759	28.3	17,141	18.1	77,558	7.0
2010	1,136,947	97,134	8.5	29,402	30.3	19,286	19.9	77,848	6.8
2011	1,170,498	97,205	8.3	30,029	30.9	20,642	21.2	76,563	6.5
2012	1,181,166	94,721	8.0	30,686	32.4	21,213	22.4	73,508	6.2
2013	1,210,474	95,884	7.9	34,579	36.1	25,570	26.7	70,314	5.8
2014	1,227,039	93,539	7.6	31,686	33.9	23,460	25.1	70,079	5.7
2015	1,260,565	98,655	7.8	31,762	32.2	23,428	23.7	75,227	6.0
2016	1,309,774	102,554	7.8	37,180	36.3	27,761	27.1	74,793	5.7
2017	1,312,664	101,358	7.7	40,728	40.2	30,595	30.2	70,763	5.4

*Deaths IDCD/all causes. All R codes were included except for R50.2, R78.0, R78.1, R78.2, R78.3, R78.4, R78.5, R95 codes.

** Deaths/IDCD before investigation.

Table 2 shows the proportions of IDCD reclassified according to age, regions, place of death and sources of investigation. Proportions were relatively similar in both sexes and age groups, particularly in the last period, which presented more than 25% of cases reclassified in each category. There was a higher increase in the proportion of reclassified causes in younger age groups in the period.

In general, reclassified IDCDs increased along the analyzed period in all regions. The Northeast region maintained similar proportions over time, although higher than 25% since 2006-2008. This region and the Southeast were responsible for most reclassified IDCD cases in the country. The highest proportions, however, were identified in the Mid-West and South regions in 2015-2017, with 53% and 39%, respectively.

Most investigated deaths were completed in hospitals and other health services, although most IDCD occurred in non-health facilities. Home interviews are also an important source of investigated IDCD. The SVO is the single source responsible for the highest proportion

of reclassified deaths among investigated causes, more than 85% since 2006. Proportions of reclassified ICD-10 increased over the period for all sources, except for the use of linkage between health information systems. Around 70% of deaths due to ICD-10 investigated in health service sources and at home were reclassified to other causes.

Table 2. Proportions of reclassified (%R.) deaths among ill-defined causes according to selected variables. Brazil, 2006-2017.

Variable		2006-2008		2009-2011		2012-2014		2015-2017	
		IDCD	%R	IDCD	%R	IDCD	%R	IDCD	%R
Age*	0-9	9,228	13.9	7,551	19.3	6,071	26.3	5,324	27.4
	10 to 29	12,405	16.8	13,610	24.7	14,033	30.4	14,584	30.5
	30 to 69	116,025	16.4	117,086	20.7	114,202	26.4	121,975	28.2
	70+	146,285	17.6	148,479	18.8	147,489	23.1	158,567	26.1
Place of death*	Hosp/H.fac.	89,203	16.2	93,960	21.2	94,273	26.3	106,668	28.6
	Home	169,961	18.2	166,968	19.6	161,350	24.4	167,657	26.7
	Street	10,014	12.6	11,153	18.2	11,301	25.0	10,811	27.6
	Others	14,072	9.7	14,872	13.9	15,649	19.3	16,474	20.5
Source of investigation**	Hospital	12,268	56.4	20,097	59.9	24,416	70.4	34,931	73.3
	SVO	3,190	90.7	4,199	89.4	3,732	91.3	4,620	86.9
	IML	3,741	37.6	9,298	60.0	13,111	64.2	10,299	70.0
	Linkage	617	51.2	1,813	47.5	2,515	64.7	1,997	42.8
	Home	14,918	70.6	24,887	69.4	25,593	75.4	27,280	76.3
	Committee	1,505	40.1	3,296	40.5	2,747	48.0	3,557	66.8
	Other source	1,388	55.7	2,106	54.8	2,756	67.8	3,670	77.7
	Several sources	15,608	89.0	18,550	77.1	21,013	78.8	22,196	78.6

IDCD: ill-defined causes of death; SVO: autopsy reports from death investigation services; IML: forensic institute;

*Missing IDCD data from age and place of death were not included; **%R. were calculated among IDCD investigated (n = 368,395 in 2006-2017), and missing data were not included.

Tables 3-5 show how CSMFs differs by age group. In children 0 to 9 years old, more than 1,200 deaths due to IDCD were reclassified at each period, 35% or more due to infectious diseases and malnutrition (Table 3). Neonatal-related conditions, lower respiratory infections, congenital anomalies and foreign body were the most frequent causes. Diarrhea was more common in 2006-2008, and had decreasing proportions over time, whereas drowning increased in the period among reassigned CODs.

Table 3. Percentage of specific causes of death detected among reclassified ill-defined causes in children 0-9 years old. Brazil, 2006-2017.

Cause of death (GBD 2017 list)	2006-2008	2009-2011	2012-2014	2015-2017
	(n = 1,287)	(n = 1,457)	(n = 1,595)	(n = 1,457)
Neonatal disorders	16.3	15.9	11.8	13.9
Congenital birth defects	8.6	8.8	9.7	10.7
Lower respiratory infections	8.6	12.8	13.3	10.4
Foreign body	5.7	7.6	9.2	10.3
Diarrheal diseases	9.3	4.0	4.1	4.3
Drowning	1.6	2.4	2.4	3.0
Other infectious disease	1.5	1.4	1.8	2.7
Meningitis	1.1	2.2	2.2	2.1
Other injuries	2.0	1.9	2.3	1.9
Interstitial lung disease	0.1	0.3	1.6	1.6
Cancer	1.3	1.0	0.8	1.6
Falls	1.2	0.8	2.1	1.4
Road injuries	0.9	0.9	1.0	1.3
Other diseases	14.2	13.8	14.5	12.2
GC level 1	9.8	11.6	12.0	11.5
GC level 2	4.4	4.0	5.4	4.9
GC level 3	9.9	7.1	3.6	4.3
GC level 4	3.2	3.3	2.1	2.0
Group A - Infectious, neonatal conditions	41.3	39.7	35.4	34.7
Group B - Chron.	18.5	18.5	22.0	23.7
Group C - Injuries	12.9	15.7	19.4	18.9

Teenagers and young adults aged 10 to 29 years showed a different profile of causes within reclassified IDCD (Table 4). Injuries were detected in 31% of reclassified deaths in 2006-2008, with increased proportions over time. In 2015-2017, homicide was responsible for 20% of deaths due to original IDCD, but road injuries, self-harm and drug use disorders were also important. Maternal disorders were the true causes in 4.1% deaths due to 2,083 IDCD in 2006-2008. Although this proportion decreased to 2.2% in 2015-2017, the number of reclassified IDCD (n=4,449) was higher at that time. Alcohol use disorders, ischemic heart disease and epilepsy were also important causes detected among IDCD.

Ischemic heart disease, diabetes, alcohol use disorders and stroke were the leading causes of death identified among around 108,000 deaths due to IDCD in the 30 to 69 age group. In individuals aged 70 and over, stroke and diabetes were the two leading causes detected in 129,020 of reclassified IDCDs (Table 5).

Table 4. Percentage of specific cause deaths detected among reclassified ill-defined causes in 10-29 years old. Brazil, 2006-2017.

Cause of death (GBD 2017 list)	2006-2008	2009-2011	2012-2014	2015-2017
	(n = 2,083)	(n = 3,356)	(n = 4,271)	(n = 4,449)
Interpersonal violence	10.1	13.1	15.3	19.4
Road injuries	6.2	7.2	7.3	7.0
Self-harm	5.9	5.9	4.6	4.4
Drug use disorders	1.3	1.7	3.2	4.0
Other infectious diseases	4.7	3.8	3.5	3.7
Ischemic heart disease	2.9	2.8	3.1	2.9
Epilepsy	2.9	2.8	2.9	2.8
Other injuries	3.2	2.5	3.1	2.5
Alcohol use disorders	4.1	3.5	3.6	2.4
Lower respiratory infections	2.7	4.3	2.1	2.4
Maternal disorders	4.1	2.9	2.8	2.2
Neoplasms	2.8	2.1	1.7	2.0
Drowning	2.5	2.1	2.4	1.9
Other diseases	18.0	16.1	15.4	15.7
GC level 1	14.0	14.8	16.4	16.1
GC level 2	6.6	6.6	5.7	4.8
GC level 3	4.9	3.4	3.0	2.2
GC level 4	3.0	4.4	4.1	3.5
Group A-Infectious, maternal disorders	14.3	12.9	9.9	9.6
Group B- Chron.	27.2	25.8	25.7	25.7
Group C- Injuries	30.1	32.1	35.3	38.1

Table 5. Percentage of specific cause deaths detected among reclassified ill-defined causes in adults (30-69 years old) and elderlies (70 and over). Brazil, 2006-2017.

Cause of death (GBD2017 list)	30 to 69 years				70+ years			
	2006- 2008	2009- 2011	2012- 2014	2015- 2017	2006- 2008	2009- 2011	2012- 2014	2015- 2017
Isquem. heart dis.	12.2	12.0	12.7	13.7	8.6	8.1	8.2	9.1
Diabetes	7.4	7.2	7.5	8.6	9.4	10.6	11.1	10.6
Alc. use disorders	7.0	7.6	7.8	6.9	0.8	1.0	0.8	0.7
Other neoplasms	6.0	4.5	4.6	4.9	4.3	4.0	4.1	4.4
Stroke	7.2	5.4	5.2	4.9	15.4	12.9	12.6	11.3
Cirrhosis	3.0	3.1	3.1	2.8	0.6	0.5	0.5	0.5
Int. violence	1.1	1.5	1.9	2.0	0.0	0.1	0.1	0.1
Card. circ. dis.	1.4	2.1	2.1	1.8	1.0	1.2	1.4	1.5
Chr obst. pulm d.	1.8	1.5	1.5	1.7	3.0	3.0	3.2	3.5
Road injuries	1.1	1.5	1.7	1.7	0.1	0.2	0.2	0.3
Lower resp. inf.	1.6	2.3	1.7	1.7	2.1	2.4	2.3	2.8
Hyp. heart disease	3.3	2.2	1.7	1.7	5.7	4.8	3.5	3.1
Other diseases	20.2	20.1	19.8	19.7	19.4	20.1	21.3	21.7
GC Level 1	7.2	8.4	9.1	9.4	7.9	7.2	7.1	7.1
GC Level 2	13.0	14.0	13.5	13.0	16.1	18.7	18.3	17.6
GC Level 3	5.4	5.1	4.3	4.0	5.1	4.8	4.5	4.2
GC Level 4	1.2	1.5	1.7	1.6	0.5	0.5	0.6	1.3
Group A – Inf, Mat,	6.7	6.7	5.9	5.6	5.8	5.8	5.7	5.6
Group B - Chron	60.4	56.8	57.0	58.7	62.3	59.8	59.9	60.7
Group C - Injur	6.2	7.5	8.5	7.8	2.5	3.2	3.8	3.4

Number of total deaths in 30 to 69 years and 70+ years were, respectively: 2006-2008: 19,014 and 25,757; 2009-2011: 24,204 and 27,870; 2012-2014: 30,098 and 34,018; 2015-2017: 34,362 and 41,375.

DISCUSSION

This paper examines the percentage allocation of specific causes derived from 257,367 ill-defined deaths reclassified after review in health services or home investigation in Brazil from

2006 to 2017. This is, to our knowledge, the first study that has quantified the real proportional distribution of reassigned causes within IDCD in this extent. The similarity and plausibility of proportions of specific causes detected among IDCD after investigation during the last 11 years indicate reliable data quality of the review process.

Apart from the lower proportion of IDCD in recent years (less than 6% of the total on average in 2013-2017), these causes are still important in numbers – 70,763 deaths in 2017⁽¹⁾ (IDCD classified according the GBD 2017 list of causes), and could affect our interpretation of some specific trends. In the GBD study, IDCD and other GCs are reallocated to valid and specific underlying COD using various redistribution algorithms^{4,7,8}. IDCD classified as ill-defined causes among these so-called GCs were redistributed onto all causes on a pro-rata basis⁴, probably on the assumption that the proportional redistribution reflects the most likely probability for correct attribution to a specific cause. However, previous studies comparing proportional distributions of causes derived from IDCD reclassified after investigation in Brazil and causes reported in the death registration system show different distributions⁹⁻¹¹. Thus, this study has been built and expanded on a previous work, and may contribute to further research in the evaluation of the garbage code redistribution method employed by the GBD study for those deaths. The IDCD investigation, in addition to direct results on improving quality of vital statistics, may also be considered for redistribution and perhaps is more accurate than indirect methods. It should be mentioned, however, that results for more disaggregated age groups, sex and locations, which are important for redistribution purposes⁸, are not presented in this study. Further detailed analysis of these data would lead to more accurate assessments of the IDCD investigation results.

Our results show that proportions of investigated and reclassified IDCD in the country increased over time, with a slight reverse in the trend in 2014-2015, and a major increase in 2016-2017. The latter increase may have been due to the Bloomberg Data for Health Initiative (D4H project), started in 2016 with a pilot study in seven cities and launched in 2017 in 60 municipalities of the country. This initiative included several interventions, such as the investigation of a shortlist of GCs (comprising around 80% of occurrences of GCs in the country) in health services to assign the real cause-of-death, the training of physicians in hospitals in medical certification of COD, the production of an app for smartphones for on-line training called AtestaDO, as well as the holding of several technical meetings about quality of information on COD. The process indicators we mentioned above seemed inversely correlated with the outcome of interest, the percentage of IDCD after investigation that decreases from 7.7% in 2006 to 5.4% in 2017.

The initial reduction of deaths due to IDCD over time could be related in part to improvements in the socioeconomic level¹². There is evidence of association at state level of higher per capita Gross Domestic Product (GDP), lower levels of inequality (represented by Gini index) and higher education to lower proportions of IDCD from 1998 to 2012¹³. This proportion, also considered an indicator of the quality of medical care provided and access to health services¹⁴, most likely could be related to the expansion of the Family Health Strategy,

⁽¹⁾ According the GBD 2017 list of causes.

which increased basic healthcare services offered to residents in poorer areas^{15,16}. It is noteworthy that, besides the decrease in the proportions of deaths assigned to IDCD before investigation, increasing proportions reassigned to other causes after investigation had a direct and important contribution to the final decrease, particularly in 2016-2017.

Although 55.4% (n = 167,657) of deaths due to IDCD occurred at home, hospital and autopsy data from SVOs and forensic institutes were the most important sources used in the investigation process. Besides, we could observe an increase over time in hospital revisions and also in home deaths. In fact, the proportions of home interviews would have been even higher if we consider that they might have been included among those investigated through several sources⁶. Although causes of death detected through verbal autopsy are considered less valid⁸, most home interviews have information of the deceased's health experience¹⁷. Higher proportion of reclassified deaths were found after investigations at death verification services (SVO), which probably was due to more detailed data of a higher quality, as had been expected.

In general, considering all age groups, IDCD were reclassified to chronic diseases (56.6%), injuries (7.2%), and infectious (5.2%) or neonatal, maternal, malnutrition (1.7%) from the GBD Level 1 cause list. Although about one-third of reclassified IDCD were reassigned to other garbage codes, only 8.4% of these (discounting R codes) were reclassified to a different Level 1 category (data not presented). Garbage codes from Levels 2-4 are considered less serious than Level 1 codes in terms of introducing bias with respect to the true pattern of CODs in the population¹⁸.

In children, our results show that a substantial proportion of IDCD was reclassified to neonatal disorders, congenital birth defects and lower respiratory infections, in all four calendar periods. It is important to note that foreign body caused about 10% of child deaths detected among IDCD in the last period, and this important proportion was different from the lower proportion observed in the routine death registration before investigation. The investigation results are in accordance with other findings indicating aspiration as one of the most important causes of accidental death of children¹⁹.

A very different profile of causes emerged among reclassified IDCD in teenagers and young adults. Proportions of deaths due to injuries among those aged 10 to 29 years increased from 31.3% in 2006-2008 to 42.1% in 2015-2017. Had they not been included in vital statistics, interpersonal violence, road injuries, self-harm and drug use disorders would be the most important misdiagnosed causes. These findings reinforce previous results and the advocacy of the non-use of proportional redistribution of IDCD excluding external causes^{9,11}.

Alcohol use disorders were the third leading cause detected among IDCD reclassified in adults aged 30 to 69 years in all four calendar periods. Cirrhosis, another important disease attributed to alcohol consumption, with the incidence associated with the amount of alcohol consumed even at moderate levels of consumption²⁰, was also among the leading causes detected in this age group.

Oldest adults in 70 or older age groups had around one third of IDCD reassigned to diabetes, stroke and ischemic heart disease after investigation. Although the investigation

may overestimate the importance of diabetes, as found in a study in Thailand using verbal autopsy²¹, the proportions were similar in all four calendar periods indicating no procedural changes in the investigation process over time. A comparative analysis of findings from this study and the GBD estimates in 2017²² present some important differences in this age group. The most notable is that lower respiratory infections were ranked as the fourth leading COD, responsible for around 8% of total deaths, while our findings show lower proportion in 2015-2017, 2.8%. Lower numbers of deaths assigned to pneumonia in the IDCD investigation may be because reviewers of medical records were probably more likely to attribute an underlying chronic condition or other infectious/external disease. In contrast, neoplasms were assigned in a lower proportion (5.8%) among IDCD than among routine death registration, confirming previous findings that cancer is less likely to be classified to ill-defined categories, compared with other natural causes^{9,11}.

This study has some limitations. First of all, the validity of causes of death reassigned after 368,395 investigations over 11 years could be subject to information bias, as several professional health workers from more than 5,000 municipalities in the country collected information from health service records and at home. In fact, the investigation process of IDCD was not homogeneous in the country, and the municipalities could start the investigation using different data sources, except for verbal autopsy, which was proposed as the last source to be used³. Besides different procedures for data collection, such as non-official forms of verbal autopsy⁶, different quality of those sources could have occurred in several states and over time. Although medical records in hospitals might facilitate the adequate ascertainment of underlying CODs, sometimes they do not contain sufficient data to support clinical diagnoses, particularly when the death occurs soon after admission²³. Moreover, a previous study found that in some municipalities physicians were not responsible for the certification of causes of death after verbal autopsy interviews¹⁷. In addition, it is well known that verbal autopsy data has a varying precision over time²⁴. Despite these issues, our data seemed plausible, as proportions of specific causes of death among IDCD after investigation were relatively stable during the past 11 years.

CONCLUSIONS

The significant number of investigations of IDCD performed in Brazil in 2006-2017 and the high proportion of deaths reassigned to other more informative causes indicate the success of this approach in correcting misclassification of important proportions of IDCD data in the country. Records from health facilities and home interviews in Brazil are of a sufficient quality in providing the necessary information for defining specific underlying COD. Thus, there is still room for improvement. Appropriate physician training on the importance of death certification for public health purposes along with more extensive access to health services and better quality of medical care provided would lead to further improvements in cause of death data in the country.

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):171-7.
2. Santo AH. Ill-defined causes of death and unattended deaths, Brazil, 2003. *Rev Assoc Med Bras.* 2008;54(1):23-8. <http://dx.doi.org/10.1590/S0104-42302008000100016>
3. Brasil. Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Análise de Situação em Saúde. Manual para investigação do óbito com causa mal definida. Brasília, DF; 2008.
4. GBD 2017 Causes of Death Collaborators. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet.* 2018;392(10159):1736-88. [https://doi.org/10.1016/S0140-6736\(18\)32203-7](https://doi.org/10.1016/S0140-6736(18)32203-7)
5. Campos D, França E, Loshi RH, Souza MFM. Verbal autopsy for investigating deaths from ill-defined causes in Minas Gerais State, Brazil. *Cad Saúde Pública.* 2010;26(6):1221-33. <http://dx.doi.org/10.1590/S0102-311X2010000600015>
6. Cunha CCD, Vasconcelos AMN, Souza MFM, França E. Assessment of the investigation of deaths from ill-defined causes in the state of Bahia in 2010. *Ciênc Saúde Coletiva.* 2019;30:24(5):1831-44. <http://dx.doi.org/10.1590/1413-81232018245.14852017>
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. *Popul Health Metr.* 2010 May 10;8:9. <http://dx.doi.org/10.1186/1478-7954-8-9>
8. GBD 2016 Causes of Death Collaborators. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet.* 2017;390(10100):1151-210. [http://dx.doi.org/10.1016/S0140-6736\(17\)32152-9](http://dx.doi.org/10.1016/S0140-6736(17)32152-9)
9. Mello Jorge MHP, Gotlieb SLD, Laurenti R. The national mortality information system: problems and proposals for solving them I – Deaths by natural causes. *Rev Bras Epidemiol.* 2002;5(2):197-223. <http://dx.doi.org/10.1590/S1415-790X2002000200007>
10. França EB, Souza FM, Ishitani LH, Teixeira R, Szwarcwald CL. Strengthening vital statistics in Brazil: investigation of ill-defined causes of death and implications on mortality statistics. *Lancet.* 2013;381(Suppl 2):S51. [http://dx.doi.org/10.1016/S0140-6736\(13\)61305-7](http://dx.doi.org/10.1016/S0140-6736(13)61305-7)
11. França E, Teixeira R, Ishitani L, Duncan BB, Cortez-Escalante JJ, Morais Neto OL, Szwarcwald CL. 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>
12. World Bank. Brazil overview [Internet]. 2016 [cited 2016 June 4]. Available from: <http://www.worldbank.org/en/country/brazil/overview>
13. Rodrigues NCP, Daumas RP, Almeida AS, O'Dwyer G, Andrade MKN, Flynn MB, Lino VTS. Risk factors for the ill-defined causes of death in the Brazilian states: a multilevel analysis. *Ciênc Saúde Coletiva.* 2018;23(11):3979-88. <http://dx.doi.org/10.1590/1413-812320182311.27182016>
14. Martins DF Jr, Costa TM, Lordelo MS, Felzemburg RDM. Trends of mortality from ill-defined causes in the Northeast region of Brazil, 1979-2009. *Rev Assoc Med Bras.* 2011;57(3):338-46. <http://dx.doi.org/10.1590/S0104-42302011000300019>
15. Macinko J, Harris MJ. Brazil's Family Health Strategy – delivering community-based primary care in a universal health system. *N Engl J Med.* 2015;372(23):2177-81.
16. 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 June 29. 10:380. <http://dx.doi.org/10.1186/1471-2458-10-380>
17. França EB, Cunha CC, Vasconcelos AMN, Cortez Escalante JJ, Abreu DX, Lima RB, Morais Neto OL. 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):119-34. <http://dx.doi.org/10.1590/1415-790X201400010010ENG>
18. Mikkelsen L, Richards N, Lopez AD. Redefining 'garbage codes' for public health policy: Report on the expert group meeting, 27-28 February 2017 [Internet]. Melbourne: University of Melbourne; 2018. Available from: <https://crvs.gateway.info/file/9816/276>
19. Rovin JD, Rodgers BM. Pediatric Foreign Body Aspiration. *Pediatr Rev.* 2000;21(3):86-90. <http://dx.doi.org/10.1542/pir.21-3-86>
20. Simpson RF, Hermon C, Liu B, Green J, Reeves GK, Beral V, et al. Alcohol drinking patterns and liver cirrhosis risk: analysis of the prospective UK Million Women Study. *Lancet.* 2019;4(1):e41-8. [https://doi.org/10.1016/S2468-2667\(18\)30230-5](https://doi.org/10.1016/S2468-2667(18)30230-5)
21. Polprasert W, Rao C, Adair T, Pattaraarchachai J, Porapakkham Y, Lopez AD. Cause-of-death ascertainment for deaths that occur outside hospitals

- in Thailand: application of verbal autopsy methods. *Popul Health Metr.* 2010 May 18;8:13. <https://doi.org/10.1186/1478-7954-8-13>
22. Institute for Health Metrics and Evaluation. GBD Compare Data Visualization [Internet]. 2018 [cited 2019 June 8]. Available from: <http://vizhub.healthdata.org/gbd-compare>.
23. Pattaraarchachai J, Rao C, Polprasert W, Porapakkham Y, Pao-In W, Singwerathum N, Lopez AD. Cause-specific mortality patterns among hospital deaths in Thailand: validating routine death certification. *Popul Health Metr.* 2010 May 18;8:12. <https://doi.org/10.1186/1478-7954-8-12>
24. Setel PW, Sankoh O, Rao C, Velkoff VA, Mathers C, Gonghuan Y, et al. Sample registration of vital events with

verbal autopsy: a renewed commitment to measuring and monitoring vital statistics. *Bull World Health Organ.* 2005;83(8):611-7. Epub 2005 Sept 22. <https://dx.doi.org/S0042-96862005000800015>

Received: 06/14/2019

Final version presented: 08/07/2019

Accepted: 08/08/2019

Authors' contribution: França EF and Ishitani LH contributed with data analysis and wrote the manuscript. Teixeira RA, Cunha CC and Marinho MF contributed with interpretation of results and reviewed the manuscript. All authors read and approved the final manuscript.

