

Redistribution of heart failure deaths using two methods: linkage of hospital records with death certificate data and multiple causes of death data

Redistribuição de óbitos por insuficiência cardíaca usando dois métodos: *linkage* de registros hospitalares com dados de atestados de óbito e de causas múltiplas de morte

Redistribución de las muertes por fallo cardíaco usando dos métodos: enlace a registros hospitalarios con datos del certificado de defunción y sobre causas múltiples de la muerte

Ana Luiza Bierrenbach ¹
Gizelson Pereira Alencar ²
Cátia Martinez ³
Maria de Fátima Marinho de Souza ⁴
Gabriela Moreira Policena ⁵
Elisabeth Barboza França ⁶

doi: 10.1590/0102-311X00135617

Abstract

Heart failure is considered a garbage code when assigned as the underlying cause of death. Reassigning garbage codes to plausible causes reduces bias and increases comparability of mortality data. Two redistribution methods were applied to Brazilian data, from 2008 to 2012, for decedents aged 55 years and older. In the multiple causes of death method, heart failure deaths were redistributed based on the proportion of underlying causes found in matched deaths that had heart failure listed as an intermediate cause. In the hospitalization data method, heart failure deaths were redistributed based on data from the decedents' corresponding hospitalization record. There were 123,269 (3.7%) heart failure deaths. The method with multiple causes of death redistributed 25.3% to hypertensive heart and kidney diseases, 22.6% to coronary heart diseases and 9.6% to diabetes. The total of 41,324 heart failure deaths were linked to hospitalization records. Heart failure was listed as the principal diagnosis in 45.8% of the corresponding hospitalization records. For those, no redistribution occurred. For the remaining ones, the hospitalization data method redistributed 21.2% to a group with other (non-cardiac) diseases, 6.5% to lower respiratory infections and 9.3% to other garbage codes. Heart failure is a frequently used garbage code in Brazil. We used two redistribution methods, which were straightforwardly applied but led to different results. These methods need to be validated, which can be done in the wake of a recent national study that will investigate a big sample of hospital deaths with garbage codes listed as underlying causes.

Heart Failure; Death Certificates; Mortality; Medical Record Linkage; International Classification of Diseases

Correspondence

A. L. Bierrenbach
Rua Dona Adma Jafet 91, São Paulo, SP 01308-050, Brasil.
albierrenbach@yahoo.com.br

¹ Instituto Sírio-Libanês de Ensino e Pesquisa, São Paulo, Brasil.

² Escola de Saúde Pública, Universidade de São Paulo, São Paulo, Brasil.

³ Centro de Informações Estratégicas em Saúde, Secretaria de Saúde de São Paulo, São Paulo, Brasil.

⁴ Secretaria de Vigilância em Saúde, Ministério da Saúde, Brasília, Brasil.

⁵ Instituto de Patologia Tropical e Saúde Pública, Universidade Federal de Goiás, Goiânia, Brasil.

⁶ Faculdade de Medicina, Universidade de Minas Gerais, Belo Horizonte, Brasil.



Introduction

To establish public health priorities, accurate information on leading causes of death is of great importance ¹. However, to discriminate whether the patterns observed across time or demographic groups reflect true epidemiological differences or are simply variations in completeness and quality of cause of death assignment can be challenging when interpreting cause-of-death statistics ².

The widespread use of non-specific codes – the so-called “garbage codes” listed as the underlying cause of death on death certificates – is a major limitation to extract such information from vital registration mortality data ³. Heart failure is considered a garbage code when assigned as the underlying cause in death certificates. Heart failure is the end stage of many cardiac and non-cardiac diseases, which should be assigned as the underlying cause, and heart failure should only be included as an intermediate cause of death.

Approaches to better understand heart failure and other ill-defined causes to the actual or probable underlying causes of death have redistributed such deaths using algorithms and statistical methods, using linked patient data from in-patient hospitalizations and actively investigating their medical history and circumstances of death ^{4,5,6,7,8,9,10,11,12}. Knowledge about the actual or probable underlying causes of heart failure would improve the estimates of the mortality risk from such causes and, therefore, provide better inputs to public health planning and policy making ^{13,14,15}.

In this study, we use two different methods to redistribute heart failure deaths to specific causes of death. The first one – method with multiple causes of death – is based on the proportion of underlying causes found in death certificates that have heart failure listed as an intermediate cause, as proposed by Stevens et al. ¹⁰ and Snyder et al. ⁹. The second one – the hospitalization data method – redistributes heart failure deaths based on the information of the hospitalization records of decedents, similarly to the method used by Teixeira et al. ¹¹ and Cascão et al. ¹⁶. We then compared the results of both methods aiming to understand which of them would be appropriate for the routine mortality data analysis in the Brazilian context.

Methodology

We used the nationwide databases of the Brazilian Mortality Information System (SIM) and the Brazilian Hospital Information System (SIH) from 2008 to 2012. The SIM database was obtained directly from the Brazilian Ministry of Health, with personally identifiable information. Data are based on death certificates, which are completed by physicians. Death certificates in Brazil use the international form recommended by the World Health Organization (WHO), into which the “multiple causes of death” have to be stated, starting with the “immediate cause,” following the logic sequence of “intermediate causes” up to the “underlying cause of death,” i.e. the disease or injury that started the sequence of events leading to death. SIM has data on these multiple causes of death coded with the International Classification of Diseases, 10th revision (ICD-10, <http://apps.who.int/classifications/icd10/browse/2010/en>). It also has data on age, sex, education, race, date of death and residential address. SIM coverage and information quality have rapidly increased over the last two decades. Coverage has been over 95% since 2000 ¹⁷.

The SIH database was also obtained directly from the Ministry of Health and contained personally identifiable information. It encompasses all hospitalizations funded by the Brazilian Unified National Health System (SUS), which represented 65.7% of all hospitalizations in the country in 2013 ¹⁸. As the SIH database is primarily used for reimbursement purposes, the likelihood that hospitalizations would be underreported is small ¹⁹. The variables include age, sex, admission and discharge dates, residential address, hospital code, and ICD-10 codes for primary and secondary diagnoses, as well the cause of death. Trained hospital clerks are responsible for extracting such information from the hospital charts. The primary diagnosis, which is the one used in this study, is the condition that caused the hospitalization (DATASUS – Brazilian Health Informatics Department. Notas técnicas. Morbidade hospitalar no SUS. <http://tabnet.datasus.gov.br/cgi/sih/midescr.htm>).

The database linkage between SIM and SIH was performed in a two-step process. First, an in-house probabilistic linkage algorithm was developed, in which the probability that two records

belong to the same patient was based on his/her name and date of birth. Municipality of residence was used as a blocking variable. For each record, a Bloom Filter was constructed following the methods developed by Schnell et al.²⁰ A secret random permutation was applied to the result, using two cryptography hash algorithms. The two linear functions were the simplest possible, the first was the result of applying one of the algorithms and the second was the sum of both results. To increase the specificity of the matches found using the first step, a deterministic linkage was performed with an algorithm similar to the ones validated by Pacheco et al.²¹ and Oliveira et al.²² This algorithm was developed using a set of rules based on the combination of several key variables (name, date of birth, address, zip code, hospital code, telephone, etc.) and their fragments, including those obtained with Soundex transformations and substrings.

Among 5,829,167 mortality records available on SIM from 2008 to 2012, we excluded the following: (1) 1,773,624 with age < 55 years; (2) 179,589 with missing age; (3) 69,379 with underlying cause of death due to external causes (ICD-10 V00-Y98); and (4) 483,719 with the following ill-defined/garbage codes as underlying causes of death (ICD-10 R00-R99, N17-N19, C76, C80, C97, I10, I46, I47.2, I49.0, I51.4, I51.5, I51.6, I51.9, I70.9). After these exclusions, 3,322,856 mortality records remained on the database for analysis.

Two methods were applied to redistribute heart failure deaths. In the method of multiple causes of death, we replicated the methodology proposed by Stevens et al.¹⁰ and Snyder et al.⁹, introducing only some small adaptations. We excluded decedents < 55 years and those with missing information on age. We also excluded decedents with underlying causes of death (UCD) due to external causes (ICD-10 V00-Y98) and due to certain ill-defined/garbage code causes (ICD-10 R00-R99, N17-N19, C76, C80, C97, I10, I46, I47.2, I49.0, I51.4, I51.5, I51.6, I51.9, I70.9), as they would not be plausible UCDs of a heart failure death.

Each of the records with heart failure listed as UCD was matched to the records that had heart failure listed as one of the multiple causes of death, by exact matching. Records were matched by five-year age groups (55 to 94 years old, and then 95+ years old), sex, race (white, black/brown and other/unknown), education (up to 7 years, 8 years and above, unknown), and macroregion of residence (North, Northeast, Central, Southeast and South).

The following grouping of causes of death proposed by Snyder et al.⁹ were used, and we also included Chagas disease: coronary heart disease (CHD; I20-I25), chronic obstructive pulmonary disease (COPD; J40-J44), diabetes (E10-E14), hypertensive heart and kidney disease (I11-I13), cardiomyopathy (I42-I43), cerebrovascular disease (I60-I69), other cardiovascular diseases (I00-I99, except I50), lower respiratory infections (J10-J18, J20-J22), cancers (C00-C97), digestive diseases (K20-K92), dementia (F01, F03), Alzheimer's disease (G30), Chagas disease (B57), other diseases (remaining causes)⁹.

Then, each heart failure death was proportionally redistributed to the frequency of the UCD observed in the matched records. For example, if a heart failure record of a given profile of age/sex/race/education and region matched to ten CHD deaths of the same profile with 3 of them having CHD, 3 hypertensive disease, 2 Chagas, 1 cardiomyopathy and 1 diabetes as UCD; the UCD for the heart failure record was distributed as 0.3 CHD, 0.3 hypertensive disease, 0.2 Chagas, 0.1 cardiomyopathy, and 0.1 diabetes.

In the method of hospitalization data, each case of heart failure death that occurred inside a hospital and was linked to a SIH record with a date of exit that was equal or had a one-day difference with the SIM date of death was redistributed based on the ICD-10 code listed as the SIH principal diagnosis or cause of death. All SIH records generated as a result of hospital discharge have a principal diagnosis. A diagnosis for the cause of death is available for patients that died during hospitalization.

Differently from the causes of death on SIM records, which are based on death certificates completed by physicians, the principal diagnosis and cause of death on SIH records are extracted from hospital charts by trained clerks, who usually do not have special training on the ICD-10 coding and logic of UCD selection.

The principal diagnosis and the cause of death diagnosis were grouped into four additional groups: Garbage codes (ICD-10 N17-N19, C76, C80, C97, I10, I46, I47.2, I49.0, I51.4, I51.5, I51.6, I51.9, I70.9 – the same codes used in the exclusion process), Ill-defined codes (ICD-10 chapter XVIII – the same codes used in the exclusion process), Codes not allowed as cause of death (ICD-10 chapters

XIX, XXI, XXII, plus isolated codes) and Missing codes, as there were some linked records that did not have a cause of death diagnosis (DATASUS. CID-10 sub categorias. <http://www.datasus.gov.br/cid10/V2008/descrcsv.htm>).

This study was included into the GBD Brazil 2015 (Global Burden of Disease) study approved by the Research Ethics Committee of the Universidade Federal de Minas Gerais (CAAE Project 62893316.7.0000.5149).

Results

Our study population comprises 3,322,856 decedents older than 55 years who had SIM records from 2008 and 2012 and for whom the UCD was not an external cause or an ill-defined/garbage code. Heart failure was listed as the UCD in 123,268 (3.7%) of these records and as a multiple cause of death in 233,197 (7%). Heart failure was not listed as the underlying cause of death or as a multiple cause of death in 2,966,391 records (Table 1).

Table 1 shows that deaths that occurred in hospitals comprised 74% of all records of our study population. The proportion of deaths that occurred in hospitals was higher for deaths with heart failure listed as the UCD (72.2%) in comparison with deaths with heart failure listed as a multiple cause of death (65.9%). Deaths linked to hospitalization records with a date of exit that was equal or had a one-day difference in comparison with the date of death, comprised 43.7% of hospital deaths. The proportion of linked records was higher for deaths with heart failure listed as the UCD (46.4%) in comparison with deaths with heart failure listed as a multiple cause of death (43.3%).

Among the 123,268 records with heart failure listed as the UCD, the method of multiple causes of death redistributed 25.3% to hypertensive heart and kidney diseases, 22.6% to coronary heart diseases and 9.6% to diabetes. The other groups of causes contributed less. Even though the proportion of records redistributed to hypertensive heart and kidney diseases and to coronary heart diseases decreased when the 89,001 deaths that occurred in hospitals and the 41,324 deaths that occurred in hospitals and that were linked to SIH records were considered, these two groups of causes were still the most predominant ones (Table 2).

Table 3 shows two sets of redistribution results based on the principal diagnosis and on the cause of death of SIH for the 41,324 heart failure deaths that died in hospitals and whose records were linked to hospitalization, each one on a separate column. The first data column shows that heart failure was listed as the principal diagnosis in 45.8% of these records. No redistribution occurred for such records. For the remaining ones, data from the principal diagnosis in the corresponding hospitalization record redistributed 21.2% to the group of other diseases, 6.5% to lower respiratory infections and 9.3% to a mix of garbage codes or other ill-defined causes or to codes that are not allowed as

Table 1

Distribution of heart failure deaths listed as the underlying cause of death, as a multiple cause of death or not mentioned as a cause of death, by site of occurrence and linkage status with hospitalization records.

Variables	Deaths that occurred in all sites (%) *	Deaths that occurred in hospitals (%) **	Hospital deaths linked to hospitalization records (%) ***
Heart failure as underlying cause of death	123,268 (3.7)	89,001 (72.2)	41,324 (46.4)
Heart failure as multiple causes of death	233,197 (7.0)	153,578 (65.9)	66,558 (43.3)
Heart failure not mentioned as a cause of death	2,966,391 (89.3)	2,217,787 (74.8)	967,761 (43.6)
Total	3,322,856	2,460,366 (74.0)	1,075,643 (43.7)

* Denominator = total deaths;

** Denominator = deaths that occurred in all sites;

*** Denominator = deaths that occurred in hospitals.

Table 2

Redistribution proportions among death records with heart failure listed as the underlying cause of death, by site of occurrence and linkage status with hospitalization records. Redistribution method based on multiple causes of death.

Groups of causes of death	Deaths that occurred in all site (%) [n = 123,328]	Deaths that occurred in hospitals (%) [n = 89,001]	Hospital deaths linked to hospitalization records (%) [n = 41,324]
Alzheimer's disease	0.7	0.5	0.4
Cancers	3.2	3.4	3.4
Cardiomyopathy	7.0	7.9	8.4
Cerebrovascular diseases	4.0	4.1	4.4
Chagas disease	3.0	3.3	3.6
Chronic obstructive pulmonary disease	5.8	6.3	7.1
Coronary heart disease	22.6	21.4	18.4
Dementia	0.1	0.1	0.1
Diabetes	9.6	9.7	9.7
Digestive diseases	1.7	2.2	2.5
Hypertensive heart and kidney diseases	25.3	21.2	21.0
Lower respiratory infections	3.0	3.8	4.5
Other cardiovascular diseases	7.8	9.1	9.2
Other diseases	6.2	7.0	7.3
Total	100.0	100.0	100.0

Table 3

Redistribution proportions among the 41,324 death records with heart failure listed as the underlying cause of death and whose records were linked to hospitalization records. Redistribution method based on data from corresponding hospitalization records.

Groups of causes	Redistribution based on the principal diagnosis of the SIH record (%)	Redistribution based on the cause of death of the SIH record (%)
Heart failure	45.8	32.6
Alzheimer's disease	0.0	0.0
Cancers	0.6	0.4
Cardiomyopathy	1.6	1.1
Cerebrovascular diseases	2.8	2.0
Chagas disease	0.0	0.0
Chronic obstructive pulmonary disease	2.9	1.7
Coronary heart disease	2.1	1.3
Dementia	0.0	0.0
Diabetes	1.0	0.6
Digestive diseases	2.3	1.2
Hypertensive heart and kidney diseases	0.1	0.2
Lower respiratory infections	10.2	6.5
Other cardiovascular diseases	3.4	2.5
Other diseases	17.8	21.2
Garbage codes	3.1	8.4
Ill-defined	3.2	11.9
Codes not allowed	3.1	1.2
Missing codes	0.0	7.2
Total	100.0	100.0

SIH: Brazilian Hospital Information System.

causes of deaths (such as codes from Chapter XXI). The other groups of causes contributed considerably less. Overall, only 44.8% of 41,324 deaths were redistributed to a defined cause of death: no heart failure, no garbage/ill-defined, no codes that are not allowed as causes of death (Table 3).

The second data column in Table 3 shows that when data from the cause of death in the hospitalization records was used instead of data from the principal diagnosis, heart failure was listed in 32.6% of records. For the remaining ones, data from the principal diagnosis in the corresponding hospitalization record redistributed 17.8% to the group of other diseases, 10.2% to lower respiratory infections and 21.5% to a mix of garbage codes or other ill-defined causes or to codes that are not allowed as causes of deaths. Also, 7.2% of these linked records did not have any data on the cause of death. The other groups of causes contributed considerably less. Overall, only 38.7% of 41,324 deaths were redistributed to a defined cause of death: no heart failure, no garbage/ill-defined, no codes that are not allowed as causes of death, no missing data (Table 3).

Discussion

Heart failure is a major global health issue, which is associated with high morbidity and mortality and represents a substantial burden to health-care systems. The harmful consequence of the use of heart failure as the UCD is that the information about the true underlying conditions remains obscure. Such information is critical for planning effective public health policies, aiming to decrease the number of such deaths by adequate prevention, early detection and treatment of the underlying conditions ^{6,23,24}.

In Brazil, from 2008 to 2012, heart failure was listed as the UCD in 123,328 deaths of decedents ≥ 55 years old. For 72% of the deaths that occurred in hospitals, it is possible that information about the underlying causes that led to cardiac failure were available from medical charts and assumed to be represented in the principal diagnosis or the cause of death of the hospital information system data. The profile of UCD of the mortality records on SIM that had heart failure listed as one of the multiple causes of death is presumably similar to the profile of true UCD of the records that had heart failure listed as the UCD. The two methods used in this study are based on these assumptions. They are systematic and replicable, which is useful for policymakers to make better use of the existing mortality data. However, they led to very different results, which made us question their mutual validity.

The method of multiple cause of death produced different results from those observed by Snyder et al. in the USA, using a very similar methodology, since over 37% of heart failure deaths were redistributed to CHD and only 4.5% were redistributed to hypertensive heart and kidney diseases ⁹. In our study, 25.3% and 22.6% of the total numbers of these deaths were redistributed to hypertensive heart/kidney diseases and CHD, respectively. These proportions decreased to 21.2% and 21.4% when only hospital deaths were considered, and to 18.4 and 21% when only hospital deaths linked to hospitalization records were considered. Stevens et al. ¹⁰ had already shown these differences in the redistribution patterns of heart failure deaths, with more deaths redistributed to hypertensive heart and kidney diseases in Mexico and Brazil in comparison with the USA.

More importantly, the method of multiple cause of death produced quite different results from those observed with the method that used data from the corresponding hospitalization records. Of the total number of heart failure deaths, we were able to find a corresponding hospitalization record and reclassify only 44.8% to a defined cause of death with the principal diagnosis, and even less, only 38.7%, with the cause of death of SIH. Using SIH data to reclassify UCD, the groups of "other diseases" and "lower respiratory diseases" predominated, while the groups of hypertensive heart/kidney diseases and CHD contributed less.

One possibility to explain these findings is that these hospital deaths with heart failure garbage code listed as the UCD could actually be very peculiar deaths, not due to cardiac and non-cardiac causes, which usually lead to fatal heart failure, such as CHD. If this were true, it would be wrong to use the method of multiple cause of death to redistribute these deaths.

Another possibility is that the physician who certifies the deaths had no or very little information to give his/her diagnosis of the sequence of events that culminated in the patient's death, which reflects on the hospital charts and on SIH records. If this were true, the SIH method would be useless to reclassify these deaths, at least as much as heart failure deaths are concerned

Another possibility is that the information was actually available from the hospital charts for both the physician responsible for completing the death certificate and the hospital clerk responsible for filling out the SIH records, but they had no appropriate motivation or training to perform their respective responsibilities.

SIH has been increasingly used in analyses of important public health issues, even though the reliability of its data for epidemiological use is uncertain and does not represent the totality of hospitalizations in the country¹⁹. Studies conducted in Rio de Janeiro, Brazil, have used hospital charts and/or SIH data to qualify the cause of death data of SIM. Melo et al.²⁵ analyzed acute myocardial infarction deaths that occurred in two hospitals affiliated to SUS. By reviewing hospital charts and SIH records, they confirmed the diagnosis of 67.1% of such deaths. For the remaining ones, either there was no agreement, no corresponding SIH records, or the records did not mention that the patient had died²⁵. Cascão et al.¹⁶ also found an overall good agreement (72.1%) between SIM and the principal diagnosis of SIH when comparing cardiovascular deaths for individuals older than 60 years. However, only 22.6% of the ill-defined deaths could be reclassified to a defined cause based on data from the corresponding SIH records¹⁶. Teixeira et al.¹¹ found a similar proportion of 20%. These proportions are lower than those presented here. Besides other methodological differences, these studies aimed to reclassify all ill-defined deaths, not just the heart failure ones.

It is noteworthy that the use of SIH data to qualify information regarding the underlying causes of natural deaths can be biased by the fact that SIH does not include deaths that happened outside hospitals funded by SUS. Deaths that happened in private hospitals or outside hospitals may have a different epidemiological pattern than those from public hospitals. Our results show that the proportion of deaths due to CHD is lower for hospital deaths than for all deaths regardless of place of occurrence, which is logical, as such deaths can be very sudden. Also, the method is constrained by the limitations of the linkage procedures, which depend on the quality of the nominal information available from both databases. Cascão et al.¹⁶ showed in their study that linked and not-linked records were not that different regarding both demographic characteristics and distribution of causes of death across ICD-10 chapters.

The method of multiple causes of death redistribution presumes the distribution of true underlying causes is similar for deaths with heart failure listed as the underlying cause and as one of the multiple causes of death. While both Stevens et al. and Snyder et al.⁹ validated their findings against primary data, these validations had small numbers, and Stevens et al.¹⁰ only considered their validation as preliminary. Also, the quality of health care and medical certification is much more varied in such a diverse country as Brazil, and such variability would most likely be an important influence in the geographical redistribution of heart failure deaths²⁶.

In the Brazilian scenario, the Ministry of Health is just about to conduct a study in many municipalities, where hospital deaths with garbage codes listed as UCD will be fully investigated. Research assistants will extract information from hospital charts using standardized questionnaires and interview the decedents' family members with verbal autopsy questionnaires, which will then be reviewed by physicians to identify the true UCD. Then, it will be possible to validate the two automated methods used in this study in comparison with our results.

The Ministry of Health has put a lot of effort over the last decades to improve the data from SIM aiming to adequately measure the mortality burden²⁶, and it is imperative to improve the quality of data registered in SIH for epidemiological analyses, using standardized criteria for issuing hospital admission forms and training hospital staff on the proper use of the ICD-10 codes.

Thus, heart failure is a frequently used garbage code in Brazil, which should be reassigned to specific causes of death to increase the validity of mortality statistics. We used two redistribution methods, which proved to be practical, but led to very different results, therefore producing disputable evidence. There is an opportunity in Brazil to validate these methods with data from a study that will review records of many hospital deaths with garbage codes listed as UCD.

Contributors

A. L. Bierrenbach and E. B. França contributed in the conception and design, data acquisition, analysis, and interpretation, drafting of the article, final approval of the version to be published and are responsible for all aspects of the study. G. P. Alencar, C. Martinez and G. M. Policena contributed in the data interpretation, critical review for important intellectual content, final approval of the version to be published and are responsible for all aspects of the study. M. F. M. Souza contributed in the data acquisition and interpretation, critical review for important intellectual content, final approval of the version to be published and is responsible for all aspects of the study.

Additional informations

ORCID: Ana Luiza Bierrenbach (0000-0002-6837-0636); Gizelton Pereira Alencar (0000-0002-2354-9050); Cátia Martinez (0000-0001-6936-1152); Maria de Fátima Marinho de Souza (0000-0003-3287-9163); Gabriela Moreira Policena (0000-0002-3191-1435); Elisabeth Barboza França (0000-0001-6984-0233).

Acknowledgments

We thank Antony Stevens for the linkage procedures that enabled this study (Estudo GBD 2015 Brasil – Fundo Nacional de Saúde process n. 25000192049/2014-14).

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; 83:171-7.
2. AbouZahr C, Boerma T. Health information systems: the foundations of public health. *Bull World Health Organ* 2005; 83:578-83.
3. 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; 8:9.
4. Ahern RM, Lozano R, Naghavi M, Foreman K, Gakidou E, Murray CJ. Improving the public health utility of global cardiovascular mortality data: the rise of ischemic heart disease. *Popul Health Metr* 2011; 9:8.
5. Cavalini LT, Ponce de Leon AC. Correção de sub-registros de óbitos e proporção de internações por causas mal definidas. *Revista Saúde Pública* 2007; 41:85-93.
6. Coady SA, Sorlie PD, Cooper LS, Folsom AR, Rosamond WD, Conwill DE. Validation of death certificate diagnosis for coronary heart disease: the Atherosclerosis Risk in Communities (ARIC) Study. *J Clin Epidemiol* 2001; 54:40-50.
7. Foreman KJ, Naghavi M, Ezzati M. Improving the usefulness of US mortality data: new methods for reclassification of underlying cause of death. *Popul Health Metr* 2016; 14:14.
8. Saito I. Review of death certificate diagnosis of coronary heart disease and heart failure in Japan. *Nihon Koshu Eisei Zasshi* 2004; 51:909-16.
9. Snyder ML, Love SA, Sorlie PD, Rosamond WD, Antini C, Metcalf PA, et al. Redistribution of heart failure as the cause of death: the Atherosclerosis Risk in Communities Study. *Popul Health Metr* 2014; 12:10.
10. Stevens GA, King G, Shibuya K. Deaths from heart failure: using coarsened exact matching to correct cause-of-death statistics. *Popul Health Metr* 2010; 8:6.
11. Teixeira CL, Klein CH, Bloch KV, Coeli CM. Reclassificação dos grupos de causas prováveis dos óbitos de causa mal definida, com base nas Autorizações de Internação Hospitalar no Sistema Único de Saúde, Estado do Rio de Janeiro, Brasil. *Cad Saúde Pública* 2006; 22:1315-24.
12. Franca 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; 48:671-81.
13. Murray CJ, Kulkarni SC, Ezzati M. Understanding the coronary heart disease versus total cardiovascular mortality paradox: a method to enhance the comparability of cardiovascular death statistics in the United States. *Circulation* 2006; 113:2071-81.

14. Khosravi A, Rao C, Naghavi M, Taylor R, Jafari N, Lopez AD. Impact of misclassification on measures of cardiovascular disease mortality in the Islamic Republic of Iran: a cross-sectional study. *Bull World Health Organ* 2008; 86:688-96.
15. Franca E, Rao C, Abreu DM, Souza MF, Lopez AD. Comparison of crude and adjusted mortality rates from leading causes of death in northeastern Brazil. *Rev Panam Salud Pública* 2012; 31:275-82.
16. Cascao AM, Jorge MH, Costa AJ, Kale PL. Use of primary diagnosis during hospitalization in the Unified Health System (Sistema Unico de Saude) to qualify information regarding the underlying cause of natural deaths among the elderly. *Rev Bras Epidemiol* 2016; 19:713-26.
17. Lima EEC, Queiroz BL. Evolution of the deaths registry system in Brazil: associations with changes in the mortality profile, under-registration of death 11 counts, and ill-defined causes of death. *Cad Saúde Pública* 2014; 30:1721-30.
18. Instituto Brasileiro de Geografia e Estatística. Pesquisa Nacional de Saúde 2013: acesso e utilização dos serviços de saúde, acidentes e violências – Brasil, grandes regiões e Unidades da Federação. Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística; 2015.
19. Bittencourt SA, Camacho LAB, Leal MC. O Sistema de Informação Hospitalar e sua aplicação na saúde coletiva. *Cad Saúde Pública* 2006; 22:19-30.
20. Schnell R, Bachteler T, Reiher J. Privacy-preserving record linkage using Bloom filters. *BMC Med Inform Decis Mak* 2009; 9:41.
21. Pacheco AG, Saraceni V, Tuboi SH, Moulton LH, Chaisson RE, Cavalcante SC, et al. Validation of a hierarchical deterministic record-linkage algorithm using data from 2 different cohorts of human immunodeficiency virus-infected persons and mortality databases in Brazil. *Am J Epidemiol* 2008; 168:1326-32.
22. Oliveira GP, Bierrenbach AL, Camargo KRJ, Coeli CM, Pinheiro RS. Accuracy of probabilistic and deterministic record linkage: the case of tuberculosis. *Rev Saúde Pública* 2016; 50:49.
23. Goraya TY, Jacobsen SJ, Belau PG, Weston SA, Kottke TE, Roger VL. Validation of death certificate diagnosis of out-of-hospital coronary heart disease deaths in Olmsted County, Minnesota. *Mayo Clin Proc* 2000; 75:681-7.
24. Mendez GF, Cowie MR. The epidemiological features of heart failure in developing countries: a review of the literature. *Int J Cardiol* 2001; 80:213-9.
25. Melo ECP, Travassos C, Carvalho MS. Qualidade dos dados sobre óbitos por infarto agudo do miocárdio, Rio de Janeiro. *Rev Saúde Pública* 2004; 38:385-91.
26. França E, Abreu DX, Rao C, Lopez AD. Evaluation of cause-of-death statistics for Brazil, 2002-2004. *Int J Epidemiol* 2008; 37:891-901.

Resumo

A insuficiência cardíaca, quando atribuída como a causa básica de morte, é considerada um código lixo. A reatribuição de códigos lixo a causas plausíveis tem por objetivo reduzir viés e aumentar a comparabilidade de dados sobre mortalidade. Dois modelos de redistribuição foram aplicados a dados brasileiros de 2008 a 2012, para pacientes falecidos de 55 anos de idade ou mais. No modelo de causas múltiplas de morte, óbitos por insuficiência cardíaca foram redistribuídos com base na proporção de causas básicas identificadas em óbitos pareados que tinham insuficiência cardíaca listada como causa intermediária. No método de dados hospitalares, óbitos por insuficiência cardíaca foram redistribuídos com base nos dados dos registros de hospitalização dos pacientes falecidos. Houve 123.269 (3,7%) óbitos por insuficiência cardíaca. O método de causas múltiplas de morte redistribuiu 25,3% para doenças cardíacas hipertensivas e doenças renais, 22,6% para doenças cardíacas coronarianas e 9,6% para diabetes. Houve 41.324 óbitos por insuficiência cardíaca relacionados com registros de hospitalização. A insuficiência cardíaca foi listada como o diagnóstico principal em 45,8% dos registros de hospitalização correspondentes. Para estes, não foi feita redistribuição. Para os óbitos remanescentes, o método de dados hospitalares redistribuiu 21,2% para outras doenças (não cardíacas), 6,5% para infecções das vias aéreas inferiores e 9,3% para outros códigos lixo. A insuficiência cardíaca é um código lixo frequentemente usado no Brasil. Nós usamos dois métodos de redistribuição, aplicados de forma simples, mas que levaram a resultados distintos. É importante que esses métodos sejam validados, o que pode ser feito a partir de um estudo nacional recente que investigará uma grande amostra de óbitos hospitalares com códigos lixo listados como causas básicas.

Insuficiência Cardíaca; Atestado de Óbito; Mortalidade; Registro Médico Coordenado; Classificação Internacional de Doenças

Resumen

El fallo cardíaco, cuando es asignado como causa subyacente de la muerte está considerado como código basura. El objetivo de este estudio es reasignar códigos basura de fallecimiento, con el fin de reducir sesgos e incrementar la comparabilidad de los datos de mortalidad. Se aplicaron dos métodos de redistribución en los datos brasileños de 2008 a 2012, para fallecidos de 55 años y mayores. En el método de causas múltiples de muerte, las muertes por fallo cardíaco fueron redistribuidas basándose en la proporción de causas subyacentes encontradas en las muertes compatibles que contaban con un fallo cardíaco descrito como causa intermedia. En el método de datos de hospitalización, las muertes por fallo cardíaco fueron redistribuidas basándose en datos del historial de hospitalización de los fallecimientos. Hubo 123.269 (3,7%) muertes por fallo cardíaco. El método de múltiples causas de fallecimiento redistribuyó un 25,3% a problemas de hipertensión cardíaca y enfermedades de riñón, un 22,6% a enfermedades coronarias de corazón y un 9,6% a diabetes. Hubo 41.324 muertes por fallos cardíacos vinculadas a los registros de hospitalización. El fallo cardíaco fue listado como diagnóstico principal en un 45,8% de los registros de hospitalización correspondientes. Para estos últimos, no se produjo redistribución. En el caso de los restantes, el método de datos de hospitalización redistribuyó un 21,2% a un grupo con otras enfermedades (no-cardíacas), un 6,5% a infecciones en las vías respiratorias bajas y un 9,3% a otros códigos basura. El fallo cardíaco es frecuentemente usado en Brasil como código basura. Usamos dos métodos de redistribución, que fueron directamente aplicados, pero que condujeron a resultados diferentes. Es importante validar estos métodos, que como consecuencia de un estudio nacional recientemente iniciado es posible que se pueda hacer, además de investigar una gran muestra de muertes hospitalarias registradas con códigos basura procedentes de causas subyacentes.

Insuficiencia Cardíaca; Certificado de Defunción; Mortalidad; Registro Médico Coordinado; Clasificación Internacional de Enfermedades

Submitted on 14/Aug/2017

Final version resubmitted on 26/Nov/2018

Approved on 30/Nov/2018