

Reliability of cause of death coding: an international comparison

Concordancia en la codificación de causas de
muerte: una comparación internacional

Confiabilidade de codificação das causas de óbito:
uma comparação internacional

Carmen Antini ¹

Danuta Rajs ²

María Teresa Muñoz-Quezada ³

Boris Andrés Lucero Mondaca ³

Gerardo Heiss ⁴

Abstract

This study evaluates the agreement of nosologic coding of cardiovascular causes of death between a Chilean coder and one in the United States, in a stratified random sample of death certificates of persons aged ≥ 60 , issued in 2008 in the Valparaíso and Metropolitan regions, Chile. All causes of death were converted to ICD-10 codes in parallel by both coders. Concordance was analyzed with inter-coder agreement and Cohen's kappa coefficient by level of specification ICD-10 code for the underlying cause and the total causes of death coding. Inter-coder agreement was 76.4% for all causes of death and 80.6% for the underlying cause (agreement at the four-digit level), with differences by the level of specification of the ICD-10 code, by line of the death certificate, and by number of causes of death per certificate. Cohen's kappa coefficient was 0.76 (95%CI: 0.68-0.84) for the underlying cause and 0.75 (95%CI: 0.74-0.77) for the total causes of death. In conclusion, causes of death coding and inter-coder agreement for cardiovascular diseases in two regions of Chile are comparable to an external benchmark and with reports from other countries.

Mortality; Cause of Death; Clinical Coding; Reproducibility of Results

Resumo

Este estudo avaliou a confiabilidade na codificação das causas de óbitos cardiovasculares entre um codificador no Chile e outro nos Estados Unidos, em uma amostra aleatória estratificada de declarações de óbito de pessoas ≥ 60 anos, emitidas em 2008 nas regiões de Valparaíso e Metropolitana do Chile. Todas as causas da morte foram convertidas em códigos CID-10 em paralelo por ambos os codificadores. A confiabilidade foi avaliada de acordo com o intercodificador e o coeficiente kappa de Cohen, segundo o nível de especificação do código CID-10 para a codificação de causa básica e para todas as causas de óbito. A concordância intercodificador foi de 76,4% para todas as causas de morte e 80,6% para a causa básica (acordo no nível de quatro dígitos), com diferenças por nível de especificação do código CID-10, linha da declaração de óbito, e por número de causas de morte por declaração de óbito. O coeficiente kappa foi 0,76 (IC95%: 0,68-0,84) para a causa básica e 0,75 (IC95%: 0,74-0,77) para todas as causas de óbito. Em conclusão, a codificação das causas de morte cardiovasculares e acordo intercodificador em duas regiões do Chile são comparáveis a uma referência externa e com os relatórios de outros países.

Mortalidade; Causas de Morte; Codificação Clínica; Reprodutibilidade de Resultados

¹ Facultad de Medicina, Universidad de Chile, Santiago, Chile.

² Comité Asesor sobre Estadísticas de Salud, Organización Panamericana de la Salud, Santiago, Chile.

³ Facultad de Ciencias de la Salud, Universidad Católica de Maule, Talca, Chile.

⁴ School of Public Health, University of North Carolina, Chapel Hill, U.S.A.

Correspondence

C. Antini
Escuela de Salud Pública,
Facultad de Medicina,
Universidad de Chile,
Independencia 939, Santiago
8380453, Chile.
cantini@med.uchile.cl

Introduction

Mortality statistics are a cornerstone of health planning and research ^{1,2,3}. Their usefulness hinges on the quality of death certification and on the reliability of coding in line with the International Classification of Diseases – 10th revision (ICD-10) ⁴. Despite extensive standardization efforts, the reliability of nosological coding of causes of death is often a cause for concern ^{3,5,6}.

Apart from possible errors associated with the medical certification of cause of death, a number of factors influence the quality of nosological coding of causes of death. Some derive from ambiguities and inconsistencies of the ICD-10, but arguably most relate to differences in training and skills, and the interpretation of the ICD guidelines by coders ^{3,7,8,9}. Studies of intercoder agreement show varying levels of agreement in relation to code classification level, the number of causes of death recorded on the death certificate, and the age of the deceased ^{3,7,10}.

Chile has made significant efforts in recent decades to improve the quality of its vital records ¹¹ and the international comparability of its causes of death statistics, but there is currently no information available on the quality of recent coding process in the country. The aim of this study was therefore to quantify the level of international agreement between a coder from Chile and a coder from the United States using a representative sample of death certificates issued in 2008 in the Valparaíso and Metropolitan regions, Chile.

Method

As part of a study on mortality due to heart failure, a stratified random sample was drawn from death certificates of people aged 60 years and over issued in the Valparaíso and Metropolitan regions in 2008 available in the Department of Health Statistics and Information (DEIS, acronym in Spanish) of the Chilean Ministry of Health (MINSAL, acronym in Spanish). These regions were selected because they account for approximately 50% of all annual deaths in the country.

Given the low prevalence of heart failure as an underlying cause of death (approximately 2% per year), the sample universe was restricted to the group of ICD-10 codes which corresponds to pathologies frequently associated with heart failure (see codes below) and which also accounts for 47% of the ICD-10 codes for cardiovascular disease (codes I00 to I52). This randomly selected study sample was stratified by sex, age (60 to 69 years, 70 to 79 years, 80 years and over) and

place of death (hospital, home). Based on the age distribution of deaths due to heart failure ¹², the death certificates of individuals aged under 60 years were excluded from the analysis. Deaths due to external causes and those that occurred in a place of death listed as “other” (prisons, nursing homes, lodgings) were also excluded given the greater probability of misclassifications in such cases that could lead to selection bias. Given the significant variation in prevalence of heart failure by age ¹², a stratified random sampling design using Neyman optimal allocation was chosen, with similar sampling fractions that accurately reflect the prevalence of heart failure as the underlying cause of death in each age-sex stratum. Based on the total number of deaths (31,112) in the Valparaíso and Metropolitan regions, a prevalence of heart failure as the underlying cause of death of 6.5%, and a variance of 0.0195, the minimum sample size required to attain the precision target was 437 death certificates. The actual study sample for this analysis was made up of 515 certificates ($\alpha = 0,05$; $Z = 1,96$; $\beta = 80\%$; number of strata = 12).

Chile's mortality database is maintained jointly by the National Statistics Institute, Civil and Identification Registry and the DEIS ¹¹. These institutions collect and verify the information recorded on death certificates, which is standardized according to World Health Organization (WHO) ¹³ recommendations into two parts: Part I (Lines A, B and C), which records the sequence of diseases that led directly to death; and Part II, which includes all other morbid conditions that may have contributed to the death, but are unrelated to the underlying cause ^{4,13}.

Based on the information recorded on the death certificate, the DEIS staff manually select and code the underlying cause of death based on the ICD-10 guidelines, translating the diagnostic terms recorded on the death certificate from words into alphanumeric code. The ICD code consists of a four-character string (with a letter in the first position and numbers in the second, third and fourth positions ¹⁰) and 21 chapters associated with the first character of the code (e.g., I – diseases of the cardiovascular system). The chapters are subdivided into homogeneous “blocks” of three-character categories (e.g., I50) that may represent a singular condition or group of diseases which have common features. Most blocks of three-character categories are subdivided by means of a fourth, numeric character after a decimal point (e.g., I50.9) ⁴. According to WHO recommendations, this fourth character is not required internationally ⁴. Therefore, a condition may be coded at the four-character, three-character or chapter (one-character) level ⁹.

In accordance with the ICD-10 coding rules, each cause of death recorded on the death certificates was independently translated into ICD-10 codes by a coder from the DEIS and a certified nosological coding instructor from the United States. The analysis considered the frequency of non translation of causes of death into ICD coding by both coders. Demographic information about the deceased (age, sex, marital status, education level, region of residence, place of death) and the type of medical certifier (physician, pathologist, other doctors) was recorded.

Level of intercoder agreement (%) at the four-character, three-character or chapter (one-character) level was assessed. Cohen's kappa coefficient and corresponding 95% confidence intervals (95%CI) were calculated. The chi-square test of independence was used to test whether differences reached nominal levels of statistical significance at $p = 0.05$. The data was analysed using SAS 9.2 software (SAS Inst., Cary, USA).

The study was approved by the Human Research Ethics Committee of the Faculty of Medicine at the University of Chile.

Results

The underlying causes of death recorded in within the group of ICD-10 codes assessed by this study accounted for 21% of the total deaths recorded at regional level. The study sample comprised 515 death certificates (women = 59.2%; men = 40.8%; age = 81 ± 9 years) (Table 1), which registered a total of 1,725 causes of death. The number of causes of death per death certificate ranged from one (5.8%) to eight (0.4%), and was higher among women (60.5%), the 80 to 105 years age group (55%), and on Line A (99.8%). Of the 1,725 causes of death, 97.2% were translated into ICD-10 codes by both coders, 2.3% were coded by only one coder, and 0.6% was not coded. Intercoder agreement was affirmed in a total of 1,715 causes of death.

Underlying cause of death

Table 2 shows the underlying causes of death and intercoder agreement by code classification level. The most common ICD-10 codes used for the 515 underlying causes of death were I21.(0-9), I11.0 and I25.(0-9) (33.8%, 14% and 10.1%, respectively). The lowest level of agreement at the four-character level occurred with ICD-10 codes I50.(0-9) (50%), I42.(0-9) (68.4%) and E11.(0-9) (60%). Level of agreement was also low for codes I09.(1-9) (50%), but frequency was also low.

Level of intercoder agreement with respect to the underlying cause of death at the four-character, three-character and chapter (one-character) level was 80.6%, 86.6% and 94.1%, respectively, and was inversely related to the code classification level (Table 3). Differences were also observed in the number of causes of death per death certificate at the four ($p < 0.01$) and three-character ($p < 0.01$) level. The lowest level of agreement (73.7%) occurred at the four-character level in certificates with four or more causes of death, while the highest level of agreement (96.7%) was observed at the chapter-level in certificates with a single cause of death. In general, level of agreement decreased with increasing code classification level, from 80.6% at the four-character level, to 94.2% at the chapter level. No difference in level of agreement was found across age categories in any of the code classification levels. Cohen's kappa coefficients of coding of underlying cause of death at the four-character, three-character and chapter level were 0.77 (95%CI: 0.75-0.79), 0.84 (95%CI: 0.82-0.86) and 0.76 (95%CI: 0.73-0.79), respectively.

Total causes of death

Level of intercoder agreement for the total causes of death coded by at least one of the coders at the four-character, three-character and chapter level was 76.4%, 81.7% and 86%, respectively (Table 4), with differences depending in the line of the death certificate ($p < 0.01$). The lowest level of agreement across all code classification levels was in Line A (53.3%, 63.1% and 67.25%, respectively) and the highest level of agreement was in Part II of the death certificate (89.3%, 91.4% and 95.9%, respectively). Cohen's kappa coefficients of the nosological coding of the total causes of death at the four-character, three-character and chapter level were 0.75 (95%CI: 0.75-0.77), 0.81 (95%CI: 0.80-0.82) and 0.76 (95%CI: 0.74-0.79), respectively.

A pattern of intercoder differences was observed for three pairs of codes on Line A: codes J96.9 vs. R06.8 were assigned the term "respiratory failure" (3.1%); codes I50.0 vs. I50.9 were used for "decompensated heart failure" (9.7%); and R99 vs. I46.9 were assigned the term "cardiac arrest" (26.6%). No pairs of codes with noteworthy frequency were found for the remaining lines.

Discussion

The results show that the level of agreement in cause of death coding between the Chilean coder and the coding instructor from the United States

Table 1

Number of deaths recorded in Valparaíso and Metropolitan regions *, Chile, during 2008 and demographic characteristics of the study sample.

	Valparaíso and Metropolitan regions		Study sample					
	Total (n ** = 31,112)		Total (n *** = 515)		Men (n *** = 210)		Women (n *** = 305)	
	n	% #	n	% #	n	% ##	n	% ##
Age (years)								
60-69	6,375	20.5	67	13.0	37	55.2	30	44.8
70-79	9,791	31.5	164	31.8	79	48.2	85	51.8
80 and over	14,941	48.0	284	55.2	94	33.1	190	66.9
Marital status								
Single	8,067	25.9	155	30.1	43	27.7	112	72.3
Married	21,550	69.3	334	64.9	158	47.3	176	52.7
Widowed	1,451	4.7	25	4.8	8	32.0	17	68.0
Unknown	44	0.1	1	0.2	1	100.0	-	-
Area of residence								
Urban	30,148	96.9	514	99.8	210	40.8	304	59.4
Rural	964	3.1	1	0.2	-	-	1	100.0
Education level								
Tertiary	1,424	4.6	17	3.3	12	70.6	5	29.4
Secondary	10,224	32.9	134	26.0	64	47.8	70	52.2
Primary	17,665	56.8	333	64.7	125	37.5	208	62.5
No education	1,792	5.8	31	6.0	9	29.0	22	71.0
Medical certifier								
Physician	10,096	32.5	143	27.8	58	40.6	85	59.4
Forensic pathologist	674	2.2	18	3.5	8	44.4	10	55.6
Other doctors	20,299	65.2	354	68.7	144	40.7	210	59.3
Witness	43	0.1	-	-	-	-	-	-
Place of death								
Hospital	14,202	45.6	200	38.8	89	44.5	111	55.5
Home	16,910	54.4	315	61.2	121	38.4	194	61.6

Source: Department of Health Statistics and Information, Chilean Ministry of Health (MINSAL).

* Deaths tabulated on the Chilean mortality database. Deaths attributed to external causes and where place of death was categorized as "others" were excluded from the study sample;

** Number of deaths in the Valparaíso and Metropolitan regions;

*** Number of deaths in the study sample;

Column percentage;

Row percentage.

was adequate. With respect to underlying cause of death, intercoder agreement was inversely related to the number of causes of death per certificate and to the code classification level (4-character, 3-character and chapter level), while in the case of total causes of death, agreement varied according to code classification level and line of the death certificate. Cohen's kappa coefficient values that demonstrate strength of agreement were considered "moderate" to "almost perfect" ¹⁴.

Table 5 shows the level of agreement between expert coders and regular coders found by calibration studies of nosological coding based on automated systems ^{10,15,16}. Our findings are similar to those with an agreement level above 80% across all code classification levels. Based on an earlier version of the ICD, Curb et al. ⁵ reported a level of agreement of 97.1% for the group relating to cardiovascular causes of death (ICD-8: 390-458, 746), which is slightly higher than our

Table 2

Underlying causes of death among the study sample and level of intercoder agreement (%) by ICD-10 code classification level. Valparaíso and Metropolitan regions, Chile, 2008.

ICD-10 codes	Number of deaths * (n = 515)		Level of intercoder agreement		
	n	%	4-character level (%)	3-character level (%)	Chapter of ICD-10 (%)
B57.2	1	0.2	100.0	100.0	100.0
E11.(0-9)	28	5.4	60.0	80.0	92.0
I05.(0-9)	3	0.6	100.0	100.0	100.0
I08.(0-9)	4	0.8	100.0	100.0	100.0
I09.(1-9)	2	0.4	50.0	100.0	100.0
I10	32	6.2	-	78.1	87.5
I11.0	72	14.0	83.3	87.5	93.8
I13.(0-2)	6	1.2	80.0	80.0	80.0
I21.(0-9)	174	33.8	92.0	95.4	97.7
I25.(0-9)	52	10.1	76.1	76.1	91.3
I34.(0-9)	3	0.6	100.0	100.0	100.0
I35.(0-9)	6	1.2	75.0	75.0	75.0
I42.(0-9)	20	3.9	68.4	79.0	100.0
I48	24	4.8	-	77.3	81.8
I50.(0-9)	49	9.5	57.1	77.6	98.0
I51.7	2	0.4	100.0	100.0	100.0
J44.9	37	7.2	78.4	89.2	97.3

ICD-10: International Classification of Diseases, 10th revision.

* Number of death certificates randomly sampled from death certificates issued in the Valparaíso and Metropolitan regions in 2008. Deaths attributed to external causes and where place of death was categorized as "others" were excluded from the study sample.

findings. Harteloh et al.⁹ reported lower levels of agreement for the codes ICD-10: I00-I99 relating to cardiovascular diseases at the four-character, three-character and chapter level (74.6%, 78.2% and 91.5%, respectively). These differences in level of agreement could be explained by the higher number of coders assessed by these studies which tends to lead to an increase in intercoder variability: Harteloh et al.⁹ assessed four coders, while Curb et al.⁵ analysed three coders.

Some studies have reported to varying levels of intercoder agreement related to age, number of causes of death per death certificate, and code classification level^{9,10}. Our results show variations only in relation to the last two variables, ranging from 73.7% in death certificates with four or more causes of death, to 86.7% in those with only one cause of death at the four-character level. A study of agreement between an expert coder and regular coder conducted in Taiwan¹⁰, which included deaths of all ages, reported agreement levels at the three-character and two-character level of 80.9% and 83.9%, respectively and showed that level of agreement was inversely re-

lated to the age of the deceased and the number of reported causes of death.

Considering the complexity of nosological coding of cause of death the WHO has developed rules to facilitate and standardize the classification process and make it more informative^{17,18}. However, some critics have pointed out a number of ambiguities and inconsistencies in these rules^{8,16,19,20} and concerns about the quality of the coding process persist^{5,6} given the observed influence of factors such as the coder's skill level and the level and continuity of coder training. The literature also suggests that differences in interpretation of coding rules result in discrepancies in intercoder agreement^{5,7,10}.

Although this investigation assessed coders from different countries, the level of agreement observed by this study was high, indicating comparable interpretation of the ICD-10 rules, in contrast to the findings of the European Community sponsored EURODIAB study¹⁶, which reported considerable differences between the interpretations of ICD rules across the nine countries included in the study. Another study which

Table 3

Underlying cause of death: level of intercoder agreement (%) by ICD-10 code classification level by age and sex of deceased and number of reported causes of death per death certificate. Valparaíso and Metropolitan regions, Chile, 2008 (n = 515).

Variable	ICD-10 code classification level								
	4-character level			3-character level			Chapter ICD-10 (1-character level)		
	n *	% (95%CI)	Kappa (95%CI)	n *	% (95%CI)	Kappa (95%CI)	n *	% (95%CI)	Kappa (95%CI)
Total	415	80.6 (77.2-84.2)	0.77 (0.75-0.79)	446	86.6 (83.7-89.5)	0.84 (0.82-0.86)	485	94.2 (92.2-96.2)	0.76 (0.73-0.79)
Sex									
Male	176	83.8 (78.8-88.8)		188	89.5 (85.4-93.6)		198	94.3 (91.2-97.4)	
Female	239	78.4 (73.8-83)		258	84.6 (80.5-88.7)		287	94.1 (91.5-96.7)	
Age (years)									
60-69	59	88.1 (80.3-95.9)		61	91.0 (84.1-97.9)		65	97.0 (92.9-100.0)	
70-79	135	82.3 (76.5-88.1)		143	87.2 (82.1-92.3)		154	93.7 (90.0-97.4)	
80-105	221	77.8 (73.0-82.1)		242	85.2 (81.1-89.3)		266	94.2 (91.5-96.9)	
Number of causes of death per death certificate									
1	26	86.7 (74.5-98.9)		27	90.0 (79.3-100.0)		29	96.7 (90.3-100.0)	
2-3	229	85.4 (81.2-89.6)		243	90.7 (87.2-94.2)		256	95.5 (93.0-98.0)	
≥ 4	160	73.7 (67.8-79.6)		176	81.1 (75.9-86.3)		200	92.2 (88.6-95.8)	

95%CI: 95% confidence interval; ICD-10: International Classification of Diseases, 10th revision.

Difference by sex: 4-character: p = 0.12; 3-character: p = 0.11; Chapter: p = 0.92.

Difference by age: 4-character: p = 0.1; 3-character: p = 0.4; Chapter: p = 0.6.

Difference by number of causes: 4-character: p < 0.01; 3-character: p < 0.01; Chapter: p = 0.244.

* Number of death certificates with intercoder agreement on underlying cause of death.

assessed nosological coding of underlying cause of death by a member of a research team and local coders showed a level of agreement of 55% at the three-character level¹⁵, which increased to 67% when the analysis was restricted to large groups of diseases. The authors explained this high level of disagreement due to a tendency by local coders to code the immediate cause of death instead of the underlying cause¹⁵. As stated by Winkler et al.³, it could be argued that the ICD-10 has not achieved the desired improvements with respect to the reliability of the coding process. These authors compared the low levels of agreement with the findings of previous research²¹ which report-

ed levels of agreement at the three and four-character level of 67.7% and 61.5%, respectively. Winkler et al.³ reported levels of agreement at the three-character and four-character level of 56% and 46%, respectively, and a Cohen's kappa coefficient of 0.69 (95%CI: 0.63-0.73) at the chapter level. One of the problems reported by this group was the limited information available on the death certificates, which adversely impacts the quality of nosological coding. Another problem concerns the quality of information on the death certificate. Ill-defined causes of death, modes of death or improper causal sequences, for example^{2,8,18}, negatively affect the coding process, since

Table 4

Level of intercoder agreement (%) by ICD-10 code classification level based on the line of the death certificate and total causes of death coded in the study sample *. Valparaíso and Metropolitan regions, Chile, 2008 (n ** = 1,715).

Line of the death certificate	Total causes of death coded by at least one of the coders (n = 1,715) n ***	n **	Level of intercoder agreement					
			4-character level		3-character level		Chapter of ICD-10 (1-character level)	
			% #	Kappa (95%CI)	%	Kappa (95%CI)	%	Kappa (95%CI)
Line A	525	280	53.3	0.50 (0.48-0.52)	63.1	0.59 (0.57-0.61)	67.2	0.46 (0.41-0.51)
Line B	487	420	86.2	0.86 (0.84-0.88)	90.4	0.89 (0.86-0.92)	95.3	0.88 (0.82-0.93)
Line C	292	243	83.2	0.82 (0.79-0.85)	87.0	0.86 (0.83-0.89)	90.8	0.82 (0.76-0.89)
Part II	411	367	89.3	0.87 (0.85-0.89)	91.4	0.90 (0.89-0.93)	95.9	0.95 (0.90-0.99)
Parts I and II	1,715	1,310	76.4	0.75 (0.74-0.77)	81.7	0.81 (0.80-0.82)	86.0	0.76 (0.74-0.79)

95%CI: 95% confidence interval; ICD-10: International Classification of Diseases, 10th revision.

Differences in the line of the death certificate: p-value < 0.01.

* Number of death certificates randomly sampled from death certificates issued in the Valparaíso and Metropolitan regions in 2008. Deaths attributed to external causes and where place of death was categorized as "others" were excluded from the study sample;

** Number of causes of death with intercoder agreement based on the line of the death certificate;

*** Number of causes of death coded in the study sample by at least one of the coders based on the line of the death certificate. Uncoded causes of death were removed from the denominator;

Percentage agreement.

it involves the post hoc interpretation of a sequence of events that characterize the fatal outcome which often involves subjective judgments.

The level intercoder agreement was lower for Line A of the death certificate (53.3%) than for the other two lines and Part II, where level of agreement was over 80%. This difference was attributable to three pairs of discordant codes (J96.9 – R06.8), (I50.0 – I50.9) and (R99 – I46.9). In this respect, although ICD guidelines emphasise that the mechanism of death does not constitute a cause of death²⁰, physicians often record the former on Line A¹⁷. Further, nosological coding relies on the coder's interpretation of the ICD rules to resolve such errors. In the first pair of discordant codes mentioned above (I50.0 – I50.9) disagreement occurs at the fourth-character level. According to Surján et al.¹⁹, a poor level of agreement at the fourth-character level is considered less severe than in the whole code. Since both codes have high-order semantic relations, this discrepancy should not have a major effect on causes of death statistics. Our findings corroborate the results of studies that show that ambiguities at the four-character level are particularly

common in coding deaths due to diabetes mellitus^{9,10,16}. The other two pairs of discordant codes (R99 – I46.9 and J96.9 – R06.8) are not semantically related and therefore affect agreement as well as the quality of mortality data. It should be noted that the examples described here reflect errors in certification of causes of death involving the use of modes of death instead of independent nosological entities.

One of the limitations of this study is that it only evaluated agreement between two coders, as opposed to agreement among several coders, or agreement between repeated coding of the same material, precluding an examination of intra and intercoder variability. A possible constraint to the generalizability of our results was, in contrast with other countries, the lack of an automated coding system, which meant that agreement was assessed based on manual coding. Given the impact of nosological coding on mortality statistics, and therefore on international comparisons, we chose coders from different countries in order to provide for a wider interpretation of the ICD-10 coding rules. To our knowledge, this is the first study of this kind in Chile.

Table 5

Level of intercoder agreement (%) in different studies addressing cause of death coding. Adapted from Harteloh et al. ⁹.

ICD-10 code classification level	Present study, Chile ^{*,**} (n = 515)	Winkler et al. ³ , Germany, 2010 ^{***,#} (n = 126)	Harteloh et al. ⁹ , Netherlands, 2010 ^{###} (n = 3,522)	Bradshaw et al. ¹⁵ , South Africa, 2006 ^{§,§§} (n = 33)	Lu et al. ¹⁰ , Taiwan, 2000 ^{§§§,†} (n = 632)	Curb et al. ⁵ , USA, 1983 ^{††,†††} (n = 446)
	%	%	%	%	%	%
4-character	80.6		74.6			
3-character	86.6	72.8	78.2		85.8	
Chapter(1 digit)	94.2		91.5			
Large group of disease of ICD				76.0 ‡		97.1

ICD: International Classification of Diseases.

* Agreement between two coders from different countries;

** ICD-10 codes included in the study sample: B57.2, E11.0, E11.1, E11.2, E11.7, E11.9, I01, I05.(0-9), I06.(0-9), I07.(0-9), I08.(0-9), I09.1, I09.2, I09.9, I10.X, I11.0, I13.0, I13.2, I20.0, I20.1, I20.8, I20.9, I21.(0-9), I22.(0-8), I23.(0-8), I25.(0-9), I31.0, I31.1, I31.8, I34.(0-9), I35.(0-9), I36.(0-9), I37.(0-9), I42.(0-9), I43.1, I48.X, I50.0, I50.1, I50.9, I51.4, I51.5, I51.7, J44.1, J44.9, R57.0;

*** Agreement between coders from two federal offices from Germany;

ICD-10 codes: I10-I89;

ICD-10 codes: I00-I99;

Agreement between four coders;

§ Agreement between an expert and a regular coder;

§§ ICD-9 codes: 410-414;

§§§ Agreement between an expert coder and coders from Taiwan's Provincial Department of Health Statistics Office;

† ICD-9 codes: 390-398, 410-414, 420-429;

†† Agreement between three coders;

††† ICD-8 (adapted) codes: 390-458+746;

‡ Agreement at the ICD-9 basic tabulation list level.

Furthermore, the fact that this study was restricted to a particular group of ICD-10 codes could also affect the generalizability of our findings. However, it is important to note that the leading cause of death in Chile is cardiovascular disease and, as mentioned in Methods section, the codes included in our study account for almost half of this large group of diseases and around a quarter of the total deaths recorded in Valparaíso and Metropolitan regions in 2008.

Despite these limitations, our findings provide important information that was previously unavailable in Chile and scarce in other Latin American countries. It also raises a number of important new questions concerning the assessment of intra and intercoder agreement among coders in Chile and coders of other Latin American countries, and the assessment of agreement in ICD-10 coding process relating to deceased of all ages. Health policies rely in large part on vital statistics ²². Reliable, valid and timely information is one of the keys to improving the health care and epidemiological research ²³. A greater understanding of the quality

of the nosological coding process contributes to the accurate interpretation and comparison of mortality statistics ⁹.

Cause-of-death coding procedures should be modernized, since system automation has a positive impact on the quality of vital statistics ^{24,25} because it reduces human intervention. However, the efficiency and accuracy of automated systems also depends on the quality of medical death certification, as reflected in the need to manually code between 15% and 20% of all death certificates, which in turn results in variability in the coding process ^{9,24,25}. Thus, improving training in medical certification of cause of death is critical to the integrity of vital statistics.

In conclusion, our results show that the level of agreement in coding causes of death due to cardiovascular disease in a representative sample of death certificates from two regions in Chile was adequate and that level of agreement was inversely related to the number of causes of death per death certificate and ICD-10 code classification level.

Resumen

Este estudio evalúa la concordancia en la codificación de causas de muerte cardiovasculares entre un codificador en Chile y otro en EEUU en una muestra aleatoria estratificada de certificados de defunción de personas \geq 60 años, emitidos el 2008 en las Regiones de Valparaíso y Metropolitana de Chile. Todas las causas de muerte fueron convertidas a códigos CIE-10 en paralelo por ambas codificadoras. La concordancia se analizó con el acuerdo inter-codificador y el coeficiente kappa de Cohen, según nivel de especificación del código CIE-10 para la codificación de la causa básica y para el total de causas de muerte. El acuerdo inter-codificador fue 76,4% para el total de causas de muerte y 80,6% para la causa básica (acuerdo a nivel de cuatro dígitos), con diferencias según nivel de especificación del código CIE, línea del certificado y número de causas de muerte por certificado. El coeficiente kappa de Cohen fue 0,76 (IC95%: 0,68-0,84) para la causa básica y 0,75 (IC95%: 0,74-0,77) para el total de causas de muerte. En conclusión, la codificación de causas de muerte cardiovasculares y el acuerdo inter-codificador en dos regiones de Chile son comparables a una referencia externa y a informes internacionales.

Mortalidad; Causas de Muerte; Codificación Clínica; Reproductibilidad de Resultados

Contributors

C. Antini collaborated in the research conception and design; data analysis and interpretation; statistical analysis; drafting of the manuscript; supervision; critical revision and final approval of the manuscript. D. Rajs collaborated in the research conception and design; data collection; critical revision of the manuscript for important intellectual content. M. T. Muñoz-Quezada participated in the research design; data analysis and interpretation; drafting of the manuscript. B. A. L. Mondaca participated in the statistical analysis; drafting of the manuscript. G. Heiss collaborated research conception and design; data analysis and interpretation; critical revision of the manuscript for important intellectual content.

Acknowledgments

We would like to thank the staff at the Department of Health Statistics and Information of the Chilean Ministry of Health, for their valuable contribution to this research.

References

1. Lu TH, Shau WY, Shih TP, Lee MC, Chou MC, Lin CK. Factors associated with errors in death certificate completion. A national study in Taiwan. *J Clin Epidemiol* 2001; 54:232-8.
2. Sibai AM. Mortality certification and cause-of-death reporting in developing countries. *Bull World Health Organ* 2004; 82:83.
3. Winkler V, Ott JJ, Becher H. Reliability of coding causes of death with ICD-10 in Germany. *Int J Public Health* 2010; 55:43-8.
4. World Health Organization. International statistical classification of disease and related health problems. Tenth revision. 4th Ed. Geneva: World Health Organization; 2010.
5. Curb JD, Babcock C, Pressel S, Tung B, Remington RD, Hawkins CM. Nosological coding of cause of death. *Am J Epidemiol* 1983; 118:122-8.
6. Mackenbach JP, Van Duyn WM, Kelson MC. Certification and coding of two underlying causes of death in The Netherlands and other countries of the European Community. *J Epidemiol Community Health* 1987; 41:156-60.
7. Walker S. Capturing health information: a coding perspective. *Bull World Health Organ* 2006; 84:254-5.
8. Stausberg J, Lehmann N, Kaczmarek D, Stein M. Reliability of diagnoses coding with ICD-10. *Int J Med Inform* 2008; 77:50-7.
9. Harteloh P, de Bruin K, Kardaun J. The reliability of cause-of-death coding in The Netherlands. *Eur J Epidemiol* 2010; 25:531-8.
10. Lu TH, Lee MC, Chou MC. Accuracy of cause-of-death coding in Taiwan: types of miscoding and effects on mortality statistics. *Int J Epidemiol* 2000; 29:336-43.
11. Nunez FM, Icaza NM. Calidad de las estadísticas de mortalidad en Chile, 1997-2003. *Rev Med Chil* 2006; 134:1191-6.
12. Cowie MR, Mosterd A, Wood DA, Deckers JW, Poole-Wilson PA, Sutton GC, et al. The epidemiology of heart failure. *Eur Heart J* 1997; 18:208-25.

13. World Health Organization. Certificación médica de causa de defunción: instrucciones para los médicos sobre el empleo del modelo internacional del certificado médico de causa de defunción. 1980. <http://whqlibdoc.who.int/publications/924356062X.pdf> (accessed on 14/Jan/2014).
14. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33:159-74.
15. Bradshaw D, Groenewald P, Bourne DE, Mahomed H, Nojilana B, Daniels J, et al. Making COD statistics useful for public health at local level in the city of Cape Town. *Bull World Health Organ* 2006; 84:211-7.
16. Jouglé E, Papoz L, Balkau B, Maguin P, Hatton F. Death certificate coding practices related to diabetes in European countries: the 'EURODIAB Sub-area C' Study. *Int J Epidemiol* 1992; 21:343-51.
17. Anderson RN, Rosenberg HM. Disease classification: measuring the effect of the tenth revision of the International Classification of Diseases on cause-of-death data in the United States. *Stat Med* 2003; 22:1551-70.
18. 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.
19. Surjan G. Questions on validity of International Classification of Diseases-coded diagnoses. *Int J Med Inform* 1999; 54:77-95.
20. Lindahl BI, Glatte E, Lahti R, Magnusson G, Mosbech J. The WHO principles for registering causes of death: suggestions for improvement. *J Clin Epidemiol* 1990; 43:467-74.
21. Giersiepen K, Greiser E. Coding of cause of death for mortality statistics: a comparison with results of coding by various statistical offices of West Germany and West Berlin. *Offentl Gesundheitswes* 1989; 51:40-7.
22. Hernandez B, Ramirez-Villalobos D, Romero M, Gomez S, Atkinson C, Lozano R. Assessing quality of medical death certification: concordance between gold standard diagnosis and underlying cause of death in selected Mexican hospitals. *Popul Health Metr* 2011; 9:38.
23. 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.
24. Lu TH, Tsau SM, Wu TC. The Automated Classification of Medical Entities (ACME) system objectively assessed the appropriateness of underlying cause-of-death certification and assignment. *J Clin Epidemiol* 2005; 58:1277-81.
25. Lu TH. Using ACME (Automatic Classification of Medical Entry) software to monitor and improve the quality of cause of death statistics. *J Epidemiol Community Health* 2003; 57:470-1.

Submitted on 29/Jun/2014

Final version resubmitted on 17/Nov/2014

Approved on 21/Jan/2015