Int J Public Health (2010) 55:43–48
DOI 10.1007/s00038-009-0053-7

ORIGINAL ARTICLE

Reliability of coding causes of death with ICD-10 in Germany

Volker Winkler · Jördis J. Ott · Heiko Becher

Received: 29 May 2008 / Revised: 18 September 2008 / Accepted: 25 March 2009 / Published online: 28 July 2009

Abstract
Objectives The international classification of diseases (ICD) is used to code death worldwide uniformly and comparably. This study investigates the reliability of national ICD-10 coding practice by assessing agreement of two official coding offices in Germany.
Methods Inter-observer agreement for coding of 372 quasi-randomly selected death certificates is measured by percentages of agreement and simple kappa coefficients.
Results Only 209 (56%) deaths were coded with the same 3-digit ICD-10 code. Agreement of the main chapters according to ICD-10 is higher with 78.2% and a kappa statistic of 0.69 (CI 95%, 0.63–0.75). Examples show that the coding rules correctly applied to the information given on the death certificates may lead to different conclusions.
Conclusions Data show good agreement in the marginal distribution, and thus population frequencies of causes of death may be more reliable despite limited agreement between the two coding offices.

Keywords Reliability · Coding causes of death · Mortality statistics · ICD-10

Introduction

Since 1948 the international classification of diseases (ICD) published by the World Health Organization (WHO) is the global standard diagnostic classification for health management and is applied worldwide. It is used to classify diseases recorded on different types of health and vital records including death certificates. The ICD also provides the basis for the compilation of national mortality and morbidity statistics published by the WHO. In addition, it enables the storage and retrieval of information on mortality and morbidity for epidemiological purposes. Data coded with the tenth revision of the ICD (ICD-10) is available for Germany from 1998 onwards. Although an adjusted version (ICD-10-GM) for diagnosis and ambulant and inpatient care for Germany exists since 2004, causes of death are still classified by the most recent WHO version (World Health Organization 2008).

Mortality statistics are of great importance for national and international health policy makers. A clear picture about the burden from specific diseases in terms of incidence, prevalence, and mortality is needed to assign budgets and resources, to design specific programmes for treatment and prevention, and to evaluate prevention programmes.

Mortality statistics are known to suffer from various limitations. Different coding practices, socio-cultural milieus, and individual socio-demographic characteristics such as age are most likely to influence the cause of death assigned on the death certificate. Problems may also arise from systematic bias in diagnosis, incomplete death certificates, and selection of underlying causes of deaths in each country (Parkin and Khlat 1996; WHO 2003; Lindahl et al. 1990).

Furthermore, studies have demonstrated that information on death certificates is error-prone by itself (Surján 1999). In 1989 a study showed a lower inter-observer-agreement for ICD-9 coded death in Germany compared to other countries (Giersiepen and Greiser 1989). The study

V. Winkler · J. J. Ott · H. Becher (✉)
Unit of Epidemiology and Biostatistics,
Institute of Public Health, University of Heidelberg,
Im Neuenheimer Feld 324, 69120 Heidelberg, Germany
e-mail: Heiko.Becher@urz.uni-heidelberg.de

compared inter-observer variation for 1,136 death certificates among seven nosologists from six different State Statistical Offices and one Federal Statistical Office. An agreement of 67.7% was achieved when looking at 3-digit ICD codes, and of 61.5% for 4-digit ICD codes. In contrast, studies in the US have shown a high level of inter-observer agreement of about 90% regardless of using 3- or 4-digit ICD codes by using an automated coding system (Curb et al. 1985). In 2001 a Polish study found the best overall agreement for neoplasms (83.1%); for cardiovascular diseases the overall agreement was 65.5% (Jedrychowski et al. 2001).

This study analyses the reliability of coding causes of death with the tenth revision of the ICD in Germany.

Methods

Inter-observer agreement for causes of death coded with ICD-10 was analysed. The study was embedded into a large cohort study ($n = 34{,}393$) on the mortality of ethnic German migrants from the former Soviet Union (Becher et al. 2007; Ronellenfitsch et al. 2004).

The follow-up assessment of vital status from population registries yielded 2,505 deaths which occurred in North-Rhine Westphalia. Of these, 1,856 occurred after 1997. Since then official coding by the Federal State Statistical Office was done with ICD-10. We requested the Federal State Statistical Office in North-Rhine Westphalia to perform a record linkage for deceased members of the cohort with official mortality records. The linkage was done by date of birth, date of death, sex, and place of death and was successful for all deaths that occurred in that State.

In parallel, we sent additional requests for anonymized copies of death certificates to local health offices. For 372 deaths after 1997 (20% of 1,856) we received copies of death certificates within 4 months. These deaths represent the data base for the present study. The selection from all deaths can thus be considered as random with respect to cause of death and validity of coding.

For this reliability analysis, the death certificates were additionally coded by a professional coding office of another Federal State in Germany (in the following: Offices A and B). All deaths occurred between 1998 and 2006.

Causes of death were grouped into broad categories (main chapters) according to the WHO classification (see Table 1). The agreement within the two biggest categories (neoplasms and diseases of the circulatory system) was then further assessed with subgroups.

To assess level of agreement of the two coding offices, percentages of agreement and simple kappa coefficients were calculated for broad cause of death groups and for some subgroups (Cohen 1960). To analyse a change in agreement over time, broad cause of death groups were divided into two groups by year of death: 1998–2003 and 2004–2006.

For each kappa coefficient the 95% confidence interval was calculated. All calculations were performed using SAS Version 9.1 (SAS 2004).

Table 1 Causes of death groups for analysis

ICD-10	Description
A00-B99	Certain infectious and parasitic diseases
C00-D48	Neoplasms
D50-D90	Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism
E00-E90	Endocrine, nutritional and metabolic diseases
F00-F99	Mental and behavioural disorders
G00-G99	Diseases of the nervous system
I00-I99	Diseases of the circulatory system
J00-J99	Diseases of the respiratory system
K00-K93	Diseases of the digestive system
N00-N99	Diseases of the genitourinary system
R00-R99	Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified
V01-Y98	External causes of morbidity and mortality
Neoplasms	
C00-C14	Lip, oral cavity and pharynx
C15-C26	Digestive organs
C30-C39	Respiratory and intrathoracic organs
C40-C41	Bone and articular cartilage
C43-C44	Skin
C45-C49	Mesothelial and soft tissue
C50	Breast
C51-C58	Female genital organs
C60-C63	Male genital organs
C64-C68	Urinary tract
C69-C72	Eye, brain and other parts of central nervous system
C76-C80	Malignant neoplasms of ill-defined, secondary and unspecified sites
C81-C96	Malignant neoplasms, stated or presumed to be primary, of lymphoid, haematopoietic and related tissue
C97	Malignant neoplasms of independent (primary) multiple sites
D37-D48	Neoplasms of uncertain or unknown behaviour
Diseases of the circulatory system	
I10-I15	Hypertensive diseases
I20-I25	Ischaemic heart diseases
I26-I28	Pulmonary heart disease and diseases of pulmonary circulation
I30-I52	Other forms of heart disease
I60-I69	Cerebrovascular diseases
I70-I79	Diseases of arteries, arterioles and capillaries
I80-I89	Diseases of veins, lymphatic vessels and lymph nodes, not elsewhere classified

Table 2 Comparison of codes for broad cause of death groups

		Office A												
	ICD-10	A00-B99	C00-D48	D50-D90	E00-E90	F00-F99	G00-G99	I00-I99	J00-J99	K00-K93	N00-N99	R00-R99	V01-Y98	Total
Office B	A00-B99	**2**	1		2			1	1	1				**8**
	C00-D48		**110**			1		4		1	1			**117**
	D50-D90			**1**										**1**
	E00-E90				**3**			3						**6**
	F00-F99					**0**		4		1			1	**6**
	G00-G99						**1**		1					**2**
	I00-I99	1	11		1			**126**	7	3		5	4	**158**
	J00-J99		2				1	5	**15**			1	1	**25**
	K00-K93	1	2					6		**10**				**19**
	N00-N99		1					1			**2**			**4**
	R00-R99							3				**8**	1	**12**
	V01-Y98		1										**13**	**14**
	Total	**4**	**128**	**1**	**6**	**1**	**2**	**153**	**24**	**16**	**3**	**14**	**20**	**372**

Results

Comparing the results of the two coding offices showed that 209 (56%) out of 372 deaths had the same 3-digit ICD-10 code. However, for 4-digit ICD-10 codes agreement was only 46% (171 cases). The evaluation of broad cause of death groups yielded an agreement of 78.2% and a kappa statistic of 0.69 (CI 95%, 0.63–0.75) which can be considered as a good agreement (Altman 1991) (see Table 2). Most non-agreeing classification occurred in both directions, except for neoplasms (ICD-10: C00-D48). Office A classified cancer more often as diseases of the circulatory (ICD-10: I00-I99), respiratory (ICD-10: J00-J99) or digestive (ICD-10: K00-K93) system. There were only few non-agreeing classifications in the category of ill-defined causes of death (ICD-10: R00-R99).

The comparison of the kappa statistics over time showed similar results. From 1998 to 2003 there were 156 deaths with an agreement of 52.6% and a kappa of 0.68 (CI 95%, 0.59–0.77) and for 2004–2006 there were 215 deaths with an agreement of 59.1% and a kappa of 0.70 (CI 95%, 0.63–0.78).

To pinpoint the non-agreeing classification found in the study two detailed examples are provided.

Example (1)

Information on death certificate (handwritten, readable):

- Immediate cause of death: hypostatic pneumonia (Stauungspneumonie)
- As a consequence of:
  (a) stroke in the right hemisphere (rechtshirniger Schlaganfall)
  (b) cardiac insufficiency (Herzinsuffizienz)
- Underlying disease: ischaemic cardiomyopathy (ischämische Kardiomyopathie)

ICD-10 codes given:

Office A: 'I25.5' (ischaemic cardiomyopathy)

Office B: 'I64' (stroke, not specified as haemorrhage or infarction).

Example (2)

Information on death certificate (handwritten, readable):

- Immediate cause of death: acute respiratory insufficiency (acute respiratorische Insuffizienz)
- As a consequence of: global, not compensable cardiac insufficiency (global dekompensierte Herzinsuffizienz)
- Underlying disease: chronic obstructive pulmonary disease (COPD), metastasizing prostate cancer (metastasiertes Prostata-Ca)

ICD-10 codes given:

Office A: 'C61' (malignant neoplasm of prostate)

Office B: 'J44.9' (chronic obstructive pulmonary disease, unspecified).

These two examples only give an impression of a much wider range of inconsistencies.

Table 3 shows the group of malignant neoplasms in more detail: 135 cases were coded as cancer at least by one office (sum of numbers in the box), and of these 110 as cancer (not necessarily the same cancer site) by both offices (numbers in italics). This results in a moderate value for the conditional kappa of 0.51 (CI 95%, 0.42–0.60). Here, 14 (12.7%) of these causes were coded as different cancers by 3-digit code yielding an agreement of 87.3%. Nine of the deaths coded as neoplasm of digestive (ICD-10: C15-C26) and of respiratory and intrathoracic organs (ICD-10: C30-C39) by Office A were classified as neoplasms of ill-defined (ICD-10: C76-C80) or multiple sites (ICD-10: C97) or of uncertain behaviour (ICD-10: D37-D48) by Office B.

Table 4 demonstrates the agreement for deaths due to diseases of the circulatory system: 185 cases were categorized into this group at least once (sum of numbers in the box) and 126 out of them were coded as being part of this

Table 3 Comparison of codes for cancer sites

Office B \ Office A	C00-C14	C15-C26	C30-C39	C40-C41	C43-C44	C45-C49	C50	C51-C58	C60-C63	C64-C68	C69-C72	C76-C80	C81-C96	C97	D37-D48	Other	Total
C00-C14	***1***		*1*														**2**
C15-C26		***40***										*2*		*2*	*2*	1	**47**
C30-C39			***23***										*1*	*1*	*1*	2	**28**
C40-C41				***1***													**1**
C43-C44					***2***												**2**
C45-C49		*1*				***0***											**1**
C50							***3***										**3**
C51-C58								***7***									**7**
C60-C63									***2***					*1*			**3**
C64-C68										***5***						2	**7**
C69-C72											***3***					1	**4**
C76-C80									*1*			***1***			*1*		**3**
C81-C96													***7***			1	**8**
C97														***0***			**0**
D37-D48															***1***		**1**
Other	1	7	2		1			1	1	1		1	2		**1**	**237**	**255**
Total	**2**	**48**	**26**	**1**	**3**	**0**	**3**	**8**	**4**	**6**	**3**	**4**	**10**	**4**	**6**	**244**	**372**

Table 4 Comparison of codes for diseases of the circulatory system

Office B \ Office A	I10-I15	I20-I25	I26-I28	I30-I52	I60-I69	I70-I79	I80-I89	Other	Total
I10-I15	***3***	*1*	*1*	*2*				1	**8**
I20-I25	*2*	***39***			*1*	*2*		6	**50**
I26-I28			***1***				*1*	3	**5**
I30-I52	*2*	*6*	*1*	***21***	*2*		*1*	17	**50**
I60-I69		*3*		*6*	***21***	*1*		4	**35**
I70-I79				*1*	*3*	***3***		**1**	**8**
I80-I89					*1*		***1***		**2**
Other	1	10		8	5	2	**1**	**187**	**214**
Total	**8**	**59**	**3**	**38**	**33**	**8**	**4**	**219**	**372**

group by both offices (numbers in italics). A comparison of these 126 deaths shows an agreement of the 3-digit ICD-10 code in 72.8% and yields a conditional kappa coefficient of 0.60 (CI 95%, 0.50–0.71) which indicates a moderate agreement. The non-agreeing classification was almost evenly distributed in both directions, i.e. the total number of cases was similar in both offices (158 and 153, respectively).

Discussion

Overall, the agreement observed in this study was lower than in the study of Giersiepen and colleagues comparing ICD-9 codes (Giersiepen and Greiser 1989). In 1995 another study published by the German Ministry of Health showed similar results to Giersiepen and colleagues (Jahn et al. 1995). This indicates that ICD-10 does not provide an improvement with regard to reliability of cause of death statistics. Since there was no improvement over time, the low agreement cannot be explained by initial difficulties with the ICD-10. A study by Stausberg and colleagues has also considered aspects of ICD coding. They compared reliability of ICD-10 diagnoses coding from different types of medical records by different types of specialists and demonstrated that the refinement of ICD-10 created more complex guidelines for coding which lead to significant uncertainties (Stausberg et al. 2008). However, these aspects were different to the topic considered in this paper.

Our results are similar to the study by Jedrychowski et al. showing an agreement of 83.1% for neoplasms and 65.5% for cardiovascular diseases (Jedrychowski et al.

2001). A higher percentage agreement within the subgroup of neoplasms compared to cardiovascular diseases was already shown for ICD-9 (Giersiepen and Greiser 1989; Jahn et al. 1995). On the other hand, these percentages do not give the full picture since the agreement within a disease subgroup may be less relevant than the agreement between different subgroups.

The subgroups of neoplasms indicate differential misclassification. However, this cannot be answered completely due to the limited number of deaths. Furthermore, the non-agreeing classifications observed in this subgroup are comprehensible. One of the coding offices classified deaths into the not clearly defined categories (ICD-10: C76-C80; C97; D37-D48), which is easily explained by limited information given on the death certificates.

In Great Britain a study showed that for different neoplasms changes in coding practice have a direct impact on mortality statistics (Goldacre et al. 2004). This has also been observed for ischemic heart diseases in Germany (Schubert 2001).

One limitation of this study is the relatively small number of death certificates, which makes it impossible to evaluate other than the two biggest causes of death groups in more detail. The kappa coefficient itself has drawbacks. It heavily depends on the number of available categories for classification. For example, the kappa coefficient from Table 3 (neoplasms) was lower (0.51) than that from Table 4 (diseases of the circulatory system) (0.60), although its agreement was better (87.3% compared to 72.8%).

The given examples show that different coders apply different coding rules. In example 1 for the selection of the underlying cause of death Office A applied the general principle (GP "When more than one condition is entered on the certificate, the condition entered alone on the lowest used line of Part I should be selected only if it could have given rise to all the conditions entered above it.") and Office B applied selection rule 1 (SR1 "If the General Principle does not apply and there is a reported sequence terminating in the condition first entered on the certificate, select the originating cause of this sequence. If there is more than one sequence terminating in the condition mentioned first, select the originating cause of the first-mentioned sequence."). We cannot decide from the information given whether Office B is correct in not applying the general principle.

In example 2 Office A applied SR1 and SR3 ("If the condition selected by the General Principle or by Rule 1 or Rule 2 is obviously a direct consequence of another reported condition, whether in Part I or Part II, select this primary condition.") and Office B applied only SR1. Based on the information given it is difficult to decide upon the correctness of either result. We think Office A may be correct since metastasizing prostate cancer could have led to chronic obstructive pulmonary disease, however, the full medical history of the patient may tell otherwise. Therefore, the information given on the death certificate, which is probably filled in correctly, does not allow an unambiguous coding.

Better agreement can be reached if additional information about coexistent diseases is mentioned in the death certificates (Jedrychowski et al. 2001). Another study concluded that training, individual characteristics of nosologists and other factors are of great importance when comparing German mortality statistics (Schelhase and Weber 2007). Death certificates from this study confirm these observations. There were many barely filled-in death certificates with insufficient information, which do not allow a clear assignment of the underlying cause of death. In addition, the handwriting was often very poor.

Retrospective epidemiological studies often rely on the cause of death information given on death certificates. If certificates are obtained through record linkage, only the ICD codes are provided and further approval of reliability is not possible.

Therefore, it is crucial to assess how much of results from a study, e.g. given as standard mortality ratio (SMR) or any other common measurement, might be biased. Our data demonstrates that the marginal distribution is fairly comparable and, independent of the initially limited agreement between the two coding offices. This implies that if an allocation of resources is based on the relative frequency of disease groups, no major bias is introduced.

Also, an overall SMR analysis conducted from these data can thus be assumed to be trustworthy. For subgroup analysis, or dose–response assessment, this may yield a bigger problem. Although it can be assumed that coding is independent of exposure (random misclassification) so the true dose–response may be underestimated.

As already indicated by Schelhase and Weber, the introduction of an automated coding system, which is able to code all causes on death certificates and select the underlying cause automatically, could possibly solve some of the problems (Schelhase and Weber 2007). Furthermore, an automated coding system is the prerequisite for a routine multiple-cause coding, which would address problems approached in this article.

Acknowledgments The study was funded by a grant of the German Research Foundation, Graduate College 793: "Epidemiology of communicable and chronic non-communicable diseases and their interrelationships".

References

Altman DG (1991) Practical statistics for medical research. Chapman and Hall/CRC, Boca Raton

Becher H, Kyobutungi C, Laki J, Ott JJ, Razum O, Ronellenfitsch U, Winkler V (2007) Mortalität von Aussiedlern aus der ehemaligen Sowjetunion: Ergebnisse einer Kohortenstudie. Dtsch Arztebl 104:1655–1662

Cohen J (1960) A coefficient of agreement for nominal scales. Educ Psychol Meas 20:37–46

Curb JD, Ford CE, Pressel S, Palmer M, Babcock C, Hawkins CM (1985) Ascertainment of vital status through the National Death Index and the Social Security Administration. Am J Epidemiol 121:754–766

Giersiepen K, Greiser E (1989) 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 51:40–47

Goldacre MJ, Duncan ME, Cook-Mozaffari P, Griffith M (2004) Trends in mortality for cancers, comparing multiple- and underlying-cause rates, in an English population 1979–1999. Br J Cancer 90:1019–1021

Jahn I, Jöckel KH, Bocter N, Müller W (1995) Studie zur Verbesserung der Validität und Reliabilität der amtlichen Todesursachenstatistik. Schriftenreihe des Bundesministeriums für Gesundheit, Band52, Baden-Baden

Jedrychowski W, Mróz E, Wiernikowski A, Flak E (2001) Validity study on the certification and coding of underlying causes of death for the mortality statistic. Prz Epidemiol 55:313–322

Lindahl BIB, Glattre E, Lahti R, Magnusson G, Mosbech J (1990) The WHO principles for registering causes of death: suggestions for improvement. J Clin Epidemiol 43:467–474

Parkin DM, Khlat M (1996) Studies of cancer in migrants: rationale and methodology. Eur J Cancer 32A:761–771

Ronellenfitsch U, Kyobutungi C, Becher H, Razum O (2004) Large-scale population-based epidemiological studies with record linkage can be done in Germany. Eur J Epidemiol 19:1073–1074

SAS Institute Inc (2004) SAS/STAT Software Release 9.1. Cary, NC, SAS Institute Inc

Schelhase T, Weber S (2007) Mortality statistics in Germany. Problems and perspectives. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 50:969–976

Schubert A (2001) Bemerkungen zur Aussagefähigkeit der Todesursachenziffer „ischämische Herzkrankheiten". Gesundheitswesen 63:769–774

Stausberg J, Lehmann N, Kaczmarek D, Stein M (2008) Reliability of diagnoses coding with ICD-10. Int J Med Inform 77:50–57

Surján G (1999) Questions on validity of International Classification of Diseases-coded diagnoses. Int J Med Inform 54:77–95

World Health Organization (2003) ICD-10 2003 version (2nd edn) vol 2—instruction manual. Available via DIALOG. http://www.who.int/entity/classifications/icd/ICD-10_2nd_ed_volume2.pdf. Accessed 11 December 2008

World Health Organization (2008) International statistical classification of diseases and related health problems 10th revision version for 2007. Available via DIALOG. http://www.who.int/classifications/apps/icd/icd10online. Accessed 21 December 2008