Arch Public Health
2005, 63, 127-148

Mortality statistics in Belgium 1980-1997 Quality of coding

by

Aelvoet W[1,3], Molenberghs G[2], van Sprundel M[3]

Abstract

Aim: *To determine to which extent and how coding errors affect the Belgian mortality statistics.*

Methods: *According to rules and notes of the ICD-9 Classification, an automated programme was developed and applied to the Belgian 1980-1997 mortality data. In addition to this, a recoding exercise has been carried out in collaboration with the French WHO Coding Reference Centre on a systematic sample of 566 certificates.*

Main findings: *The number of errors decreased significantly over time (from about 100 ‰ errors to 10 ‰), but the pattern of errors changed in time and in magnitude, according to region of residence and age group (higher error rates in the higher age group, lower error rates*

[1] Federal Public Service Health, Food Chain Safety and Environment, Directorate-General for the Organisation of Health Care Establishments
[2] Universiteit Hasselt, Centrum voor Statistiek, Diepenbeek
[3] Universiteit Antwerpen, Vakgroep Epidemiologie en Sociale Geneeskunde, Antwerpen

Correspondence address: W. Aelvoet, Federal Public Service Health, Food Chain Safety and Environment, Directorate-General for the Organisation of Health Care Establishments, Eurostation II, Victor Hortaplein 40 bus 10, 1060 Brussels, Belgium.
E-mail: willem.aelvoet@health.fgov.be.

in Brussels). The high proportion of certificates, mentioning only one cause of death, prevented to study these phenomena according to type of pathology. Moreover, we found that the real error rate is much higher than that established by the automated programme as shown by the recoding exercise, which determined a four times higher error rate.

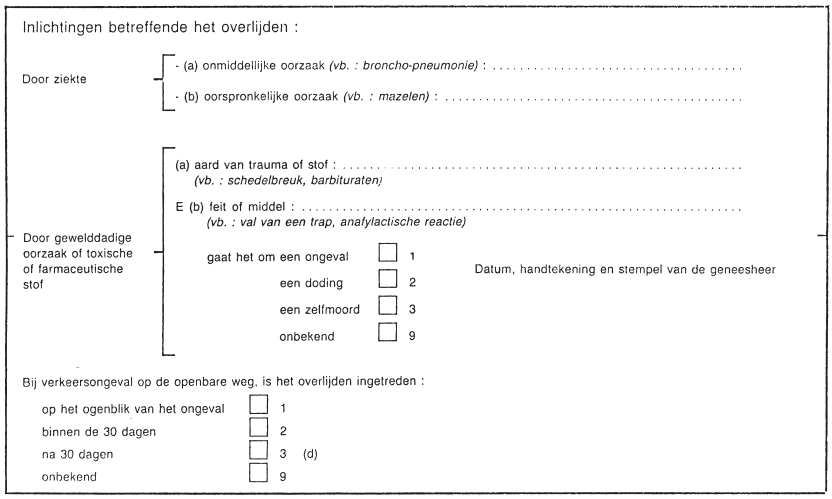
Principal conclusions: *Notwithstanding the improved quality of our mortality data, a continuous effort of quality assurance and of the timely availability of the statistics is mandatory. It aims essentially at the accuracy of the coding, in which comparisons with other registers are essential, and at the allocation of the necessary resources. Due to shortcomings of the past, one has to be cautious by the interpretation of our older mortality data and of time trends.*

Key words: *Coding, Mortality Statistics, Editing, Death Certificates.*

Introduction

Mortality statistics are a major source of information regarding public health (1-7). Between- and within-country differences in patterns of mortality have been studied. They revealed marked variations regarding very important categories of diseases, for instance diseases of the circulatory system, neoplasms and diabetes (6,8-13). Although internationally agreed instructions regarding certification and coding exist, it remains difficult to interpret these variations due to differences in certification and coding practices (14-21). Both of them have been assessed for Belgium (20). In our country additional problems consisted in the long-lasting decentralisation of the coding and the use of a certificate not in conformity with the internationally agreed upon WHO model (22,23). Whereas the latter consists of two parts with five lines to describe the immediate, intermediate, underlying and contributory causes-of-death (CODs) and provides in addition a time sequence of the onset of the different causes, the former consists only of two lines allowing the mentioning of the mere immediate and underlying CODs (Figure 1).

The 'Ninth Revision of the International Classification of Diseases' (ICD-9) contains an important chapter entitled 'Medical Certification and Rules for Classification', which is devoted to the completion and the coding of death certificates (22). The primary objective of public health in the domain of mortality being to prevent the precipitating cause from operating, the main focus of that chapter concerns the identification of the underlying cause-of-death (COD), which is defined by the WHO as



Inlichtingen betreffende het overlijden :

Door ziekte
- (a) onmiddellijke oorzaak *(vb. : broncho-pneumonie)* : .....
- (b) oorspronkelijke oorzaak *(vb. : mazelen)* : .....

Door gewelddadige oorzaak of toxische of farmaceutische stof
(a) aard van trauma of stof : .....
(vb. : schedelbreuk, barbituraten)
E (b) feit of middel : .....
(vb. : val van een trap, anafylactische reactie)
gaat het om een ongeval ☐ 1
een doding ☐ 2
een zelfmoord ☐ 3
onbekend ☐ 9

Datum, handtekening en stempel van de geneesheer

Hier plooien

Hier plooien

Bij verkeersongeval op de openbare weg, is het overlijden ingetreden :
op het ogenblik van het ongeval ☐ 1
binnen de 30 dagen ☐ 2
na 30 dagen ☐ 3 (d)
onbekend ☐ 9

CAUSE OF DEATH		Approximate interval between onset and death
I		
Disease or condition directly leading to death *	(*a*) . . . . . due to (or as a consequence of)	. . . . .
Antecedent causes Morbid conditions, if any, giving rise to the above cause, stating the underlying condition last	(*b*) . . . . . due to (or as a consequence of)	. . . . .
	(*c*)	. . . . .
II		
Other significant conditions contributing to the death, but not related to the disease or condition causing it	. . . . .	. . . . .
	. . . . .	. . . . .

* This does not mean the mode of dying, e.g., heart failure, asthenia, etc. It means the disease, injury, or complication which caused death.

Figure 1. The Belgian Certificate and the WHO model

'(a) the disease or injury which initiated the train of morbid events leading directly to death, or [in case of an unnatural death] (b) the circumstances of the accident or violence which produced the fatal injury' (24). The rules for classification naturally lend themselves to the automation of coding. Thus, at the National Center for Health Statistics of the US, an automated system for coding mortality, called Automated

Classification of Medical Entities (ACME), was developed (25-28). It aimed to introduce a consistent and rapid assignment of the underlying COD as well as to allow for multiple COD statistics. The core of the system consists in a number of 'decision tables' according to which the codes are assigned in a consistent way. These tables became 'de facto' the international coding standard (27,29). Inspired by the ACME experience and adapting it to their situation, several European countries initiated automated coding and in Belgium the first steps towards automated coding were taken as well (28,30).
In Belgium too, the quality of mortality statistics has improved by automated screening of impossible or improbable codes, given age or gender of the deceased. Unfortunately, our national screening programme does not exploit the full possibilities of the coding rules (31). The above-mentioned shortcomings in coding and certification practices made our mortality statistics rather unreliable for the study of health problems such as myocardial infarction, respiratory diseases and diabetes (18,20,32).
In this study we try to determine to which extent coding errors may affect our mortality statistics. To do so we identified coding errors through editing techniques, which test the consistency of the assigned codes. Indeed, to be valid, a code has to belong to the list of ICD codes, has to be compatible with the deceased's gender, has to be eligible to be used as underlying cause of death, according to the appropriate lists, and has to obey to the rules of the classification (28). This validation exercise is carried out for the period of 1980 to 1997 on our national mortality statistics. As this approach may not detect all coding errors, we investigate its discriminating potential through the comparison in a systematic sample of 566 certificates of the codes assigned by the Flemish Community in 1988 and reassigned by the French WHO coding reference centre.

Material and methods

The National Institute of Statistics and the Administration of Health Care of the Ministry of the Flemish Community provided us with the 1980-1997 mortality data for Belgium, for people aged 1 year and older. The data included gender; age groups, expressed as less than 25 years, 25 to 64 years, and 65 years and more; residence; and the underlying and immediate CODs. According to the existing rules the death certificates are to be destroyed once the data are processed. As a consequence we had to rely exclusively on the coded data.

The rules and notes sections of the ICD-9 manual contain instructions specifying how to select the underlying COD between several reported causes (22). Since the national certificate in force at that time consisted

of merely two lines, which were often insufficient to describe the causal chain leading to death as well as possible contributory causes, coders had to select and code an immediate and underlying COD between all by the physician mentioned CODs. In addition, the original death certificates are to be destroyed once the definitive codes have been assigned, so that the study of the quality of coding has to be carried out on the registered immediate and underlying CODs. This provision prevents the comparison of the original death certificates with the recorded CODs, restricting the field of investigation to the assessment of the correct implementation of a limited number of rules, namely those specifying how to choose the underlying cause of death between categories or sequences of causes. To do so we checked the recorded immediate and underlying CODs against the modification rules of 'Senility' and 'Ill-defined conditions'. These rules stipulate that if the selected underlying cause is classifiable to 780-799, codes belonging to the chapter entitled 'Symptoms, signs, and ill-defined conditions', and a condition classifiable among codes other than 780-799 is reported, one has to reselect the underlying cause as if the senility or the ill-defined conditions had not been reported. For instance if the code of senility is recorded as underlying cause and the code of heart failure as immediate cause it is clear that the above mentioned rule of 'Senility' has not been applied. Similarly, we checked the application of the modification rule on linkage and the thereby referred 'Notes for use in underlying cause mortality coding' as well as the 'Notes for interpretation of entries of causes of death'. These are intended to help the coder to apply the rules, especially in situations of less consistent or less complete certification. An example of this type of misclassification would be a code of atherosclerosis as underlying cause and a code of acute myocardial infarction as immediate cause.

The ICD-9 classification consists of 17 chapters and a supplementary classification for external causes. They are broad anatomical or etiologic categories encompassing the diseases. Most of these chapters are further subdivided into subchapters or blocks. According to ICD-9, the CODs are coded in three or four digits, from which the first two digits determine the chapter and subchapter to which the cause belongs, whereas the third and fourth digit are further increasingly specific descriptions of the cause to be coded.

Having programmed these selected rules and notes, it was possible to compare the published statistics with their corrected version. We rated coding disagreements as follows: 0: complete agreement between coding and re-coding; 1: error on the level of the fourth digit; 2: error on the level of the third digit; 3: error on the level of a subchapter; and, 4: error on the level of a chapter. An example hereof is given in table 1.

TABLE 1.
Rating of coding differences

Original coding			Recoding			Rating of discordance
Immediate cause of death	Description	Underlying cause of death	Description	Underlying cause of death	Description	
198.3	Secondary malignant neoplasm of the brain and spinal cord.	162.0	Malignant neoplasm of trachea	162.0	Malignant neoplasm of trachea	0
036.0	Meningococcal meningitis	036.2	Meningococcaemia	036.0	Meningococcal meningitis	1
410	Acute myocardial infarction	413	Angina pectoris	410	Acute myocardial infarction	2
304.9	Drug addiction	292.9	Drug psychosis	304.9	Drug addiction	3
571.2	Cirrhosis of liver	303	Alcohol dependence syndrome	571.2	Cirrhosis of liver	4

Since only one COD is mentioned in a considerable proportion of records, two types of analyses are carried out: a first one on the entirety of the data and a second one on the set of records mentioning more than one COD, which will from now on be referred to respectively as 'whole dataset' and 'subset' in text and tables. The automated programme is almost unable to flag records mentioning a unique COD for non-application of the coding rules – a code of metastasis as underlying COD being an example of the contrary – and underestimates the proportion of errors if a COD mentioned on a certificate has been considered unduly a contributory cause by the coder and thus has not been recorded. Conversely, an analysis of the subset may overestimate the proportion of errors since it seems sensitive to assume that in a considerable part of those records no coding error has been made, e.g. in case of a certificate mentioning only acute myocardial infarction or lung cancer.
Apart from these automated approaches on both the whole dataset and subset, we present a recoding exercise, carried out in collaboration with the French coding Reference Centre on the original certificates, including the causes of death written by the physician, in a systematic sample of 566 certificates. The sampling was carried out as follows: we randomly selected the first certificate from the then available stock of 1998 Flemish certificates and afterwards we drew each next eighth certificate. This approach, allowing to check both the assignment of the codes and the implementation of the whole arsenal of coding rules, makes it possible to appreciate the discriminating ability of the automated programme.
Mainly graphical analyses are carried out, completed by more formal statistical modelling in order to synthetically reflect determinants of coding errors as well as time trends, and to draw conclusions regarding our national and regional statistics. For this purpose, polychotomous logistic regression is used. Herein, due to the small numbers in the age group of less than 25 year, this age group and age group 25 to 64 year have been collapsed. Overdispersion, as a consequence of the correlation within the data, is corrected for by rescaling the standard errors (33-35).

Results

Study population

In table 2 the age and gender distribution of the study population is given for both the whole dataset and for the subset. This distribution is quite similar in both datasets and is characterized by the predominance of the older age groups. We notice an irregular evolution of the proportion of death certificates with a unique COD, globally and according to

TABLE 2.
Age and gender distribution of the study population, Belgium, 1980-1997

	Dataset						Subset					
Age group	**Men**		**Women**		**Total**		**Men**		**Women**		**Total**	
	N	%	N	%	N	%	N	%	N	%	N	%
01-24 year	19,465	2	8,666	1	28,131	1	16,581	3	6,672	1	23,253	2
25-64 year	249,427	26	129,335	14	378,762	20	148,143	25	74,383	13	222,526	19
> 64 year	708,439	72	805,134	85	1,513,573	79	416,720	72	499,648	86	916,368	79
Total	977,331	100	943,135	100	1,920,466	100	581,444	100	580,703	100	1,162,147	100

Source: National Institute of Statistics and Ministry of the Flemish Community

pathology, as well as quite dissimilar proportions of such certificates according to pathology (Figure 2).

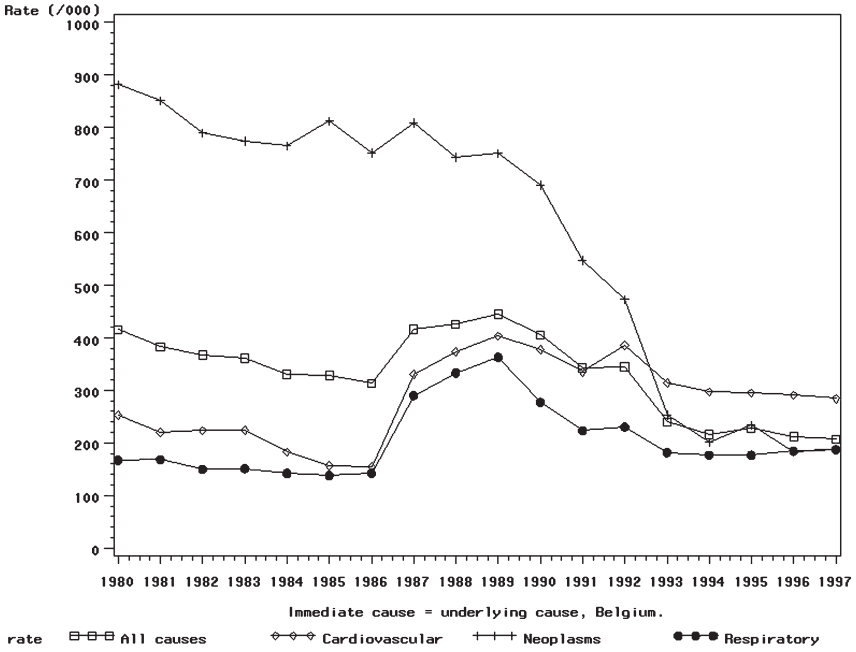


Figure 2. Evolution of the proportion of death certificates with a unique COD, globally and according to pathology, expressed as a chapter of the ICD-9 classification. Belgium, 1980-1997.

Source: National Institute of Statistics and Ministry of the Flemish Community

Evolution over time of coding errors according to residence, gender and age

The number of errors seems to decrease over time but the pattern of errors changes in time and in importance, according to region of residence (Figures 3, 4 and 5), gender and age group (Figure 6) and according to the study population (dataset or subset (Figures 3 and 4)): higher error rates in the subset analyses, higher error rates in the higher age group and dissimilar pattern of error rates according to region of residence but globally decreasing since 1986-7 and a marked improvement between 1992 and 1993.

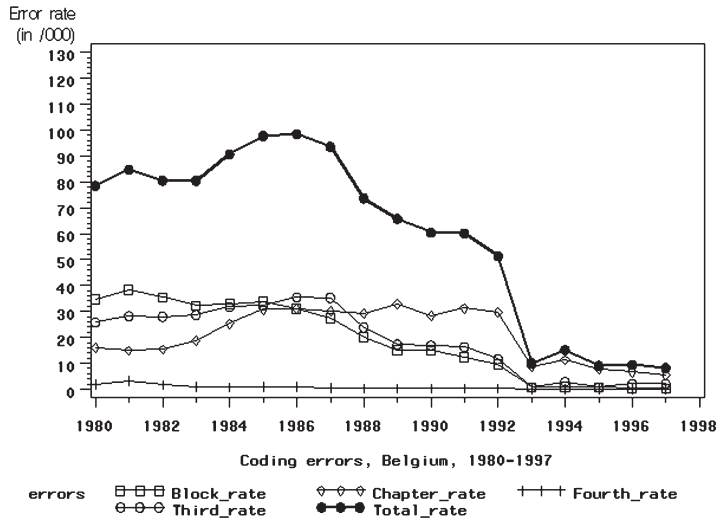


Figure 3. Evolution of coding errors over time, whole dataset. Belgium, 1980-1997.

Chapter_rate: rate of errors at chapter level; Block_rate: rate of errors at subchapter level; Third_rate: rate of errors at the third digit level; Fourth_rate: rate of errors at the fourth digit level; Total_rate: rate of errors
regardless their importance.
Source: National Institute of Statistics and Ministry of the Flemish Community

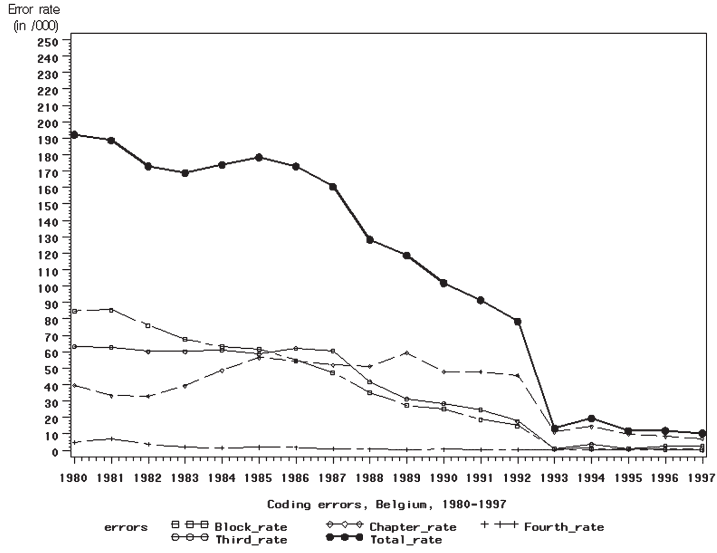


Figure 4. Evolution of coding errors, subset. Belgium, 1980-1997.

Chapter_rate: rate of errors at chapter level; Block_rate: rate of errors at subchapter level; Third_rate: rate of errors at the third digit level; Fourth_rate: rate of errors at the fourth digit level; Total_rate: rate of errors regardless their importance.
Source: National Institute of Statistics and Ministry of the Flemish Community

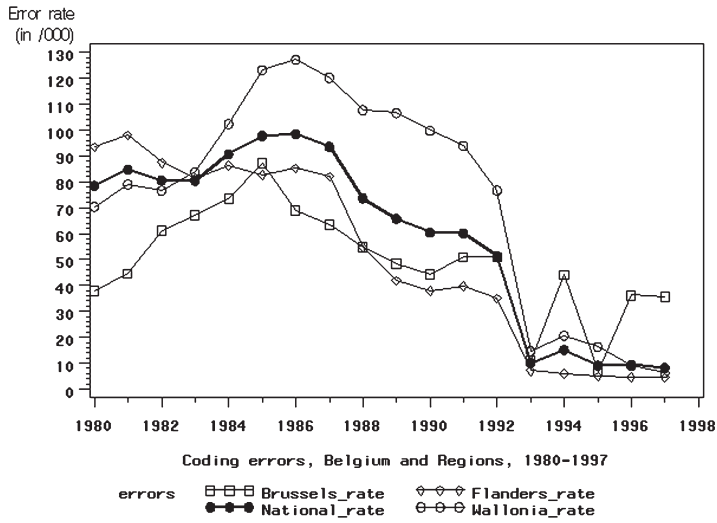


Figure 5. Evolution of coding errors, according to region. Whole dataset. Belgium, 1980-1997.

National_rate: rate of errors regardless their importance, Belgium; Brussels_rate: rate of errors regardless their importance, Brussels; Flanders_rate: rate of errors regardless their importance, Flanders; Wallonia_rate: rate of errors regardless their importance, Wallonia.
Source: National Institute of Statistics and Ministry of the Flemish Community

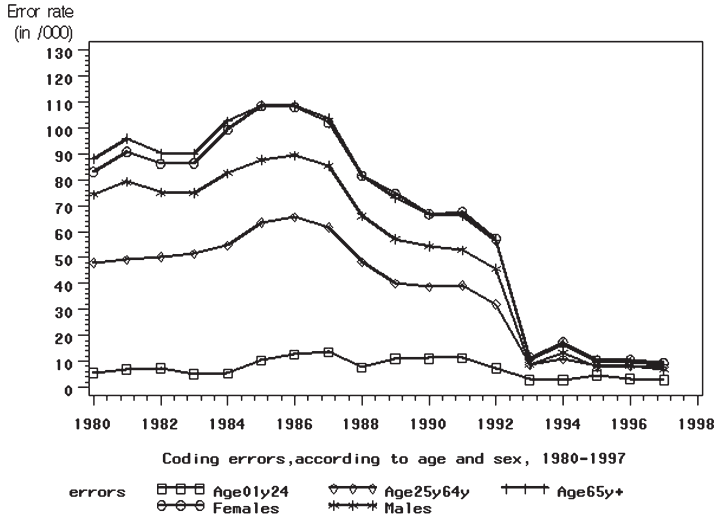


Figure 6. Evolution of coding errors, according to age and gender. Whole dataset. Belgium, 1980-1997.

Age01y24: rate of errors regardless their importance, age group 1-24 y; Age25y64y: rate of errors regardless their importance, agegroup 25-64 y; Age65y+: rate of errors regardless their importance, agegroup 65 y+; Females_rate: rate of errors regardless their importance, females; Males_rate: rate of errors regardless their importance, males.
Source: National Institute of Statistics and Ministry of the Flemish Community

Statistical analysis

Our modelling explored a main effects model as well as several models with one or more interactions. These models were fitted using conventional maximum likelihood. The main effects model did not achieve a good fit (Likelihood Ratio: chi square of 10,213.63; df 2000 and prob. <.0001). From the models with interactions, the model with all pair-wise interactions of the explanatory variables (age group, gender, year, chapter of the classification, region of residence) achieved the nicest fit (Likelihood Ratio: chi-square of 4335.77, df 10,000 and prob. 1,000) and showed extremely significant pair-wise interactions (all with prob. < .0001). These significant interactions denote a dissimilar error pattern according to residence, age group, gender, time and type of pathology. To disentangle this complex pattern we present the main results of the main effects model. From this model the main results for each explanatory variable, adjusted for the other explanatory variables, can be summarized as follows: no remarkable gender divide; less errors in the younger age groups; less important errors in Brussels as compared with Flanders and Wallonia; decrease of errors over time (a categorisation of time in three six-years groups yielded a much better fit than a model with a unique time trend); and a huge amount of coding errors in the cardiovascular and symptoms chapter (data not shown).

Recoding exercise

The age and sex distribution of this sample is quite similar to that of our study population (data not shown). For both the whole dataset and the subset the number of errors detected by the automated programme

TABLE 3.
Distribution of errors according to the automated programme or to the reference centre and to the presence or absence of more than one coded cause of death. Flanders 1988.

Error	Pgm	(%)	PgmNU	(%)	Reference	(%)	RefNU	(%)
0	504	89.0	167	77.7	315	55.7	95	44.2
1	1	0.2	1	0.5	36	6.4	14	6.5
2	14	2.5	2	0.9	74	13.1	27	12.6
3	28	4.9	28	13	66	11.7	45	20.9
4	19	3.4	17	7.9	75	13.3	34	15.8
Total	566	100.0	215	100.0	566	100.2	215	100.0

Pgm: automated correction applied to the dataset; Reference: reference centre recoding applied to the dataset
PgmNU: automated programme applied to the subset; RefNU: error rating by the reference centre applied to the subset
Source: Ministry of the Flemish Community

was far less, with a factor of about five, than that determined by the Reference Centre (Table 3). Several chapters and subchapters of the classification were substantially over- or underestimated. The symptom chapter (ICD-9:780-99) and the blocks of 'other forms of heart disease' (ICD-9:420-9), 'diseases of arteries, arterioles and capillaries' (ICD-9:440-8) are largely overestimated whereas the 'external causes' chapter (ICD-9:E800-E999) and the blocks of 'ischaemic heart disease' (ICD-9: 410-4), 'cerebrovascular disease' (ICD-9: 430-8) and 'chronic obstructive lung disease' (COLD) (ICD-9:490-6) are seriously underestimated. Apart from older age the most prominent factor that facilitated the emergence of errors, appeared to be the difficulty to conveniently complete the then applicable Belgian Death Certificate (23).

Effects of the automated recoding on the Belgian mortality statistics

At chapter level and for both genders, the automated recoding led to a considerable decrease of the importance of the Symptoms chapter mostly in favour of the chapters regarding the 'respiratory and circulatory systems' (Table 4). At subchapter level and for both genders, one sees the disappearance of the subchapter on 'ill defined and unknown causes' within the leading subchapters and the appearance of the subchapter devoted to 'pneumonia and influenza'. 'Ischaemic heart disease' and 'cerebrovascular diseases' increase in both genders as well (Table 5). These over- and underestimations gradually disappear (Figures 7, 8). The recoding influence on both the 'treatable' and 'preventable' components of 'avoidable mortality' is rather small (maximally 2%) and vanishes completely in the course of time.

Discussion

The coding of death certificates improved appreciably during the period 1980-1997. Several factors have contributed to this improvement. External comparisons through the participation in international studies (18,20); the contacts with the French WHO Coding Reference Centre; the collaboration with researchers (23,31); and the comparison of vital statistics data with those of specialised registers such as the cancer register or the Ischaemic Heart Disease Register (MONICA) undoubtedly played an important role (32,36). These efforts were completed by coding supervision; collaboration of the persons in charge of coding of the Communities (23); querying the certifiers about unclear items in the by them completed certificates; and initial and continuing training of the coders and centralisation of coding in Flanders in 1993 as well as similar initiatives in Wallonia. The coding errors of the past, chiefly observed in the older age groups, were largely due to a certificate not in confor-

TABLE 4.
Effect of recoding on mortality statistics, chapter level. Belgium, 1980-1997

		Men				Women			
Chapter	ICD-9	Original	(%)	Recoded	(%)	Original	(%)	Recoded	(%)
Infections	1-139	9323	1	9,901	1	9,424	1	10,428	1
Neoplasms	140-239	284,303	29	284,068	29	207,531	22	207,200	22
Endocrine	240-279	14,386	1	14,762	2	29,887	3	30,800	3
Blood	280-289	2,936	0	2,893	0	3,772	0	3,757	0
Mental	290-319	8,653	1	8,490	1	10,814	1	10,723	1
Nervous	320-389	21,080	2	21,216	2	31,432	3	31,532	3
Circulatory	390-459	360,296	37	369,426	38	413,525	44	427,716	45
Respiratory	460-519	100,692	10	103,048	11	58,351	6	62,268	7
Digestive	520-579	35,582	4	35,217	4	38,912	4	38,588	4
Genitourinary	580-629	14,304	1	14,565	1	17,718	2	18,218	2
Pregnancy	630-676	0	0	0	0	127	0	127	0
Skin	680-709	1,611	0	1,696	0	5,074	1	5,451	1
Musculoskeletal	710-739	1,678	0	1,674	0	4,557	0	4,546	0
Congenital	740-759	907	0	935	0	919	0	942	0
Perinatal	760-779	4	0	7	0	3	0	5	0
Symptoms	780-799	44,818	5	32,681	3	63,572	7	43,314	5
External	E800-E999	76,758	8	76,752	8	47,517	5	47,520	5

TABLE 5.
Effect of recoding on mortality statistics, ten most frequent subchapters. Belgium, 1980-1997

Ranking	Subchapter	ICD-9	N	(%)	Subchapter	ICD-9	N	(%)
	Original coding				Recoding			
1) Men								
1	IHD	410-4	134,976	14	IHD	410-4	145,096	15
2	Neo: respiratory, intrathoracic	160-5	110,506	11	Neo: respiratory, intrathoracic	160-5	110,337	11
3	Other HD	420-9	92,434	9	Other HD	420-9	95,060	10
4	Cerebrovascular	430-8	77,986	8	Cerebrovascular	430-8	86,784	9
5	Neo: disgestive, peritoneum	150-9	70,929	7	Neo: disgestive, peritoneum	150-9	70,994	7
6	COPD	490-6	51,928	5	COPD	490-6	52,008	5
7	Neo: genitourinary	170-89	47,325	5	Neo: genitourinary	170-89	47,155	5
8	Arteries, arterioles	440-8	38,505	4	Pneumonia, Influenza	480-7	28,072	3
9	Ill defined, unknown	797-9	36,481	4	Arteries, arterioles	440-8	27,603	3
10	Neo: other, unspecified	190-9	27,227	3	Neo: other, unspecified	190-9	27,283	3
2) Women								
1	Other HD	420-9	129,156	14	Other HD	420-9	136,302	14
2	Cerebrovascular	430-8	117,702	12	Cerebrovascular	430-8	131,504	14
3	IHD	410-4	101,790	11	IHD	410-4	110,498	12
4	Neo: digestive, peritoneum	150-9	66,845	7	Neo: digestive, peritoneum	150-9	66,789	7
5	Ill defined, unknown	797-9	55,713	6	Neo: bone, connective, skin, breast	170-5	45,840	5
6	Neo: bone, connective, skin, breast	170-5	45,960	5	Ill defined, unkn.	797-9	37,776	4
7	Arteries, arterioles	440-8	41,562	4	Neo: genitourinary	179-89	34,749	4
8	Neo: genitourinary	179-89	34,822	4	Pneumonia, Influenza	480-7	29,129	3
9	Pneumonia, Influenza	480-7	26,108	3	Arteries, arterioles	440-8	28,589	3
10	Neo: other, unspecified	190-9	25,717	3	Neo: other, unspecified	190-9	25,771	3

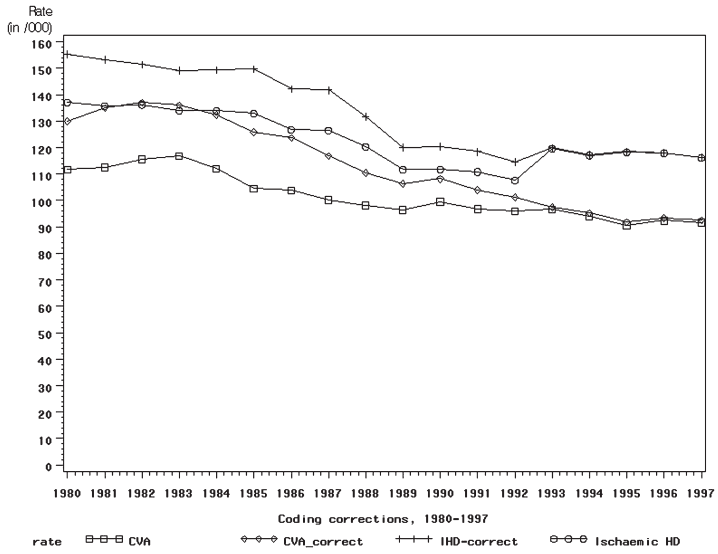


Figure 7. Effects of recoding on COD statistics. Ischemic Heart Diseases (IHD) and Cerebrovascular Diseases (CVA). Whole dataset. Belgium, 1980-1997.

CVA_correct: rate of CVA after recoding
IHD_correct: rate of IHD after recoding
Source: National Institute of Statistics and Ministry of the Flemish Community

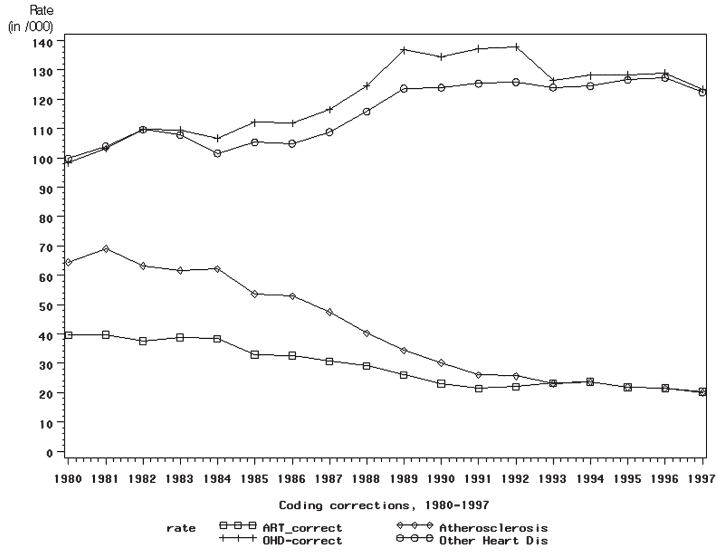


Figure 8. Effects of recoding on COD statistics. Other Heart Diseases (OHD) and atherosclerosis (ART). Whole dataset. Belgium, 1980-1997.

ART_correct: rate of ART after recoding
OHD_correct: rate of OHD after recoding
Source: National Institute of Statistics and Ministry of the Flemish Community

mity with the international model; the inadequate initial training of the coders; their multitask duties, with other priorities set than coding death certificates; and, to a sub-optimal error screening programme (31).

Overlay optimism regarding the effectiveness of our automated programme and hence the real proportion of erroneous coded certificates is unwarranted. Indeed several coding rules require either the conformity of death certificates to the WHO model or an access to the original death certificate and could not be included in our automated programme. Moreover, the presence of an important proportion of certificates that mention only one COD and that, by definition, are out of scrutiny of our automated programme, may constitute another reason of underestimation of the number of errors. The recoding exercise with the French Reference Centre, showing a considerable underestimation of the coding errors of all types, not only supports these hypotheses but also allows a more realistic approach of the real error burden, which was five times higher in 1988 than that estimated by the automated programme.
The important proportion of single COD certificates may not only lead to an underestimation of the number of coding errors but the variability of its magnitude, according to type of pathology, may prevent a firm conclusion regarding an association between type of pathology and presence of coding errors as well. The higher error rates we found regarding the 'cardiovascular and symptoms' chapters of the ICD-9 Classification may be due to this artefact.

The evolution of coding errors over time is dissimilar according to region of residence and age groups. This phenomenon hampers epidemiological and evaluative studies, especially regarding time trends and geographical patterns, as well as the planning activities of policy makers both at the level of the communities and of the country. Fortunately, an important health indicator such as 'avoidable mortality', which focuses on younger age groups and is crucial for developing and evaluating policies of preventive health (4,5,43,44) as well as for the evaluation of the quality of curative health services (45), seems less affected.

As already mentioned in the introductory section, the problem of coding errors is not a Belgian exclusivity (14-21). They are not limited to COD statistics either, as has been shown in several studies about hospital discharge abstracts (40-44). From these studies it appears that the number of errors increases by increasing detail of coding, for instance when one wants a precision up to the fourth digit of the code in stead of the third (42,43). In the European Union (EU), including our country,

several initiatives regarding automated coding have been launched, and are fostered by Eurostat (28,30). Such initiatives may contribute to overcome the distrust regarding COD statistics health care workers often have (45).

Timely availability is another reason for disregarding COD statistics, but this time from a policy makers' viewpoint, and constitutes at the same time an important characteristic of quality of data (46). Indeed, recent data are needed if one wants to stimulate the use of COD statistics as an instrument for policy making, which in turn may facilitate the release of the necessary resources to be affected to the COD statistics. The EU suggests an interval of about two years between the year of collection and the release of the statistics (30). In that respect, the backlog in the processing of the Walloon mortality data – the last available data are from 1997 – is still a major reason for concern.

Conclusions

Despite their limitations, mortality statistics are still a cornerstone of public health, necessitating a sustained effort with regard to their quality, including their timely availability (6,7). Due to shortcomings of the past, one has to be cautious with the interpretation of our older mortality data and of time trends. Indeed, the quality of those data is such that, if the older age groups are to be included in the research, we would advocate not making use of data that are farther back in time than the early nineties and to be very cautious when intending to carry out a study that requires a precision up to the fourth digit.

Although the quality of our mortality data has substantially improved, a continuous effort of quality assurance is mandatory. It consists in a whole of measures aiming at the improvement of the quality of the data provided; the accuracy of the coding, in which comparisons with other registers are essential; and the allocation of the necessary resources. Quality and use of COD statistics are mutually dependent. A better awareness of its usefulness by physicians and policymakers has to be fostered. In this respect, the several initiatives regarding the quality of mortality data that have already been taken by the EU and the new ones that are ongoing may contribute to a revaluation of our COD statistics, provided they are communicated in an appropriate way (30). From these initiatives, the implementation of automated coding in our country deserves a special mention. It will achieve a better interregional and international comparability as well as consistent coding practices over time, prerequisites for a better approach of the health problems in our country.

Acknowledgment

We would like to thank Dr. P. Maguin, Chef du Centre OMS pour la Classification des Maladies, for her invaluable participation into the recoding exercise as well as for her constant encouragements to improve the quality of coding.

Samenvatting

Doelstelling: Bepalen in welke mate en hoe codeerfouten de Belgische doodsoorzakenstatistiek beïnvloeden.

Methode: Een geautomatiseerd programmema, waarin regels en nota's van de ICD-9 classificatie werden geïntegreerd, werd uitgewerkt en toegepast op de Belgische mortaliteitsgegevens van 1980-1997. Daarenboven werd in samenwerking met het Franse WGO Codeer Referentiecentrum een hercoderingsoefening uitgevoerd op een systematische steekproef van 566 certificaten.

Resultaten: Over de jaren heen verminderde het aantal codeerfouten significant (van ongeveer 100‰ tot 10‰). Verder veranderden in de loop der jaren de omvang en het patroon van de codeerfouten in functie van het gewest van de woonplaats en de leeftijdsgroep (hogere foutencijfers in de oudere leeftijdsgroep, lagere cijfers in Brussel). De aanzienlijke proportie certificaten waarin slechts één doodsoorzaak vermeld werd, maakte een analyse in functie van het type pathologie onmogelijk. Ons baserend op de hercoderingsoefening stelden we daarenboven vast dat het geautomatiseerde programmema slechts een vierde van de codeerfouten opspoort.

Conclusies: Niettegenstaande de verbeterde kwaliteit van de overlijdensgegevens is een volgehouden inspanning vereist, zowel op het gebied van de kwaliteitsbewaking als van de tijdige terbeschikkingstelling van de gegevens. Het gaat hierbij voornamelijk over de accuratesse van de codering, waarbij de vergelijking met andere registers essentieel is, en het toekennen van de nodige middelen. Omwille van de tekortkomingen in het verleden dient men voorzichtig te zijn bij het interpreteren van oudere sterftegegevens en van tijdsreeksen.

Résumé

Objectif: Déterminer dans quelle mesure et de quelle façon les erreurs d'encodage influencent les statistiques de décès belges.

Méthode: Un programme informatisé, dans lequel des règles et des notes de la Classification Internationale des Maladies (CIM-9) ont été intégrées, a été élaboré et appliqué aux données de mortalité belges de 1980-1997. En outre et en collaboration avec le Centre de référence français OMS pour la Classification des Maladies, un exercice de réencodage a été effectué sur un échantillon systématique de 566 certificats de décès.

Résultats: Au fil des années, le nombre d'erreurs d'encodage a diminué d'une façon significative (d'environ 100‰ à 10‰). De plus, durant cette période, le nombre et le type d'erreurs ont changé en fonction de la région du lieu de résidence et du groupe d'âge (un taux d'erreurs plus élevé dans le groupe d'âge des 65 ans et plus, des chiffres plus bas

à Bruxelles). La proportion importante de certificats de décès ne mentionnant qu'une cause de décès a rendu impossible une analyse en fonction du type de pathologie. Nous basant sur l'exercice de réencodage, nous estimons que le programme informatisé ne permet le dépistage que d'un quart des erreurs de codage.

Conclusions: Bien que la qualité des données de décès se soit améliorée, un effort soutenu est requis aussi bien dans le domaine du contrôle de la qualité que dans celui de la mise à disposition des données en temps utile. Il s'agit avant tout de l'exactitude du codage, dans laquelle les comparaisons avec d'autres registres sont essentielles, ainsi que de l'attribution des ressources nécessaires. En raison des insuffisances du passé, la prudence est de mise lors de l'interprétation des données de décès anciennes et des séries temporelles.

References

1. Hill AB. The Aim of the Statistical Method. In: Hill AB, ed. A Short Textbook of Medical Statistics. London: Hodder and Stoughton, 1977:1-21.
2. Mackenbach JP. Mortality and medical care: studies of mortality by cause of death in the Netherlands and other European countries. 1-276. 1988. Rotterdam, Erasmus Universiteit.
3. Rosenberg, HM, Chevarley F, Powell-Griner E, Kochanek K, Feinleib M. Causes of Death Among the Elderly.Information From the Death Certificate. Feinleib, M. 35-54. 1991. Hyattsville,MD, NCHS. Vital and Health Statistics. Series 5.No 6. DHHS.
4. Van Oyen H, Tafforeau J, Aelvoet W, Felten G. Overlijden en oorzaakspecifiek overlijden in de Vlaamse Gemeenschap 1987. Brussel: COOV, IHE, 1991.
5. Van Oyen H, Aelvoet W, Tafforeau J. Overlijdenspatroon in het Vlaamse Gewest. Tijdschr voor Geneeskunde 1993;49:821-8.
6. Friedman GD. Primer of Epidemiology. New York: McGraw-Hill Book Company, 1980.
7. Moolgavkar SH, Lee JAH, Stevens RG. Analysis of Vital Statistics Data. In: Rothman KJ, Greenland S, eds. Modern Epidemiology. Philadelphia: Lippincott Williams & Wilkins, 1998:481-97.
8. Holland WW. European Community Atlas of Avoidable Death. Oxford: Oxford University Press, 1988.
9. Humblet P, Lagasse R, Moens G, Van de Voorde H, Wollast E. Atlas de la mortalité évitable en Belgique - Atlas van de vermijdbare sterfte in België (1974-1978). 1-94. 1986. Bruxelles, Ecole de Santé Publique.
10. Humblet PC, Lagasse R, Moens GF, et al. [Avoidable mortality in Belgium]. Soc Sci Med 1987;25:485-93.
11. Mackenbach JP, Moens GFG, Lagasse R. Verschillen tussen Nederland en België in door gezondheidszorg vermijdbare sterfte. Gezondheid en Samenleving 1985;6:265-77.
12. Moens GF, Lagasse R, Humblet P, et al. 'Vermijdbare sterfte': een evaluatie-instrument van de Gezondheidszorg in België? Tijdschr voor Geneeskunde 1986; 42: 921-31.
13. Moens GF, Haenen W, Van de Voorde H. Epidemiological aspects of suicide among the young in selected European countries. J Epidemiol Community Health 1988;42:279-85.
14. Kelson M, Farebrother M. The effect of inaccuracies in death certification and coding practices in the European Economic Community (EEC) on international cancer mortality statistics. Int J Epidemiol 1987;16:411-4.

15. Farebrother MJ, Kelson MC, Heller RF. Death certification of farmer's lung and chronic airway diseases in different countries of the EEC. Br J Dis Chest 1985;79:352-60.
16. Heller RF, Kelson MC. Respiratory disease Mortality in Agricultural workers in eight member countries of the European community. Int J Epidemiol 1982;11:170-4.
17. Evans, J. An assessment of the contribution of differences in coding and certification practices to reported levels of avoidable deaths in the European Community. 1-17. 1988. EC Workshop on Health Services on Avoidable Mortality.
18. Jougla E, Papoz L, Balkau B, et al. Death certificate coding practices related to diabetes in European countries—the 'EURODIAB Subarea C' Study. Int J Epidemiol 1992;21:343-51.
19. Mackenbach JP, Van Duyne 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.
20. Kelson MC, Heller RF. The effect of death certification and coding practices on observed differences in respiratory disease mortality in 8 E.E.C. countries. Rev Epidemiol Sante Publique 1983;31:423-32.
21. Percy C, Dolman A. Comparison of the coding of death certificates related to cancer in seven countries. Public Health Rep 1978;93:335-50.
22. WHO. Medical Certification and Rules for Classification. In: WHO, ed. International Classification of Diseases. 1975 Revision. WHO, 1977:697-741.
23. Aelvoet, W. Vital Statistics in Flanders. An Exploration of Possible Biases due to Coding and Certification Practices. 1-46. 1994. Limburgs Universitair Centrum.
24. WHO. Definitions and Recommendations. In: WHO, ed. International Classification of Diseases. 1975 Revision. WHO, 1977:763-8.
25. Israel RA, Armstrong R. An Alternative Procedure for Classifying and Analizing Mortality Data. 229-42. 1973. LIEGE, International Union for the Scientific Study of Populations.
26. Israel RA, Templeton MC, Evans MC. New approaches to Coding and Analyzing Mortality Data. 20-24. 1972. American Statistical Association.
27. 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.
28. Pavillon G, Coleman M, Johansson LA, Jougla E, Kardaun J. Coding of Causes of Death in European Community. OS/E3/98/COD/3, 1-180. 1998. Luxembourg, Eurostat.
29. Anderson RN, Minino AM, Hoyert DL, Rosenberg HM. Comparability of cause of death between ICD-9 and ICD-10: preliminary estimates. Natl Cancer Inst Monogr 2001;49:1-32.
30. Hooft, P. Verslag van de bijeenkomst van de Eurostat Technical Group on Cuases od Death Statistics. 2004.
31. Aelvoet W, Humblet PC, Lagasse R. Het in routine gebruikmaken van een geïnformatiseerd controleprogramma als middel tot kwaliteitsbewaking van de statistieken van de burgerlijke stand in Vlaanderen. Arch Public Health 1991;49:1-8.
32. De Henauw S, de Smet P, Aelvoet W, et al. Misclassification of coronary heart disease in mortality statistics. Evidence from the WHO-MONICA Ghent-Charleroi Study in Belgium. J Epidemiol Community Health 1998;52:513-9.
33. Collett D. Overdispersion. In: Collett D, ed. Modelling binary data. Boca Raton: Chapman & Hall/CRC, 1991:188-222.
34. Allison PD. Logit Analysis of Longitudinal and Other Clustered Data. In: Allison PD, ed. Logistic Regression Using the SAS System: Theory and Application. Cary, NC: SAS Institute Inc., 1999:179-216.

35. Hosmer DW, Lemeshow S. Special Topics. Logistic Regression Models for the Analysis of Correlated Data. In: Hosmer DW, Lemeshow S, eds. Applied Logistic Regression. New York: John Wiley & Sons, INC., 2000:308-30.
36. Haelterman G, Capet F. Kankerregistratie in Vlaanderen. In: Aelvoet W, Bogaerts K, Capet F, Quataert P K, eds. Gezondheidsindicatoren 1995. Brussel: Ministerie van de Vlaamse Gemeenschap, 1997.
37. Van Oyen H, Tafforeau J, Aelvoet W, Felten G. Overlijden en oorzaakspecifiek overlijden in de Vlaamse Gemeenschap 1988. Brussel: COOV, IHE, 1994.
38. Van Oyen H. Belang van gezondheidgegevens voor een gezondheidsbeleid. In: Nijs H, Denekens J, eds. Preventieve Gezondheidszorg.Verslagboek. Diegem Editorial: Kluwer, 1997:74-82.
39. Newey, C, Nolte, E., McKee, M., and Mossialos, E. Avoidable Mortality in the Enlarged European Union. 3-44. 2004. Brussels, ISS.
40. Crabbe T, Donmall M, Millar T. Validation of the University of Manchester Drug Misuse Database. J Epidemiol Community Health 1999;53:159-64.
41. Colin C, Ecochard R, Delahaye F, et al. Data quality in a DRG-based information system. Int J Qual Health Care 1994;6:275-80.
42. Meyer GS, Krakauer H. Validity of the Department of Defense Standard Inpatient Data Record for quality management and health services research. Mil Med 1998;163:461-5.
43. Fisher ES, Whaley FS, Krushat WM, et al. The accuracy of Medicare's hospital claims data: progress has been made, but problems remain. Am J Public Health 1992; 82:243-8.
44. Cattaruzzi C, Troncon MG, Agostinis L, Garcia Rodriguez LA. Positive predictive value of ICD-9th codes for upper gastrointestinal bleeding and perforation in the Sistema Informativo Sanitario Regional database. J Clin Epidemiol 1999;52:499-502.
45. Westendorp RG. Dwalingen in de methodologie. VI. Doodsoorzaken in perspectief. [Roaming through methodology. VI. Medical perspective on the causes of death]. Ned Tijdschr Geneeskd 1998;142:1950-3.
46. Wyatt JC, Liu JL. Basic concepts in medical informatics. J Epidemiol Community Health 2002;56:808-12.