



**HAL**  
open science

## **Cause-specific mortality time series analysis: a general method to detect and correct for abrupt data production changes**

Grégoire Rey, Albertine Aouba, Gérard Pavillon, Rasmus Hoffmann, Iris Plug, Ragnar Westerling, Eric Jouglu, Johan Mackenbach

### ► **To cite this version:**

Grégoire Rey, Albertine Aouba, Gérard Pavillon, Rasmus Hoffmann, Iris Plug, et al.. Cause-specific mortality time series analysis: a general method to detect and correct for abrupt data production changes. *Population Health Metrics*, 2011, 9 (1), pp.52. 10.1186/1478-7954-9-52 . inserm-00634738

**HAL Id: inserm-00634738**

**<https://www.hal.inserm.fr/inserm-00634738>**

Submitted on 22 Oct 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



RESEARCH

Open Access

# Cause-specific mortality time series analysis: a general method to detect and correct for abrupt data production changes

Grégoire Rey<sup>1\*</sup>, Albertine Aouba<sup>1</sup>, Gérard Pavillon<sup>1</sup>, Rasmus Hoffman<sup>2</sup>, Iris Plug<sup>2</sup>, Ragnar Westerling<sup>3</sup>, Eric Jougl<sup>1</sup> and Johan Mackenbach<sup>2</sup>

## Abstract

**Background:** Monitoring the time course of mortality by cause is a key public health issue. However, several mortality data production changes may affect cause-specific time trends, thus altering the interpretation. This paper proposes a statistical method that detects abrupt changes ("jumps") and estimates correction factors that may be used for further analysis.

**Methods:** The method was applied to a subset of the AMIEHS (Avoidable Mortality in the European Union, toward better Indicators for the Effectiveness of Health Systems) project mortality database and considered for six European countries and 13 selected causes of deaths. For each country and cause of death, an automated jump detection method called Polydect was applied to the log mortality rate time series. The plausibility of a data production change associated with each detected jump was evaluated through literature search or feedback obtained from the national data producers.

For each plausible jump position, the statistical significance of the between-age and between-gender jump amplitude heterogeneity was evaluated by means of a generalized additive regression model, and correction factors were deduced from the results.

**Results:** Forty-nine jumps were detected by the Polydect method from 1970 to 2005. Most of the detected jumps were found to be plausible. The age- and gender-specific amplitudes of the jumps were estimated when they were statistically heterogeneous, and they showed greater by-age heterogeneity than by-gender heterogeneity.

**Conclusion:** The method presented in this paper was successfully applied to a large set of causes of death and countries. The method appears to be an alternative to bridge coding methods when the latter are not systematically implemented because they are time- and resource-consuming.

## Background

The study of cause-specific mortality time series is one of the main sources of information for public health monitoring [1-3]. However, while demonstrative and striking use can be made of such trends when communicating with the general public, many concerns relating to the data production process have to be addressed. More specifically, it is necessary to evaluate, and, if necessary, correct artifacts due to data production

changes that may bias the interpretation of time trends over a study period.

The production processes for mortality databases have been similar in many industrialized countries (particularly in Western Europe) since the end of World War II. When a death occurs, a medical certificate based on the international form recommended by the World Health Organization (WHO) [4] is filled in by a physician. The physician reports the causes of death that directly led or contributed to the death on the death certificate. The death certificate is then forwarded to a national (e.g., France) or regional (e.g., Germany) coding office, where it is coded using the International Classification of

\* Correspondence: gregoire.rey@inserm.fr

<sup>1</sup>INSERM, CépiDc, Le Kremlin-Bicêtre, France

Full list of author information is available at the end of the article

Diseases (ICD). The ICD has been regularly reviewed and improved since the end of the 19<sup>th</sup> century [5]. For each death, an underlying cause of death is selected in compliance with the ICD rules. The underlying cause is the most commonly used in statistical analyses.

Underlying cause coding is a complex process and thus implies potential between-coder coding differences. These differences may produce coding discrepancies over time and space. This is why, in addition to ICD revisions, coding may induce variations in the causes of death by period, region, or country. This situation has resulted in countries using increasingly automated coding systems (ACS).

Variations in mortality trends may also be related to changes in death certification (death certificates, certification habits, diagnoses, etc.). However, the changes are often diffuse and take place over long periods of time, making them harder to take into account. Moreover, the methods of analyzing actual historical variations in mortality by cause of death and taking into account data production process changes are essentially related to changes in the coding process. Of these methods, three main kinds can be distinguished: bridge coding, concordance table and cause recombination, and time series analysis-based methods.

#### **Bridge coding**

The bridge coding method is used when there is a major change in the coding process (ICD version change or switch from a manual to an automatic coding system). The method consists of coding a large set of death certificates twice, applying the rules prevailing before and after the change. The ratios of numbers of death calculated by cause category before and after the change, called “comparability ratios,” generate information for trend analyses and characterization of “jumps” in mortality time series. However, analyses of long-period time series do not necessarily use comparability ratios [1,6].

Bridge coding analyses have been carried out in the United States and England and Wales for each ICD change since the eighth version of the ICD (ICD-8) [7-12]. Bridge coding analysis was also used and generated detailed results for the change from ICD-9 to ICD-10 in some countries (Scotland, Sweden, Italy, Spain, France, and Canada) [13-18]. However, to the authors’ knowledge, most European countries have not implemented bridge coding to assess ICD changes. Comparability ratios are heterogeneous between country, most likely because of variations in intra-group composition of causes of death, reporting practices, and ICD coding interpretation. Therefore, it is unlikely that a comparability ratio for one country can be inferred from the results of other countries.

#### **Concordance table and cause recombination**

The concordance table and cause recombination approach consists of determining the most consistent cause categories, under medical consideration, for two successive ICD revisions. Analysis of mortality using the resulting categories is then theoretically influenced little by coding changes. This approach typically only works well when considering the coding of any particular cause reported on the death certificate. It is often not effective when considering changes in the rules for selecting the underlying cause of death, especially when such rule changes favor the selection of one cause over another. It is often impossible to recombine codes to fully account for these changes.

The method was used on French, Dutch, and Swedish data [19-21]. This approach, complex and time-consuming when it is applied to a single country, is even more difficult to use in the context of an international study [22].

#### **Time series analysis**

The time series analysis method consists of looking for sustainable jumps, evaluating their statistical significance and amplitude, and possibly smoothing the time series by adjusting the data with correction factors. The method is easy to document, even when the volume of data considered is large (many countries, many causes of death, etc.). Furthermore, the method is necessary when the time of the change in the data production process is unknown [23]. To the authors’ knowledge, the detection of jumps in mortality data has rarely been undertaken [22,24], but, in particular for Janssen et al’s work [22], has given rise to fruitful international public health studies [25-28]. However, the methods used in these studies did not take advantage of the recent development of automated jump detection methods in indexed data analysis (by time or other variables) [29-31]. Interest in the automatic jump detection method resides in its ability to avoid the subjectivity of visual detection or a priori selection of jump positions.

The aim of this paper is to propose a complement to a time series analysis method that was previously developed by Janssen et al. [22], allowing detection of sustainable jumps attributable to changes in data production and development of correction factors by age and gender in order to enable subsequent epidemiological analyses. The method is then applied to a wide range of different mortality time series: 13 causes of death for each of six European countries participating in the AMIEHS (Avoidable Mortality in the European Union, toward better Indicators for the Effectiveness of Health Systems) project <http://amiehs.lshtm.ac.uk/>.

## Methods

### General approach

The following step-by-step approach was adopted:

1. Given a list of selected causes of death, the ICD codes to be considered were determined by nosologists based on the correspondence table method, while maintaining the medical consistency of the list of codes for the various ICD revisions.

2. An automated jump detection method was applied to the mortality rate time series for each of the selected causes of death.

3. For documented jumps (e.g., ICD changes), the available comparability ratios were compared to the amplitude of the estimated jumps. For nondocumented jumps, general information feedback was requested from the national data producers.

4. For documented or plausible jump positions, the statistical significance of the between-age and between-gender jump amplitude heterogeneity was evaluated by means of a regression model, and correction factors were deduced from the results.

### Mortality data

The mortality data were derived from the AMIEHS project dataset. Six countries were included in the analysis: Spain, France, the Netherlands, Germany, Sweden, and England and Wales (considered together). In order to simplify presentation, Estonia, which is participating in the AMIEHS project, has not been considered in this paper because Estonia used a specific coding system until 1994. The study period is 1970 to 2005. While causes of death are usually coded with four-digit ICD codes, the AMIEHS dataset only contains three-digit codes for practical reasons. Some precision in the characterization of causes has thus been lost.

Generally, deaths are coded using the same ICD revision in each calendar year. The dates of the ICD revisions used by each European country are presented in table 1.

### Code allocation

For 13 causes of death selected in the AMIEHS project, the method of allocating the ICD-8, ICD-9, and ICD-10

**Table 1 Dates of ICD change and automatic coding system (ACS) implementation for six of the AMIEHS European countries**

Country	ICD-8	ICD-9	ICD-10	ACS
England and Wales	1968	1979	2001	1993
France	1968	1979	2000	2000
Germany	1968	1979	1998	2008
Netherlands	1969	1979	1996	
Spain	1968	1980	1999	
Sweden	1969	1987	1997	1987

codes was as follows:

- When the cause was included in the Eurostat 65 causes shortlist [32], the codes defined by Eurostat were retained.
- For other causes, two nosologists independently selected the optimal three-digit codes. Then, a final choice was made in order to minimize the coding-related jumps in cause of death-specific time series analysis. Table 2 shows the related codes.

### The automatic jump detection method

The Polydict method [33] was applied to yearly log mortality rate time series for each country and cause of death.

Given that mortality analyses are often based on multiplicative assumptions, log-linear generalized models were used. Thus, the time series jump detection method was applied to the log mortality rates.

Let  $O_t$  be the number of deaths during year  $t$ ,  $p_t$  be the number of person-years, and  $L_t = \log\left(\frac{O_t}{p_t}\right)$  be the log mortality rate time series. The occurrence of jumps in the log mortality rate time series may be expressed as follows:

$$\log(E(O_t)) = \log(p_t) + g(t) + \sum_{t' \in S} d_{t'} \cdot 1_{(t > t')}$$

**Table 2 ICD-8, ICD-9 and ICD-10 codes for the 13 selected causes of death**

Cause	ICD-8	ICD-9	ICD-10
Cerebrovascular disease	430-438	430-438	I60-I69
Conditions originating in the perinatal period	760-779	760-779	P00-P96
Congenital heart disease	746	745-746	Q20-Q24
Heart failure	428-429	428-429	I50-I51
Hodgkin's disease	201	201	C81
Hypertension	400-404	401-404	I10-I13
Ischemic heart disease	410-414	410-414	I20-I25
Malignant colorectal neoplasm	153-154	153-154	C18-C21
Malignant neoplasm of cervix uteri	180	180	C53
Malignant neoplasm of testes	186	186	C62
Peptic ulcer	531, 532	531, 532	K25-K26
Renal failure	593, 792	584-586	N17-N19
Rheumatic heart disease	390-398	390-398	I00-I09

In which  $g$  is a continuous function,  $S$  is the set of jump locations, and  $\{d_t, t \in S\}$  are the corresponding jump magnitudes.

In this model,  $g$ ,  $S$ , and  $\{d_t, t \in S\}$  are all assumed unknown.

The method consists of three main steps:

1. A left and right limit of  $E(L_t)$  were estimated for each point  $t$  using two local polynomial smoothers, denoted  $P_l(t)$  and  $P_r(t)$ , fitted on  $[t - h, t)$  and  $(t, t + h]$ , respectively, where  $h$  is the bandwidth for the estimation to be estimated in further steps. If  $t \notin S$ , and the jumps location are distant from at least  $h$ , then, given that  $g$  is continuous, we expect  $E(P_l(t)) = E(P_r(t)) = g(t)$ . Else, if  $t \in S$ , we expect  $E(P_l(t)) = g(t)$  and  $E(P_r(t)) = g(t) + d_t$ .

The noise  $\sigma$  of the  $L_t$  process is estimated as:

$$\hat{\sigma} = \frac{\sum_t \min \left( (L_t - P_l(t))^2, (L_t - P_r(t))^2 \right)}{T - 1}$$

The polynomial kernel of the smoothers could, a priori, be constant, linear, or quadratic, depending on the number of observations and the curvature level of the time series. Since, in the present case, the number of observations was not greater than 40, and the time series was expected to be quite stable, a linear kernel was selected.

2. Considering  $M(t) = P_r(t) - P_l(t)$ , jump points were defined as points where the signal-to-noise ratio  $\frac{|M(t)|}{\hat{\sigma}}$  was higher than a threshold  $C_\alpha$ .

$C_\alpha$  was chosen such that, if  $t$  is not a jump point,  $P \left( \frac{|M(t)|}{\hat{\sigma}} > C_\alpha \right) \leq \alpha$ . The analytic calculation of  $C_\alpha$  is given elsewhere [33]. In the following steps,  $\alpha$  was set to  $10^{-5}$ , a low value, in order to avoid as many as possible false positive jumps.

$$\text{Then, } S = \left\{ t : \frac{|M(t)|}{\hat{\sigma}} > C_\alpha \right\} \text{ and } \left\{ \hat{d}_t = M(t), t \in \hat{S} \right\}$$

were directly estimated. When several jumps were detected in a time range less than the bandwidth, only the jump that maximized  $M(t)$  was retained.

3. The bandwidth  $h$  was estimated by minimizing the Hausdorff distance [29], defined as:

$$d_H(S, \hat{S}; h) = \max \left\{ \sup_{t_1 \in \hat{S}} \inf_{t_2 \in S} |t_1 - t_2|, \inf_{t_1 \in S} \sup_{t_2 \in \hat{S}} |t_1 - t_2| \right\},$$

in which  $d_H(S, \hat{S}; h)$  was calculated through a bootstrap procedure, setting  $B$ , the number of batches used,

equal to 1000. A full description of this method is given elsewhere [33].

For a given jump  $i$ , the multiplicative factor  $MF_i$  between before and after the jump period was calculated as:

$$MF_i = \exp(d_i),$$

in which  $d_i$  is the amplitude of the jump.

### Age- and gender-heterogeneity test

Age categories were defined as the tertile of the cause-specific death counts.

Generally, when considering  $J$  different population groups (age and gender), a generalized additive model (GAM) with an overdispersed Poisson distribution is used [34,35]. The model has the following form:

$$\log(E(O_{t,j})) = \log(p_{t,j}) + g_j(t) + \sum_{t' \in S} d_{t',j} \cdot 1_{(t > t')},$$

in which  $j$  is one of the  $J$  groups,  $g_j$  are continuous functions fitted by a thin plate penalized regression spline,  $S$  is the set of jump locations, and  $\{d_{t,j}, t \notin S\}$  are the corresponding jump magnitudes for group  $j$ .

$S$  is supposed known, and the aim is to test for each  $t \in S$ :

$$H_0: d_{t,1} = \dots = d_{t,J}$$

Backward variable selection was used to suppress, successively, age and gender from the model if their respective effects on the jump amplitude were not statistically significant at the 5% level, using Wald's test.

The MGCV (Multiple smoothing parameter estimation by Generalized Cross Validation) R package was used for this purpose [36].

### Correction factors

Correction factors were calculated for all confirmed jumps.

The correction factors were calculated for use in subsequent analyses, not discussed in this article, with a log-linear model of general form:

$$\log(E(O_t)) = \log(p_t) + c_t + f(X_t),$$

in which  $t$  is the year between  $T_1$  and  $T_2$  (respectively equal to 1970 and 2005 in this study);  $c_t$  is the correction factor and  $f(X_t)$  could be any function of independent variables to be estimated.

The correction factors were set so that the last values of the corrected mortality rates were equal to the exact

mortality rates, i.e.,  $c_{T_2} = 0$ . This choice was based on the supposed superior quality and between-country comparability of the most recent year's data.

The foregoing results in the following definition of the correction factors  $c_t$ :

For  $t \in [T_1, T_2]$ ,

$$\begin{aligned} c_t &= \sum_{t' \in S, t' < t} d_{t'} - \sum_{t' \in S} d_{t'} \\ &= - \sum_{t' \in S, t' \geq t} d_{t'}. \end{aligned}$$

The estimate of  $c_t$  was then directly obtained from the estimates of  $S$  and  $d_t$  detailed earlier.

A corrected version of the log mortality rate was then obtained as:

$$L_t^{\text{cor}} = L_t - c_t$$

## Results

### Jump detection

By applying the jump detection method to all of the time series, a set of jumps was obtained (table 3). Most of the jumps detected were concomitant with a known coding change (ICD updates or change from a manual to an automatic coding system). Some of the jumps (e.g., for heart failure and rheumatic heart disease) were of great amplitude and almost systematically observed in each country. For the former East Germany, most of the changes were concomitant with the reunification of Germany.

The answers from data producers, contacted to determine whether the jump was related to a data production issue, were consistent between countries. Most (excluding the "No answer," 42 out of 44) of the detected jumps were confirmed to be related to a coding change. The three-digit coding constraint was given as an explanation for some jumps (rheumatic heart disease in France, ischemic heart disease in Spain, etc.), especially when countries chose specific codes (as in Spain for malignant colorectal neoplasm). The 1990 and 1991 jumps in East Germany were related to a complete change of coding staff. However, most of the coding changes are not documented by a literature reference.

Given the large proportion of confirmed jumps, we decided to exclude from subsequent treatment the jumps for which we received a negative answer from data producers.

It was possible to compare a few of the multiplicative factors with the comparative ratios generated by bridge coding studies corresponding to ICD-9 to ICD-10 changes (table 4). In particular, the large multiplicative factors (e.g., for rheumatic heart disease) had no related

comparative ratios. Some coding changes were not detected by the jump detection method (Hodgkin's disease in England and Wales and Sweden and renal failure in England and Wales). However, none of the detected jumps were found unrelated to a coding change.

### Corrected mortality rate time series

Considering some of the most clear-cut time series, the profile of the corrected time series is quite different from that of the uncorrected series (Figure 1). It is noteworthy that the corrected curves do not reduce the general trends at the jump positions, which would have been the case if constant rather than linear kernel smoothing was chosen. Rather, they prolong the trends, even if the jump is in the opposite direction of the general trend.

Concerning hypertension in the Netherlands, we observe a trend shift between the periods before and after the corrected jumps. Such trend shift is not taken into account by the current method.

### Estimates of jump amplitudes by age and gender

With regard to the jump amplitude heterogeneity test by age and gender, only 19 out of 47 jumps were not statistically significantly heterogeneous (table 5). Five of the jumps were heterogeneous by gender, 15 by age, and eight by age and gender simultaneously. While the jump amplitudes are of the same order by gender, even when statistically heterogeneous, they are of different orders when considered by age group. This was particularly marked for rheumatic heart disease and heart failure.

## Discussion

The originality of the methodology reported herein mainly resides in its ability to detect jumps automatically using the Polydetect method, without a priori or visual investigation for jump positions. In addition, application of the method to a large dataset is less time-consuming and less human-dependent than any other known method.

Some methodological choices were made, such as the choice of a linear kernel smoother and the choice of the probability  $\alpha$  of detecting fake jumps. Considering a constant kernel smoother or different values of  $\alpha$  slightly affected the final set of detected jumps and only for time series in which the jump amplitudes were of an order comparable to that of the overall noise of the time series. Choosing a low value of  $\alpha$  insured a better accuracy in the jump's amplitude estimation, which is more statistically stable when the jump is of much larger amplitude than the overall noise of the time series. According to the visual inspection of time series graphs and comparable bridge coding results, a jump's

**Table 3 Jumps in the log mortality rate time series for the 13 selected causes of death from 1970 to 2006 identified by the Polydect method**

Country	Underlying cause of death	Year of the jump	Multiplicative factor <sup>1</sup>	Data producer confirmation	Proposed jump cause <sup>2</sup>
England and Wales	Cerebrovascular disease	2001	1.24	Yes	ICD-10
	Conditions originating in the perinatal period	1986	0.09	No answer	No
	Conditions originating in the perinatal period	2001	2.78	Yes	ACS
	Heart failure	1979	3.59	Yes	ICD-9
	Heart failure	1984	0.74	Yes	No
	Heart failure	1993	1.61	Yes	ACS
	Renal failure	1979	2.90	Yes	ICD-9
	Renal failure	1993	2.33	Yes	ACS
	Rheumatic heart disease	1979	0.63	Yes	ICD-9
Rheumatic heart disease	2001	0.77	Yes	ICD-10	
France	Heart failure	1979	1.99	Yes	ICD-9
	Hypertension	1979	0.71	Yes	ICD-9
	Hypertension	1998	1.26	Yes	Death certificate
	Malignant colorectal neoplasm	1979	0.92	Yes	ICD-9
	Malignant colorectal neoplasm	2000	0.94	No	ICD-10
	Peptic ulcer	1979	0.74	Yes	ICD-9
	Peptic ulcer	2000	0.63	Yes	ICD-10
	Rheumatic heart disease	1979	0.36	Yes	ICD-9
Rheumatic heart disease	2000	1.54	Yes	ICD-10	
East Germany	Cerebrovascular disease	1991	1.45	Yes	Reunion
	Conditions originating in the perinatal period	1991	0.46	Yes	Reunion
	Heart failure	1990	1.85	Yes	Reunion
	Heart failure	1991	0.52	Yes	Reunion
	Hypertension	1991	0.51	Yes	Reunion
	Ischemic heart disease	1991	1.43	Yes	Reunion
	Malignant colorectal neoplasm	1991	1.21	Yes	Reunion
	Rheumatic heart disease	1998	1.14	Yes	ICD-10
West Germany	Heart failure	1979	1.65	Yes	IC-D9
	Hodgkin's disease	1994	0.59	No	No
	Ischemic heart disease	1979	0.87	Yes	ICD-9
	Rheumatic heart disease	1979	0.51	Yes	ICD-9
Netherlands	Heart failure	1979	1.54	Yes	ICD-9
	Heart failure	1996	0.76	No answer	ICD-10
	Hypertension	1980	0.65	Yes	No
	Renal failure	1979	0.75	No answer	ICD-9
	Rheumatic heart disease	1992	5.40	No answer	No
	Rheumatic heart disease	1996	0.13	No answer	ICD-10
Spain	Conditions originating in the perinatal period	1975	2.94	Yes	Definition change
	Heart failure	1980	1.41	Yes	ICD-9
	Ischemic heart disease	1980	0.93	Yes	ICD-9
	Malignant colorectal neoplasm	1980	0.78	Yes	ICD-9
	Renal failure	1980	1.73	Yes	ICD-9
Sweden	Heart failure	1987	5.55	Yes	ICD-9
	Malignant neoplasm of testes	1974	1.33	Yes	No
	Malignant neoplasm of testes	1979	0.74	Yes	No

**Table 3 Jumps in the log mortality rate time series for the 13 selected causes of death from 1970 to 2006 identified by the Polydect method (Continued)**

Renal failure	1987	0.69	Yes	ICD-9
Renal failure	1997	1.23	Yes	ICD-10
Rheumatic heart disease	1982	0.08	Yes	Coding change
Rheumatic heart disease	1987	15.33	Yes	ICD-9

<sup>1</sup>: multiplicative factor of the mortality rate between the period before and after the detected jump

<sup>2</sup>: new coding, implementation of which may have caused the jump

amplitude estimates were reliable enough to be used in subsequent analyses.

The codes used in this study to characterize the conditions were not chosen to be used in all contexts. Indeed, they were allocated with the constraint of being comparable between three versions of the ICD and based on three-digit codes. Taking each ICD individually would certainly have led us to select other codes.

The method is designed to detect sustained jumps. Therefore, it is not sensitive to the occurrence of one-year outliers in time series data and it does not necessitate considering them separately, unlike other methods [22].

However, the proposed method is not able to detect and correct for nonabrupt data production changes. For example, if a new death certificate form, impacting certification practice and final coding, slowly spread through the population (as was the case in France between 1997 and 1999), the impact on yearly death counts would occur over several years. But, to the authors' knowledge, no general method is able to correct time-spread data production changes.

When comparable, the multiplicative factors obtained from bridge coding studies and time series methods were similar [11,15-17,37,38].

The purpose of this article is not to challenge bridge coding studies. However, bridge coding studies are not implemented in all countries, and it would be very difficult and costly to do so retrospectively for every data production change. The time series analysis methods proposed herein provide a reliable way of correcting data production changes affecting death count time trends.

Given the indirect manner in which data production changes are identified, the method necessitates feedback from data producers in order to confirm and explain the plausibility of the changes. Without that additional information, the automatic method would blindly correct any detected jumps, some of which may be related to real abrupt and sustained variations in the mortality risk. However, it is not always straightforward for a data producer to obtain a broad overview of past coding process methods in the producer's country. The reasons for the occurrence of some of the oldest jumps may have

**Table 4 Comparative ratios (CR) between bridge coding and multiplicative factors (MF) estimated by the jump detection method for ICD-9 to ICD-10 coding change**

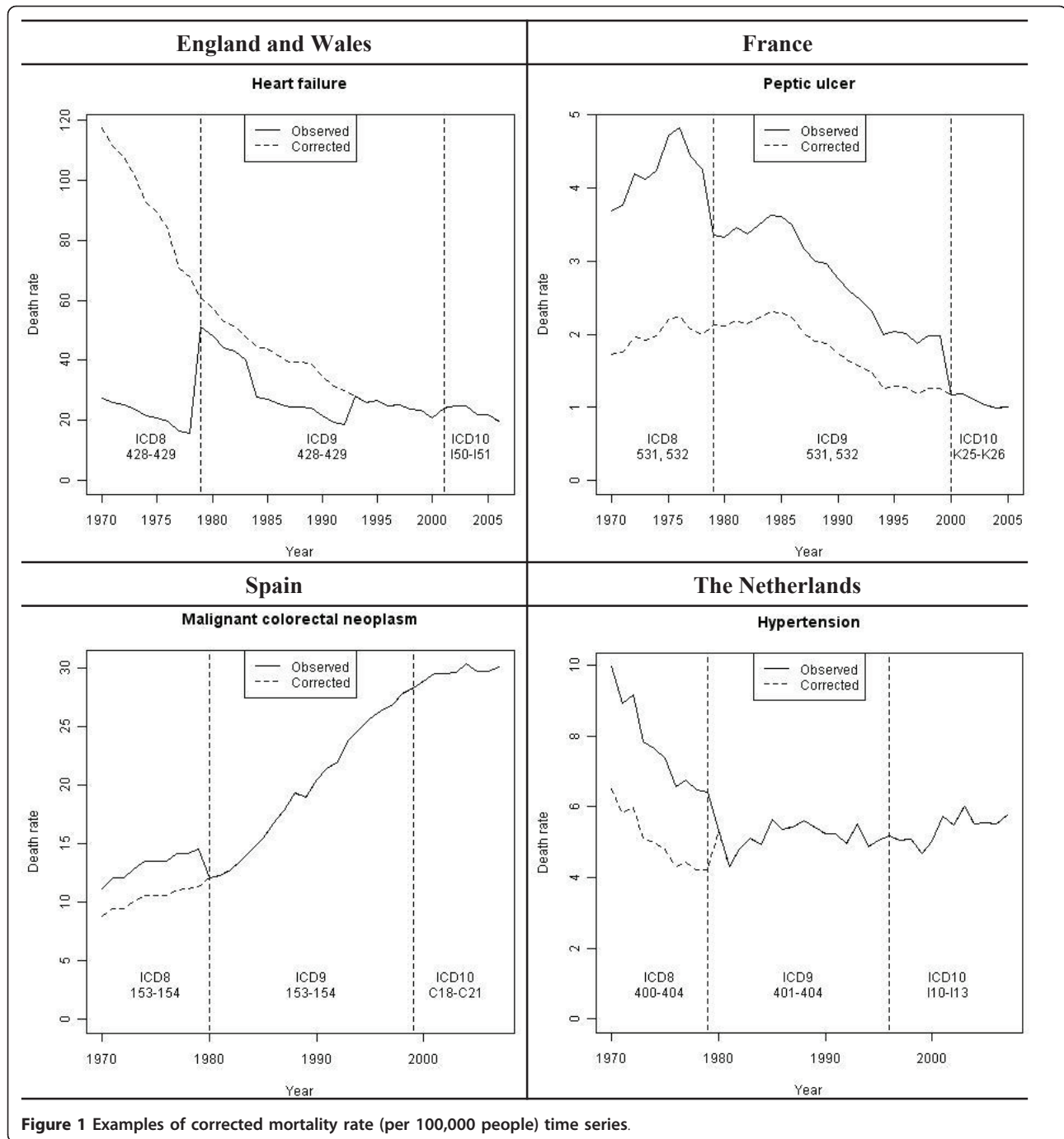
Underlying cause of death	France		England and Wales		Sweden		Spain	
	CR	MF	CR	MF	CR	MF	CR	MF
Cerebrovascular disease	0.97	1.00	1.13	1.13	1.03	1.00	-	1.00
Conditions originating in the perinatal period	1.03	1.00	-	2.24	1.02	1.00	-	1.00
Congenital heart disease	-	1.00	-	1.00	-	1.00	-	1.00
Heart failure	-	1.00	-	1.00	-	1.00	-	1.00
Hodgkin's disease	-	1.00	1.06	1.00	1.13	1.00	-	1.00
Hypertension	-	1.00	-	1.00	-	1.00	-	1.00
Ischemic heart disease	1.03	1.00	1.01	1.00	1.00	1.00	1.00	1.00
Malignant colorectal neoplasm	0.97	0.94	1.01	1.00	0.99	1.00	1.00	1.00
Malignant neoplasm of cervix uteri	0.92	1.00	1.00	1.00	1.00	1.00	-	1.00
Malignant neoplasm of testes	-	1.00	0.99	1.00	-	1.00	-	1.00
Peptic ulcer	-	0.63	1.00	1.00	-	1.00	-	1.00
Renal failure	-	1.00	1.08	1.00	-	1.23	-	1.00
Rheumatic heart disease	-	1.54	-	0.77	-	1.00	-	1.00

CR: Comparative ratios obtained from bridge coding, calculated as the ratio of numbers of death calculated by cause category before and after the ICD change

MF: Multiplicative factor of the mortality rate between the period before and after the ICD change, equal 1.00 when no jump was detected

-: CR not available





been lost. Therefore, the decision to take into account or not any detected jump that is not confirmed by the data producer will depend on the degree of confidence that the jump is not attributable to a production change.

Some jumps are of great amplitude (e.g., rheumatic heart disease). This may be observed when the cause considered is highly likely to be the result of other causes [10,23]. In that case, the death count time trend is very sensitive to changes in coding rules (e.g., ICD-

10 rule 3). However, the absence of high-amplitude jumps is not sufficient to ensure the interpretability of time trends. Time trends for some conditions like hypertension, heart failure, and renal failure have to be interpreted cautiously. Indeed, the approach chosen was to only consider the underlying cause of death, and these specific causes may be selected as underlying, due to lack of additional information about the real underlying cause on the death certificate. In these

**Table 5 Multiplicative factor by age and gender, if statistically heterogeneous, for each detected jump**

Country	Underlying cause of death	Year	Gender <sup>2</sup>	Multiplicative factor <sup>1</sup>		
				Age tertile <sup>3</sup>		
				1st	2nd	3rd
England and Wales	Cerebrovascular disease	2001	Both	1.13		
	Conditions originating in the perinatal period	1986	Both	0.10		
	Heart failure	2001	Both	2.24		
		1979	Male	4.78	3.86	3.34
			Female	4.22	3.41	2.95
		1984	Both	0.60	0.72	0.77
	Renal failure	1993	Both	1.83	1.58	1.37
		1979	Both	2.59		
	Rheumatic heart disease	1993	Both	2.93	1.53	1.50
		1979	Male	0.54	0.50	0.39
		Female	0.82	0.75	0.59	
France	Heart failure	2001	Both	0.87	0.94	0.75
		1979	Both	1.91		
	Hypertension	1979	Male	0.63		
			Female	0.70		
		1998	Both	1.14	1.10	1.25
	Malignant colorectal neoplasm	1979	Both	0.91		
	Peptic ulcer	1979	Both	0.75		
		2000	Both	0.61		
	Rheumatic heart disease	1979	Male	0.39	0.32	0.22
			Female	0.62	0.51	0.35
	2000	Both	1.53			
East Germany	Cerebrovascular disease	1991	Both	1.40		
	Conditions originating in the perinatal period	1991	Both	0.52		
	Heart failure	1990	Both	1.71	1.46	1.27
		1991	Both	0.72		
	Hypertension	1991	Male	0.45	0.51	0.58
			Female	0.52	0.58	0.67
	Ischemic heart disease	1991	Male	1.14	1.21	1.30
			Female	1.22	1.30	1.40
	Malignant colorectal neoplasm	1991	Male	1.14	1.17	1.32
			Female	1.04	1.07	1.20
West Germany	Rheumatic heart disease	1998	Male	0.92		
			Female	1.25		
	Heart failure	1979	Both	1.60		
	Ischemic heart disease	1979	Male	0.98	0.92	0.76
			Female	0.93	0.88	0.72
	Rheumatic heart disease	1979	Male	0.46		
Female			0.69			
Netherlands	Heart failure	1979	Male	2.15	1.33	1.23
			Female	1.86	1.15	1.06
		1996	Both	0.64	0.77	0.82
	Hypertension	1980	Both	0.95	0.81	0.70
	Renal failure	1979	Both	0.41	0.72	0.63
	Rheumatic heart disease	1992	Both	0.89	2.76	5.32
1996		Both	0.42			

**Table 5 Multiplicative factor by age and gender, if statistically heterogeneous, for each detected jump (Continued)**

Spain	Conditions originating in the perinatal period	1975	Male	2.85		
			Female	3.08		
	Heart failure	1980	Both	1.46		
	Ischemic heart disease	1980	Both	0.94		
	Malignant colorectal neoplasm	1980	Male	0.84		
			Female	0.75		
Renal failure	1980	Both	1.66	1.57	1.24	
Sweden	Heart failure	1987	Both	3.57	7.55	7.99
	Malignant neoplasm of testes	1974	Both	1.40		
		1979	Both	0.41	0.72	0.88
	Renal failure	1987	Both	0.65		
		1997	Both	1.59	0.88	1.23
	Rheumatic heart disease	1982	Both	0.12		
1987		Both	5.71	7.20	14.27	

<sup>1</sup>: Multiplicative factor of the mortality rate between the period before and after the detected jump

<sup>2,3</sup>: Multiplicative factor estimated by age and gender, if statistically heterogeneous at the 5% level. Age and gender categorical variables were selected with a backward procedure.

<sup>3</sup>: Age categories were defined by country and cause of death in order to include a third of the number of deaths

cases, mortality time trends could be influenced by other conditions or slowly diffused certification changes. A multiple cause approach considering each cause mentioned on the death certificate could bring very different results.

Large jumps may also be observed when a country uses very specific codes. In this study, for practical reasons, it was decided to use the same codes for all the countries. However, the same general method could have been applied to specific codes for each country.

In any event, time trends for causes with large amplitude jumps, even after correction, are to be interpreted with caution.

For some causes, jump amplitude was markedly heterogeneous by age. This result has already been observed in bridge coding studies [11,16]. This result could be attributed to three factors: first, for some causes, subcause structure is different by age, and each subcause is differentially impacted by production change; second, older age mortality is more frequently associated with multiple pathologies, and the selection of one of these as the underlying cause may change with coding rules; third, in certain cases, the same death certificate may be interpreted differently depending on the age of the deceased, and this difference may also depend on the coding rules used.

## Conclusions

The method presented in this paper was successfully applied to a large set of causes of death and countries. The set of causes considered is heterogeneous in terms of frequency of occurrence (e.g., more than a 100-fold difference between the frequencies of cerebrovascular

disease and malignant neoplasm of the testes) and sensitivity to coding change (no sensitivity for congenital heart disease and high sensitivity for heart failure).

In the future, it would be of interest to investigate the extent to which such a time series approach could be used in a spatial approach to some specific causes. The hypotheses would then be that a large and clear-cut discontinuous change in cause-specific death count, coinciding with a country's border, is attributable to data production discrepancies rather than to real underlying mortality risk variations.

## Acknowledgements

We would like to express our gratitude to Cleo Rooney and Vanessa Fearn with the Office for National Statistics, England and Wales, Lars Age Johansson with the Swedish National Board of Health and Welfare, Torsten Schelhase with the German Federal Statistical Office, Maria Del Rosario Gonzales Garcia with the Spanish National Statistics Institute, Jan Kardaun with the Dutch National Bureau of Statistics, and Gérard Pavillon and Jean Boileau with the French National Institute of Health and Medical Research, for their expertise on their respective databases.

We also are very grateful to Anton Kunst for his role in the study initiation and Andy Mullarky for his skilful assistance in the preparation of the English version of this manuscript.

The study was financially supported by the European Commission (Public Health Program, Agreement number: 2007106).

## Author details

<sup>1</sup>INSERM, CépiDc, Le Kremlin-Bicêtre, France. <sup>2</sup>Erasmus MC, Department of Public Health, Rotterdam, The Netherlands. <sup>3</sup>Department of Public Health and Caring Sciences, Social Medicine, Uppsala University, Uppsala, Sweden.

## Authors' contributions

All authors read and approved the final manuscript. GR participated in the design of the study, conducted the analysis, and drafted the manuscript. AA and GP conducted the code allocation and revised the manuscript. RH and IP provided the dataset and participated in the revision of the results and manuscript. RW revised the manuscript. EJ and JM participated in the design of the study and revised the manuscript.

### Competing interests

The authors declare that they have no competing interests.

Received: 31 January 2011 Accepted: 19 September 2011

Published: 19 September 2011

### References

1. Griffiths C, Brock A: Mortality trends in England and Wales. *Health Statistics Quarterly* 2003, **18**: 5-24.
2. Montserrat A, Devis T, Westlake S, Sumun E, Hervé A, Brückner G, et al: Health statistics - Key data on health 2002 - Data 1970-2001. Luxembourg: Office for Official Publications of the European Communities; 2002.
3. DREES: L'état de santé de la France en 2008 - indicateurs. Paris; 2008.
4. International Classification of Diseases and Related Health Problems, Tenth Revision. Geneva; 1992.
5. History of the development of the ICD: *International Statistical Classification of Diseases and Related Health Problems* Geneva: WHO; 2004, 145-158.
6. Allender S, Scarborough P, O'Flaherty M, Capewell S: Patterns of coronary heart disease mortality over the 20th century in England and Wales: Possible plateaus in the rate of decline. *BMC Public Health* 2008, **8**:148.
7. 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**(9):1551-70.
8. Klebba A, Scott J: Estimates of selected comparability ratios based on dual coding of 1976 death certificates by the eight and ninth revisions of the international classification of diseases. *Monthly Vital Statistics Report* 1980, **28**(11 suppl).
9. Brock A, Griffiths C, Rooney C: The effect of the introduction of ICD-10 on cancer mortality trends in England and Wales. *Health Statistics Quarterly* 2004, **23**.
10. Brock A, Griffiths C, Rooney C: The impact of introducing ICD-10 on analysis of respiratory mortality trends in England and Wales. *Health Statistics Quarterly* 2006, **29**.
11. Griffiths C, Brock A, Rooney C: The impact of introducing ICD-10 on trends in mortality from circulatory diseases in England and Wales. *Health Statistics Quarterly* 2004, **22**:14-20.
12. Griffiths C, Rooney C: Results of the England and Wales ICD-10 comparability study: the effect on main injury and external causes. *Meeting of WHO collaborating centres for the family of international classifications* Cologne, Germany; 2003.
13. Pace M, Bruzzone S, Frova L: Bridge coding study in Italy following ICD-9 to ICD-10 transition: evidences and international comparisons. *EAPS working group on Health, morbidity and mortality. Workshop on Individual, area and group variation in morbidity and mortality* Roma; 2007.
14. The introduction of ICD10 for cause of death coding in scotland. General Register Office for Scotland; 2001.
15. Johansson L: Mortality bridge coding ICD-9/ICD-10: preliminary results from a Statistics Sweden study. *Meeting of heads of WHO collaborating centres for the classification of diseases* Rio de Janeiro, Brazil; 2000.
16. Cano-Serral G, Perez G, Borrell C, Group C: Comparability between ICD-9 and ICD-10 for the leading causes of deaths in Spain. *Rev epidemiol Sante Publique* 2006, **54**:355-365.
17. Pavillon G, Boileau J, Renaud G, Lefèvre H, Jouglà E: Bridge coding ICD9-ICD10 and effects on French mortality data. *WHO family of International Classifications network meeting* Reykjavik, Iceland; 2004.
18. Comparability of ICD-10 and ICD-9 for mortality statistics in Canada. Statistics Canada; 2005.
19. Vallin J, Mesle F: Causes of death in France from 1925 to 1978. Reconstitution of coherent statistical series. *Cah Social Demogr Med* 1987, **27**(4):297-319.
20. Wolleswinkel-van Den Bosch JH, Van Poppel FW, Mackenbach JP: Reclassifying causes of death to study the epidemiological transition in the Netherlands, 1875-1992. *Eur J Popul* 1996, **12**(4):327-61.
21. Klassificering av dödsorsaker i svensk statistik [Classification of causes of death in Swedish statistics]. Statistics Sweden; 1990.
22. Janssen F, Kunst AE: ICD coding changes and discontinuities in trends in cause-specific mortality in six European countries, 1950-99. *Bull World Health Organ* 2004, **82**(12):904-13.
23. Jansson B, Johansson LA, Rosen M, Svanstrom L: National adaptations of the ICD rules for classification—a problem in the evaluation of cause-of-death trends. *J Clin Epidemiol* 1997, **50**(4):367-75.
24. Pearson-Nelson BJ, Raffalovich LE, Bjarnason T: The effects of changes in the World Health Organization's International Classification of Diseases on suicide rates in 71 countries, 1950-1999. *Suicide Life Threat Behav* 2004, **34**(3):328-36.
25. Janssen F, Kunst AE: Cohort patterns in mortality trends among the elderly in seven European countries, 1950-99. *Int J Epidemiol* 2005, **34**(5):1149-59.
26. Amiri M, Kunst AE, Janssen F, Mackenbach JP: Cohort-specific trends in stroke mortality in seven European countries were related to infant mortality rates. *J Clin Epidemiol* 2006, **59**(12):1295-302.
27. Janssen F, Peeters A, Mackenbach JP, Kunst AE: Relation between trends in late middle age mortality and trends in old age mortality—is there evidence for mortality selection? *J Epidemiol Community Health* 2005, **59**(9):775-81.
28. Janssen F, Mackenbach JP, Kunst AE: Trends in old-age mortality in seven European countries, 1950-1999. *J Clin Epidemiol* 2004, **57**(2):203-16.
29. Joo JH, Qiu P: Jump detection in a regression curve and its derivative 2009.
30. Qiu P, Yandell B: A local polynomial jump detection algorithm in nonparametric regression. *Technometrics* 1998, **40**:141-152.
31. Wu WB, Zhao Z: Inference of time trends in time series. *J R Stat Soc B* 2007, **69**(3):391-410.
32. Jouglà E, Pavillon G, Rossollin F, De Smedt M, Bonte J: Improvement of the quality and comparability of causes-of-death statistics inside the European Community. EUROSTAT Task Force on "causes of death statistics". *Rev Epidemiol Sante Publique* 1998, **46**(6):447-56.
33. Zhang B, Su Z, Qiu P: On jump detection in regression curves using local polynomial kernel estimation. *Pak J Stat* 2009.
34. Wood SN: Fast stable direct fitting and smoothness selection for generalized additive models. *J R Stat Soc* 2008, **70**(3):495-518.
35. Hastie T, Tibshirani R: Generalized additive models for medical research. *Stat Methods Med Res* 1995, **4**(3):187-96.
36. Wood SN: Generalized Additive Models: An Introduction with R. Boca Raton, FL: Chapman & Hall/CRC; 2006.
37. OPCS: Mortality statistics: cause 1984. *Series DH2* 1985, **11**.
38. OPCS: Mortality statistics: cause 1993 (revised) and 1994. *Series DH2* 1995, **20**.

doi:10.1186/1478-7954-9-52

Cite this article as: Rey et al.: Cause-specific mortality time series analysis: a general method to detect and correct for abrupt data production changes. *Population Health Metrics* 2011 **9**:52.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
www.biomedcentral.com/submit

