Cardiovascular diseases as causes of death: towards coherence and comparability

Agnieszka Fihel[1,2], France Meslé[1]
[1] Institut National d'études Démographiques, Paris; [2] University of Warsaw

Paper presented at the 2016 European Population Conference, 31st August - 3rd September 2016, Mainz

Abstract

Mortality data by single causes of death allow for the analysis of epidemiological trends and impacts of health policies across countries. Comparing mortality trends in international perspective makes an important methodological challenge for three particular reasons: 1) the cause-specific data lack longitudinal continuity due to changes of classifications of causes of death, 2) each country differs with regard to the way the information on causes of death are collected, coded and registered, 3) causes registered under so-called 'garbage codes' are used with different frequency and in different context across countries. In case of cardiovascular diseases, in particular, data exhibit large dissimilarities in historical and geographical dimension with regard to the scale of mortality and its structure by specific causes of death.

This study presents how the cause-specific data may be corrected and adjusted to international analyses by redistributing deaths registered under garbage codes across selected cardiovascular diseases. Possible impacts of unusual circumstances affecting registering practices, such as introduction of automatic coding, are also investigated. We study mortality: – in the period covered by the 10th revision of ICD, 1994–2013; – across four countries with different coding practices and epidemiologic situations (Czech Republic, Poland, Russia and the United Kingdom), – in adult age groups, – due to four 'garbage code' causes: atherosclerotic cardiovascular or heart disease (ICD-10 code: I25.0, .1), cardiac arrest (I46), heart failure (I50), atherosclerosis (I70). Preliminary results of adjustments prove that mortality due to garbage codes can be to some extent redistributed across well-defined cardiovascular diseases that are related to the respective garbage codes in a pathophysiological and statistical sense. Consequently, for well-defined causes of death we can smooth disruptions in time series resulting from sudden changes in coding practices.

Keywords: Causes of death, garbage codes, cardiovascular mortality

Introduction[1]

In the study concerning all countries of the world for which cause-specific mortality data is available, Mathers et al. (2005) evaluated the quality of such data with the use of three criteria. One of them referred to the frequency of so-called garbage codes (*GCs*) in the national registration: if the share of deaths assigned to garbage codes[2] exceeded 20%, the quality of country-specific information on mortality was univocally assessed as low.

Garbage codes is a term introduced by Murray and Lopez (1996) to designate all causes of death that are not useful in the analyses of public health and mortality, that is:
1) codes that cannot or should not be considered as underlying causes of death (i.e. symptoms, signs and ill-defined conditions),
2) that constitute intermediate causes of death (i.e. heart failure), or
3) that remain unspecified within larger groups of causes (i.e. malignant neoplasm of other and ill-defined sites).

The frequency of *GCs* registration varies tremendously from one country to another and from one revision of *International Statistical Classification of Diseases and Related Health Problems* (ICD) to another. According to the already-quoted study on quality of cause-of-death data, altogether in 105 countries associated with the World Health Organization (WHO) approximately 12% of deaths have been assigned to the *GCs* since 1990 (Mathers et al., 2005). In 20 countries, therein 5 in Europe[3], this percentage was equal to 20% or higher; in one of them, Poland, the highest number of deaths attributed to *GCs* concerned cardiovascular diseases, among which the most frequent remained cardiac arrest (ICD-10 code I46), heart failure (I50) and generalized and unspecified atherosclerosis (I70.9). These causes of death together with atherosclerotic cardiovascular or heart disease (ICD-10 code: I25.0, .1), another *GC*, constituted in Poland more than a half (54%) of all deaths assigned to cardiovascular diseases in 2013. Cardiovascular *GCs* have particularly low informative value and therefore, they remain useless in advanced analyses of epidemiologic situation, assessment of health policy efficiency and international comparisons (Naghavi et al., 2010).

The common and widely accepted practice in using cause-specific mortality data so far has consisted of redistributing age-, sex- and cause-specific death counts (or seldom: death rates) assigned to *GCs* proportionally across all other (well-defined)

[1] This study was financed from the project MODICOD „Le projet AXA Motality Divergence and Causes of Death" and from the project DIMOCHA „Project ANR-12-FRAL-0003-01 DIMOCHA".
[2] The following causes were included: ill defined (ICD-10 codes starting with R), ill-defined cardiovascular diseases (I47.2, I49.0, I46, I50, I51.4, I51.5, I51.6, I51.9, I70.9), neoplasms of unspecified sites (C76, C80, C97).
[3] Cyprus, Greece, Poland, Portugal and San Marino.

causes. This is a solution chosen by the WHO (2015a, 2013) in such studies as *Global Burden of Disease* and *World Health Statistics*. Such redistribution imposes a strong, unrealistic assumption that in case of each disease and each external cause of death the difficulty in recognizing the appropriate diagnose is the same. Also, this practice favors causes of death that are already large and important, even though pathophysiologic or statistical links between those causes and *GCs* may not exist at all. Finally, for some countries the proportional redistribution seems to be inaccurate and inappropriate; for instance, in case of Russia or Ukraine ill-defined causes (constituting a part of *GCs*) are assigned more often instead of cardiovascular diseases (Meslé and Vallin, 2012, 2003), whereas in Poland there is a statistically significant geographical correlation showing that ill-defined categories are assigned instead of selected well-defined causes: acute myocardial infarction, pneumonia, non-insulin-dependent diabetes mellitus (Fihel and Pechholdová, 2016). In this study we want to find a more accurate technique identifying selected well-defined categories that are usually replaced by *GCs*, and to test a more accurate way of redistributing deaths assigned to the *GCs* across selected well-defined causes. This could improve the quality of cause-specific mortality data, enhance more detailed analyses of public health situation and international comparisons.

Data and methods

The data concerned death counts by age, sex and specific causes of death derived from the World Health Organization mortality database (WHO, 2015b). We studied mortality:

– in the period covered by the 10th revision of ICD, different from one country to the other[4];

– in adult age groups, that is aged 40 and above, by different groupings (altogether, grouped in 3 large age groups: 40-59, 60-74, 75 and above, and grouped in 5-year age groups);

– across four European countries with different epidemiologic situations and different frequency of assignment of *GCs*: Czech Republic (13% of deaths due to all *GCs*[5] since 1990), Poland (25%), Russia (5%) and the United Kingdom (7%),

– due to four garbage code causes: atherosclerotic cardiovascular or heart disease (ICD-10 code: I25.0, .1), cardiac arrest (I46), heart failure (I50), atherosclerosis (I70).

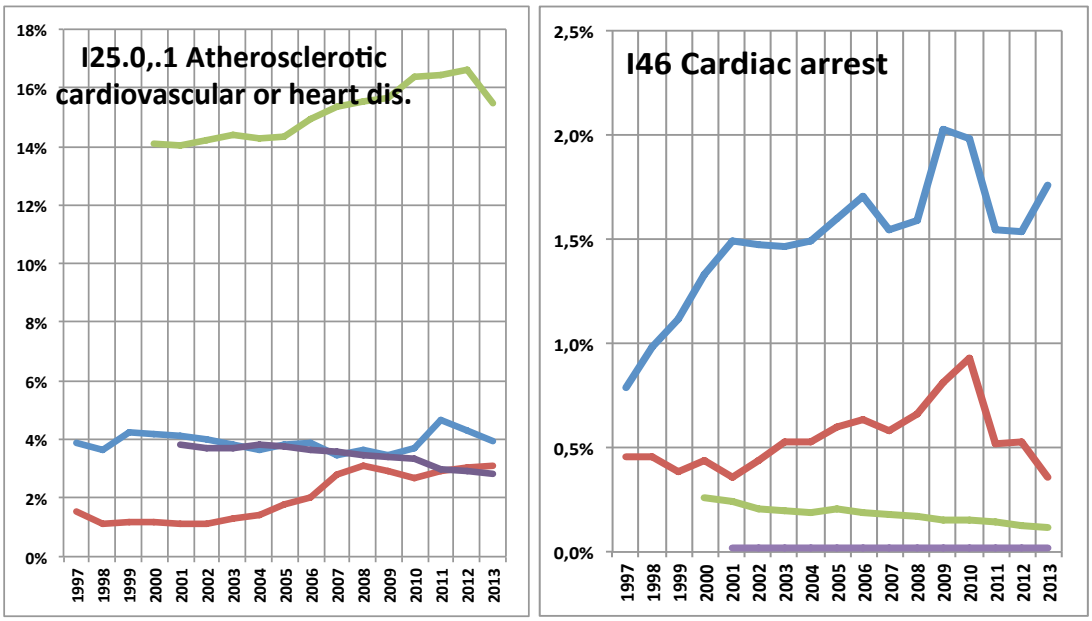
The choice of the above-listed *GCs* was due to their low informative value; atherosclerotic cardiovascular or heart disease (I25.0 and I25.1) belongs to the group of ischaemic heart diseases (I20–I25) that arise when heart tissue die from lack of oxygen due to impaired flow of blood into the heart muscle. This may lead to acute myocardial infarction or acute myocardial infarction (AMI). Cardiac arrest (I46) takes place when a heart fails to contract effectively or at all and, consequently, the blood

[4] For the Czech Republic 1994–2013, Poland 1997–2013, Russia 2000–2013, the United Kingdom 2001–2013.

[5] Defined as in the study by Mathers et al. (2005), see footnote no. 2.

stops to circulate. As cardiac arrest can result from many pathologic processes (all terminal conditions lead to the arrest of heart contractions after all), it should not be coded as an underlying cause of death according to the manual concerning the 10th ICD revision. As for heart failure (I50), it takes place when heart contractions persist, though they are insufficient; it should be considered 'an obvious consequence of other heart conditions', such as AMI, and as such coded differently (WHO, 2008, p. 42). Finally, atherosclerosis (I70) takes place when atherosclerotic plaques accumulate in arteries, thus reducing elasticity of artery walls and impairing the flow of blood. This may lead to decay of such tissues as heart (and development of AMI), brain (stroke), kidneys (renal failure).

The *GCs* were chosen also because of their high frequency in national registration (Figure 1), especially in Poland (the case of heart failure, I50, and atherosclerosis, I70), Russia (the case of atherosclerotic heart disease, ICD-10 code I25.1), and to some extent the Czech Republic (the case of atherosclerosis, I70). In Poland and Russia the frequency of use of some *GCs* – I25.1 in Russia, I50 in Poland has been even increasing recently, despite the WHO recommendations and actions implemented by the Statistical Office of Poland in order to improve the quality of cause-of-death data in this country (i.e. Cierniak, 2014). In the United Kingdom such causes of death as cardiac arrest and atherosclerosis are almost negligible, which proves that in each country the prevailing coding practices are very different and the *GCs* are assigned in a very different way.



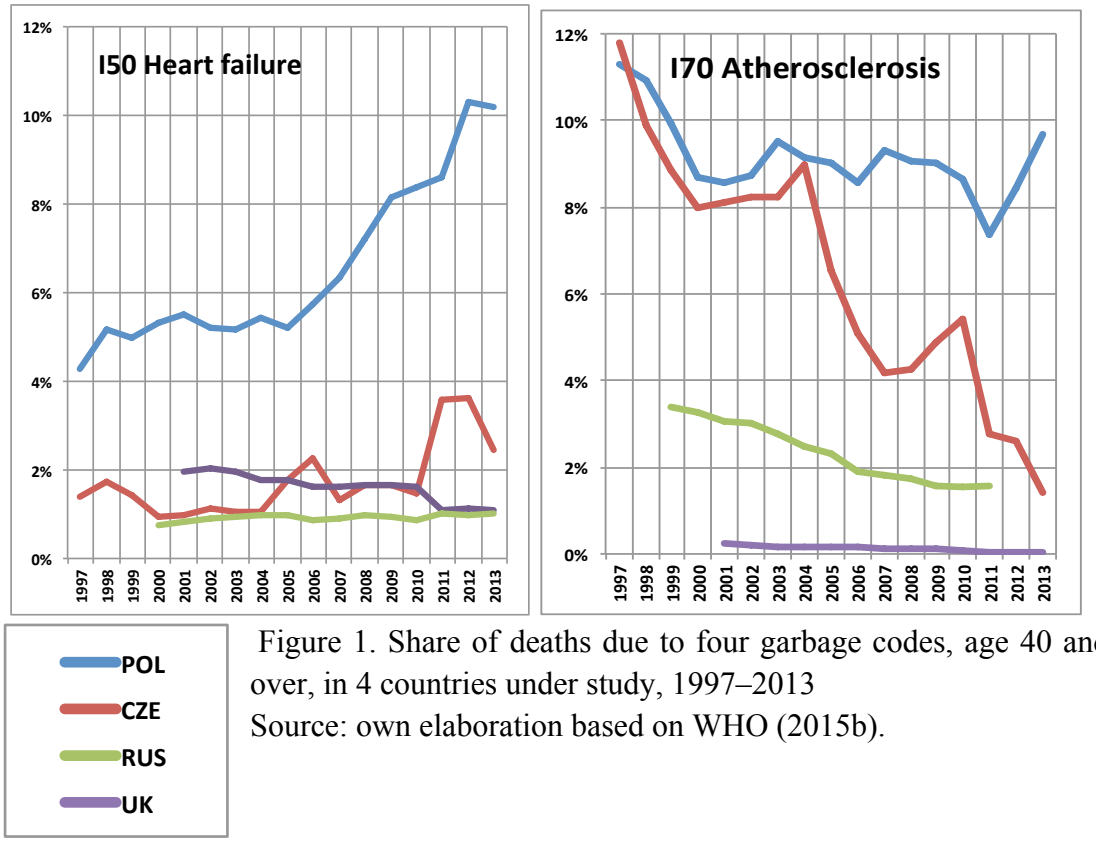


Figure 1. Share of deaths due to four garbage codes, age 40 and over, in 4 countries under study, 1997–2013
Source: own elaboration based on WHO (2015b).

In the study published as early as in 1955 Sully Ledermann (1955) analyzed in geographical dimension statistical correlations between proportions of deaths assigned to ill-defined causes and selected well-defined causes of death in France and Italy. He showed that in regions with high proportion of deaths due to ill-defined causes, proportion of deaths due to pulmonary tuberculosis was low, and *vice versa*, which proves certain exchangeability between those causes of death. Based on these results he proposed a method of redistribution of deaths due to ill-defined causes across well-defined causes for which negative, statistically significant correlations were found: redistributed proportions of deaths were calculated on the basis of slopes standing in linear regressions between ill-defined and particular well-defined causes. In a later study for France similar method was adopted to deal with deaths assigned to three large groups of ill-defined causes (Vallin and Meslé, 1988) and results obtained in this way were more credible than those of simple, proportionate redistribution across all well-defined causes.

Similarly, the method applied in this analysis consisted of seeking reverse trends in mortality due to single causes of death, that is in four *GCs* and selected well-defined causes that in a pathophysiologic sense can be linked with the former. In practice, we searched for negative, statistically significant correlations[6] in age-specific time series of death counts for each analyzed *GC* and causes of death belonging to respiratory

[6] All operations, including tests of statistical significance were made in Stata. Statistical significance was at the level of 0.1 (*), 0.05 (**) or 0.01 (***).

diseases, well-defined cardiovascular diseases, falls or complications stemming from medical interventions (see Table A1 in the Annex for detailed list of causes). In order to find an exchangeability between causes of death, we wanted to allow for annual changes and attenuate long-term trends in levels of mortality, which can be due to long-lasting epidemiologic trends and population ageing. We did so by calculating correlations for first differences in time series, that is for changes, differences between death counts registered in year *t* and year *t-1* [7]. In the hypothetical case of a *GC* that is often registered instead of (interchangeably with) a well-defined cause of death (Figure 2A), first differences resemble a reflection in a mirror (Figure 2B).

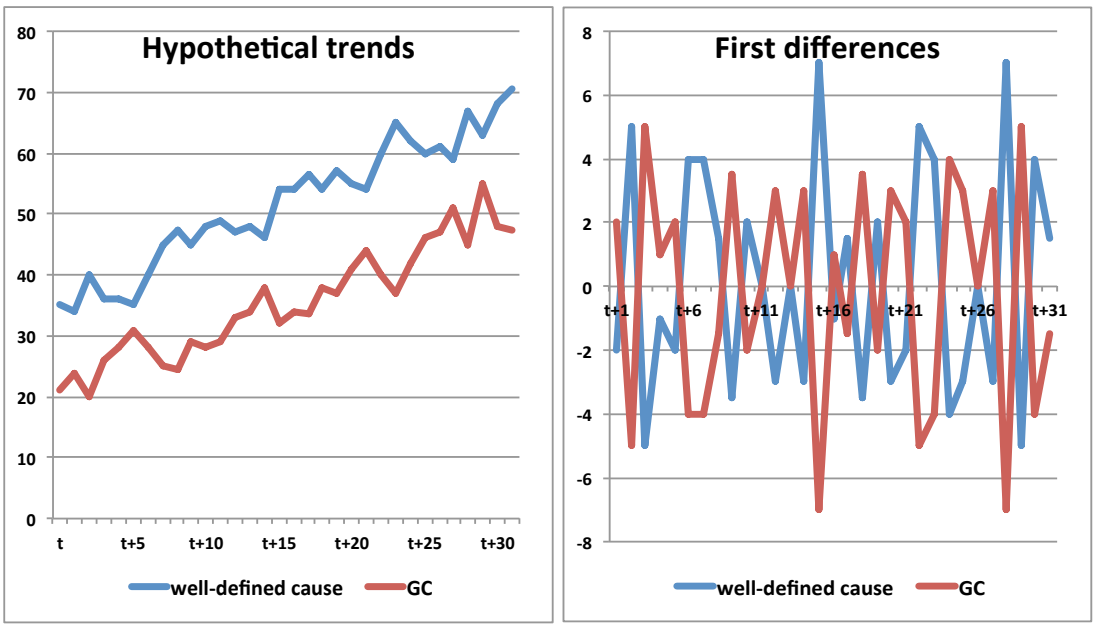


Figure 2A-B. Hypothetical trends in death counts (left) and first differences in trends (right) for a *GC* and a well-defined cause of death.

Preliminary results

1. Correlations between cardiovascular GCs and well-defined causes of death

For each *GC* under study a large number of correlations was verified because we tested different age groupings. Nevertheless, regardless of age grouping we found very few correlations that were negative and statistically significant.

For Poland, where the scale of *GC* coding is large, four *GCs* under study turned out to be negatively correlated with each other: atherosclerotic cardiovascular or heart disease (I25.0, .1) was correlated with cardiac arrest (I46, Figure 3) and atherosclerosis (I70, Figure 4). However, the *GCs* under study appeared not to be correlated with well-defined causes of deaths and this is why we decided to search for significant correlations in geographical dimension. We present results in the point 4, page 13.

[7] For each cause of death: number of deaths registered in year *t* less number of deaths registered in year *t-1*. The term 'first' signifies that we use only first lag of a time series (and ignore the antecedent periods and lags).

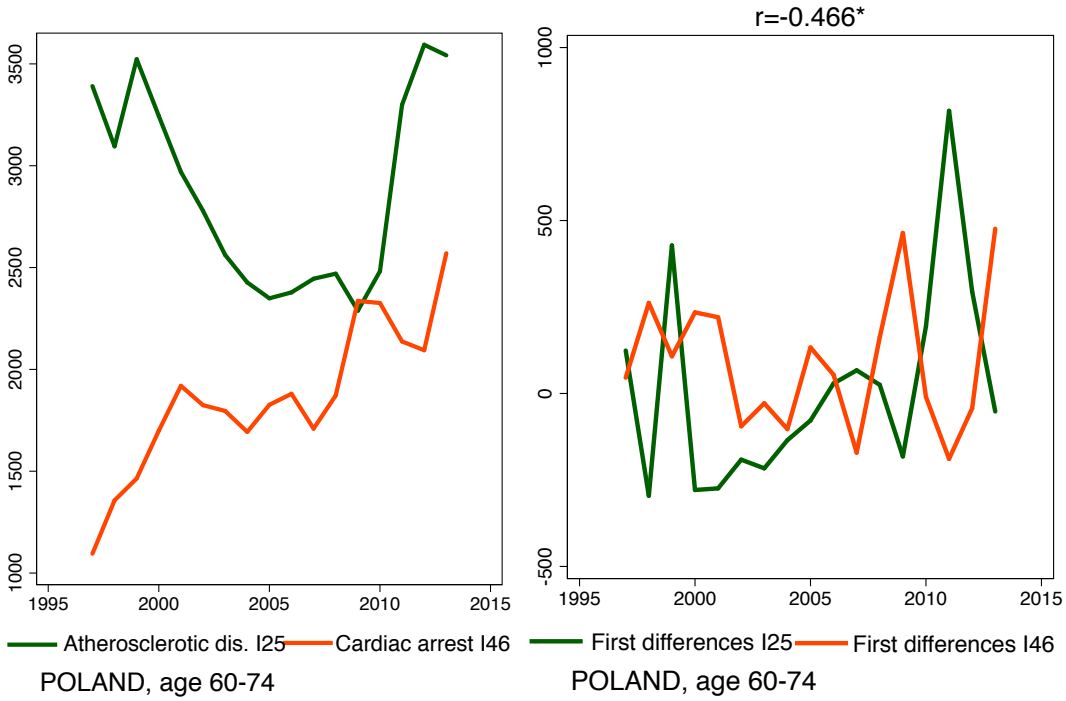


Figure 3A-B. Number of deaths at the age 60-74 (left) and first differences (right) for atherosclerotic cardiovascular or heart disease (I25.0, .1) and cardiac arrest (I46), Poland 1997–2013

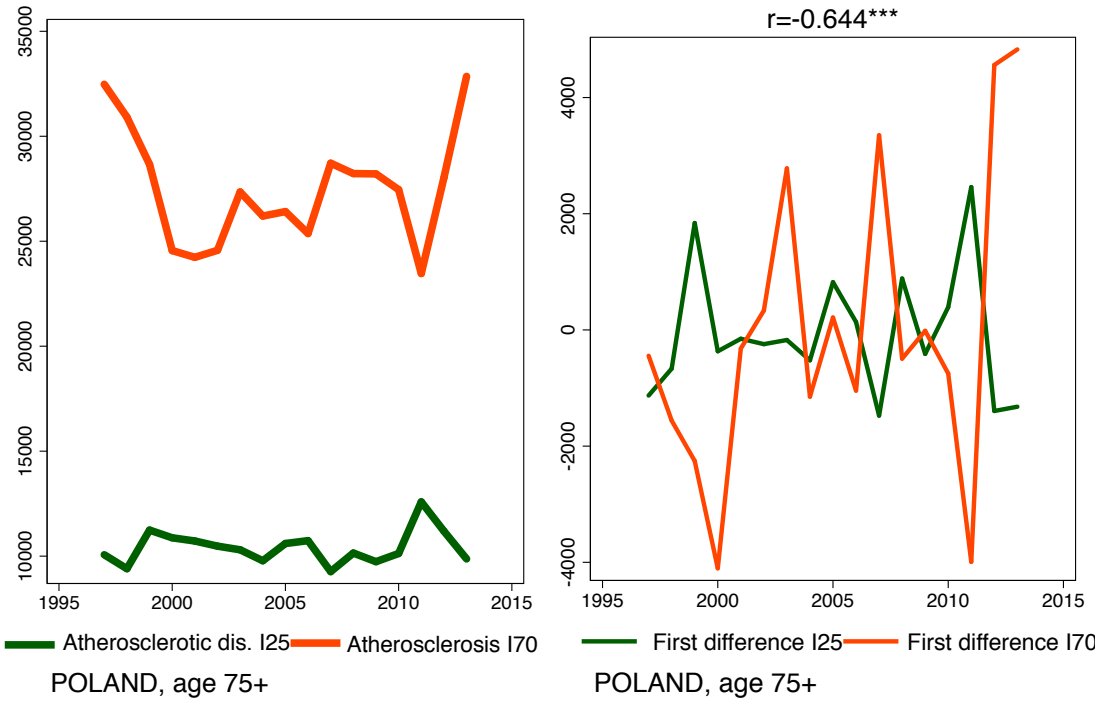


Figure 4A-B. Number of deaths at the age 75 and above (left) and first differences (right) for atherosclerotic cardiovascular or heart disease (I25.0, .1) and atherosclerosis (I70), Poland 1997–2013

The Czech Republic and the United Kingdom exhibited significant changes in *GCs* mortality due to introduction (in the Czech Republic) or up dating (in the UK) of the system of automatic coding. In both countries since 2011, when those changes took place, mortality due to vascular dementia (F01) and unspecified dementia (F03) – diseases belonging to mental and behavioral disorders – has been increasing importantly, while mortality due to some cardiovascular diseases exhibited equivalent downward disruptions.

In the United Kingdom such disruptions concerned mostly unspecified cerebrovascular disease (I67.9) but also *GCs* under study: atherosclerotic cardiovascular and heart disease (I25.0, .1), heart failure (I50) and, to a minor extent, atherosclerosis (I70). The British bridge coding study (Office for National Statistics, 2011) showed that some deaths originally registered as due to atherosclerotic disease have been anew assigned to other cardiovascular diseases (AMI, code I21 and other acute IHD, I24) and unspecified dementia (F03), but our analysis did not prove any negative relation between those disease entities; the bridge coding study suggested also that deaths due to heart failure have been anew assigned to unspecified dementia (F03), atrial fibrillation and atrial flutter (I48) or pneumonia (J18), but again we did not find any significant correlation. Both diseases, however, exhibited a significant and negative relation with vascular dementia (F01), see Figure 5 and 6.

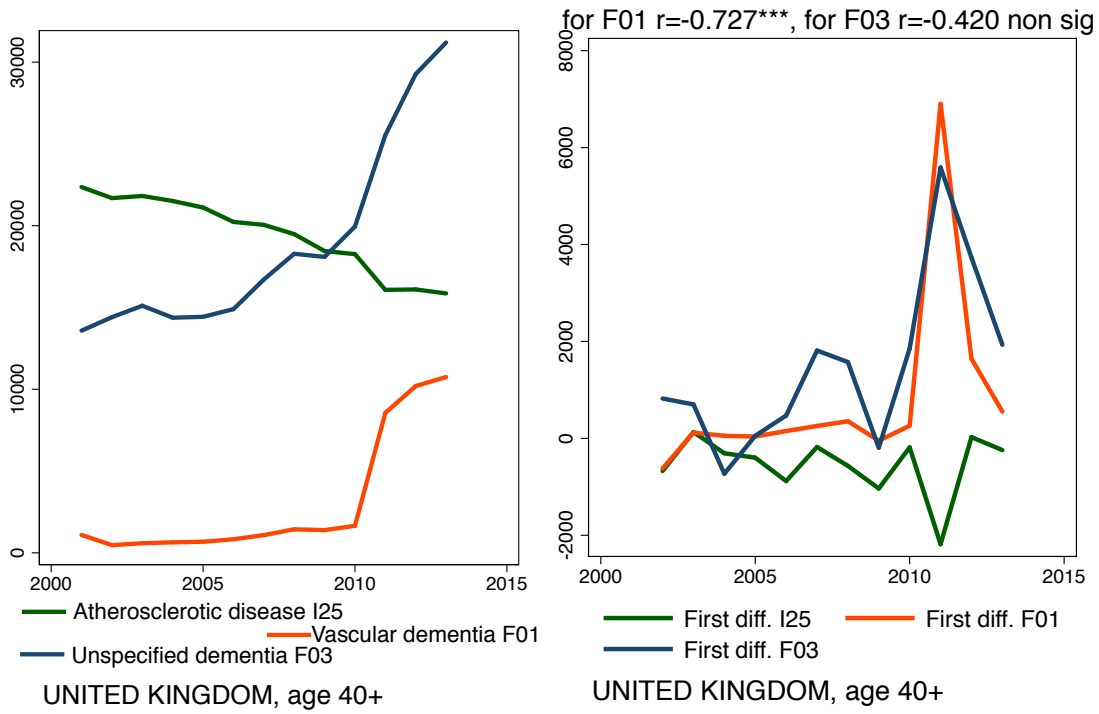


Figure 5A-B. Number of deaths at the age 40 and above (left) and first differences (right) for atherosclerotic vascular and heart disease (I25.0, .1), vascular dementia (F01) and unspecified dementia (F03), the United Kingdom 2001–2013

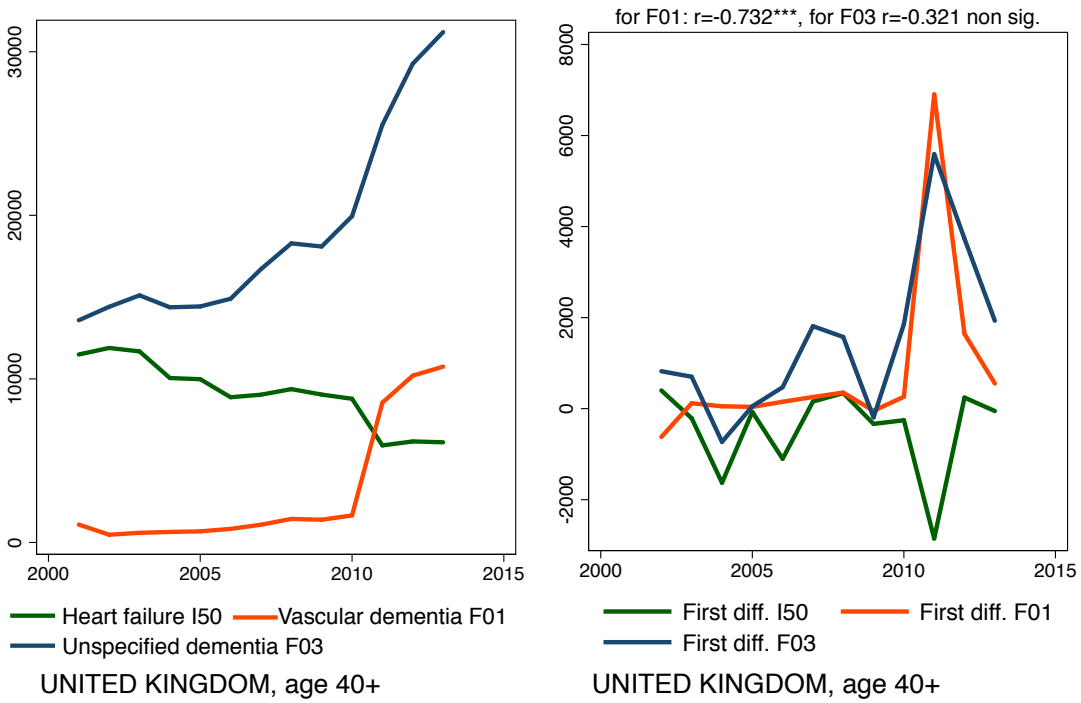


Figure 6A-B. Number of deaths at the age 40 and above (left) and first differences (right) for heart failure (I50), vascular dementia (F01) and unspecified dementia (F03), the United Kingdom 2001–2013

In the Czech Republic the introduction of automatic coding in 2011 entailed an important decrease in mortality due to atherosclerosis (I70) and, to a lesser extent, due to cardiac arrest (I46). At the same time mortality due to IHD (I20-I25, thus including atherosclerotic cardiovascular and heart disease, I25.0, .1) and due to heart failure (I50) increased. As no bridge coding study was conducted, nothing is known about possible exchanges between those cardiovascular diseases due to the change of method of coding. In 2011 mortality due to cardiovascular dementia (F01) increased considerably, and this increase turned out to be significantly and negatively correlated with the change in mortality from cardiac arrest and atherosclerosis (Figure 7 and 8). Thus, while in the United Kingdom cardiovascular dementia seemed to be coded instead of atherosclerotic disease and heart failure, in the Czech Republic it replaced two other *GCs*, namely cardiac arrest and atherosclerosis, which is in line with specific prevalence of those cardiovascular *GCs* in each country (see Figure 1). As opposed to the United Kingdom, mortality due to unspecified dementia (F03) in the Czech Republic has remained negligible since 2011.

Finally, for Russian mortality correlations were checked between the cardiovascular *GCs* and groups of well-defined cardiovascular diseases, respiratory diseases and falls (one large category). No statistical significant correlation was found, which raises doubts about the method of analysis that we incorporated (see discussion).

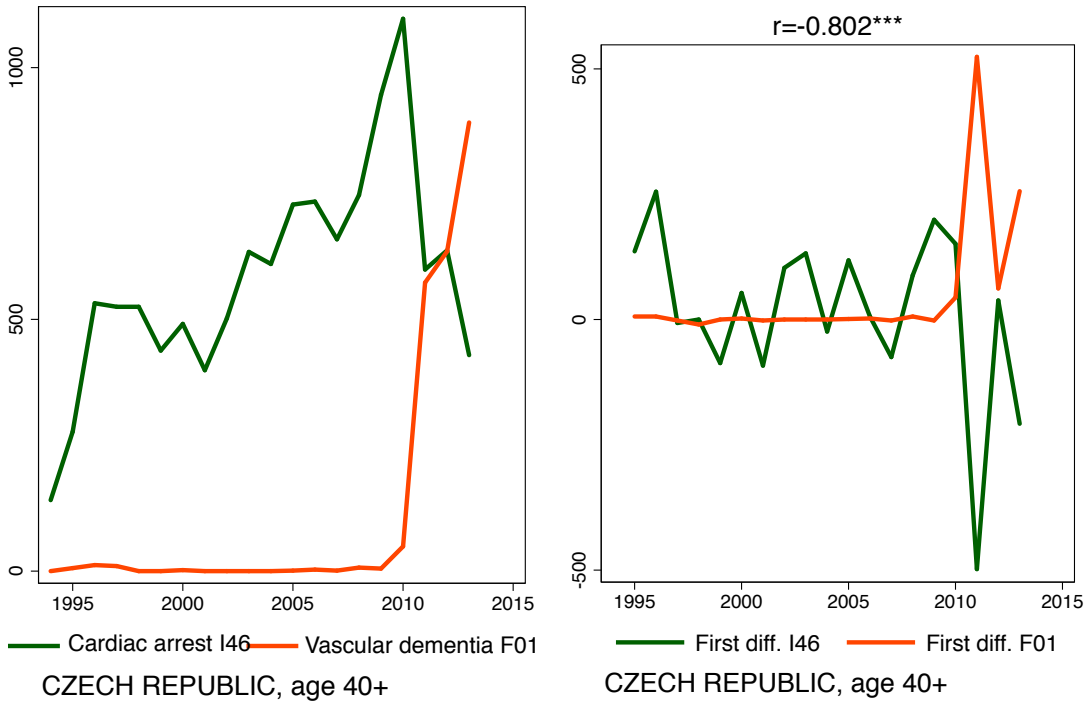


Figure 7A-B. Number of deaths at the age 40 and above (left) and first differences (right) for cardiac arrest (I46) and vascular dementia (F01), the Czech Republic 1994–2013

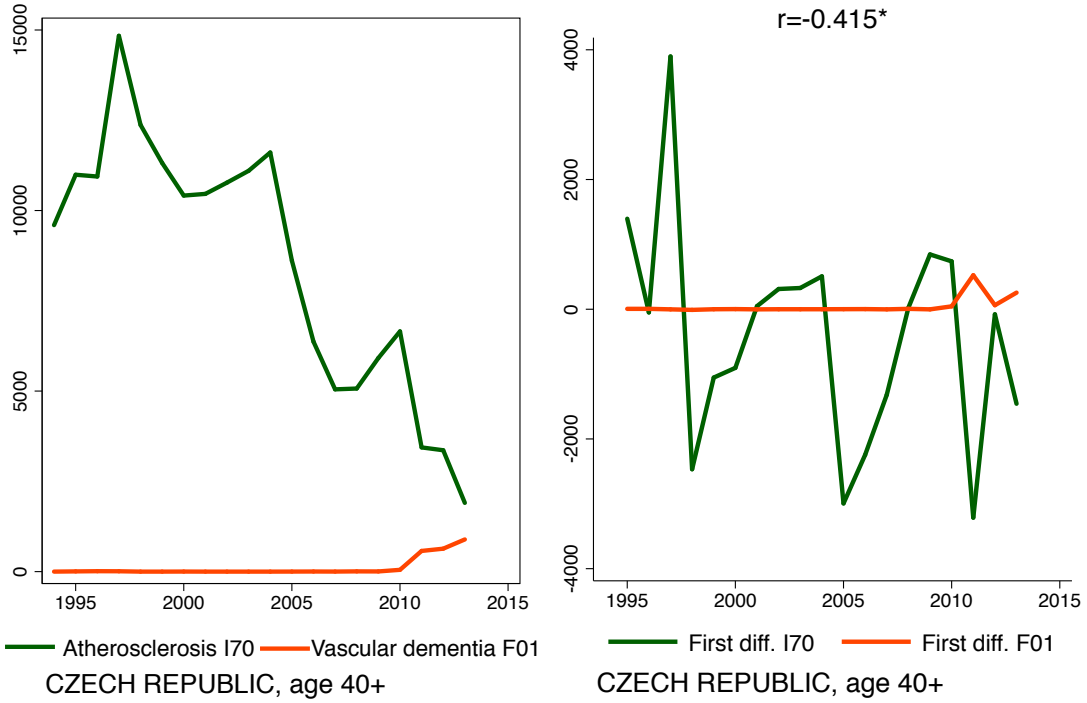


Figure 8A-B. Number of deaths at the age 40 and above (left) and first differences (right) for atherosclerosis (I70) and vascular dementia (F01), the Czech Republic 1994–2013

2. Redistribution of cardiovascular GCs

We decided to redistribute the cardiovascular *GCs* for the Czech Republic and the United Kingdom, for which negative, statistically significant correlations were found between the *GCs* and well-defined causes of death. The adjustments included:

- for the Czech Republic: redistribution of cardiac arrest, I46 across vascular dementia, F01. In this case we ignored the fact that another *GC*, atherosclerosis (I70) was also correlated with vascular dementia. Disruptions in time series of mortality due to I46 and F01 seemed to be equal and an additional redistribution of atherosclerosis (an item assigned to thousands of deaths each year) across F01 would considerably increase the latter and contribute to another disruption in time series;
- for the United Kingdom: redistribution of atherosclerotic disease, I25.0, .1 and heart failure, I50 across vascular and unspecified dementia, F01 and F03. In this case we included unspecified dementia, though we did not find a statistically significant correlation, as the British bridge coding study showed an exchange between this disease, atherosclerotic disease and heart failure.

Adjustments were made at the age-specific level (for simplicity we show below only aggregate results) and for the periods prior to 2011, the first year of new coding practices. Firstly, we assumed for each *GC* that the number of death count registered in the year preceding the disruption should be equal to that registered in 2011. On this basis we calculated proportion of deaths that were 'redundant' and should be moved from each *GC* to a well-defined cause. Second, we applied the afore-mentioned proportion to the whole period preceding the change of coding system. If there were more than one well-defined cause, as in the case of United Kingdom, we redistributed 'redundant' death counts proportionally between specific well-defined causes, according to their scale registered in 2011.

Obviously, in the United Kingdom where vascular and unspecified dementia became important in the aftermath of changes in coding practices in 2011, the adjustments were relatively large for all causes of death concerned by redistribution: thousands of deaths were moved from the two *GCs* under study to vascular and unspecified dementia in the period prior to 2011 (Figure 9). As a result of redistribution, mortality due to atherosclerotic disease and heart failure became smaller in this country by 12 and 32%, respectively (2001–2010). The adjusted time series for vascular dementia still exhibit a clear disruption between 2010 and 2011, as obviously there were also other (than the two *GCs*) causes of deaths that before 2011 used to be assigned instead of vascular dementia. In the Czech Republic, where mortality from vascular dementia remained relatively small after the introduction of automatic coding, the results of adjustments turned out to be less important (Figure 10) but still, mortality due to cardiac arrest diminished by 43% (1994–2010).

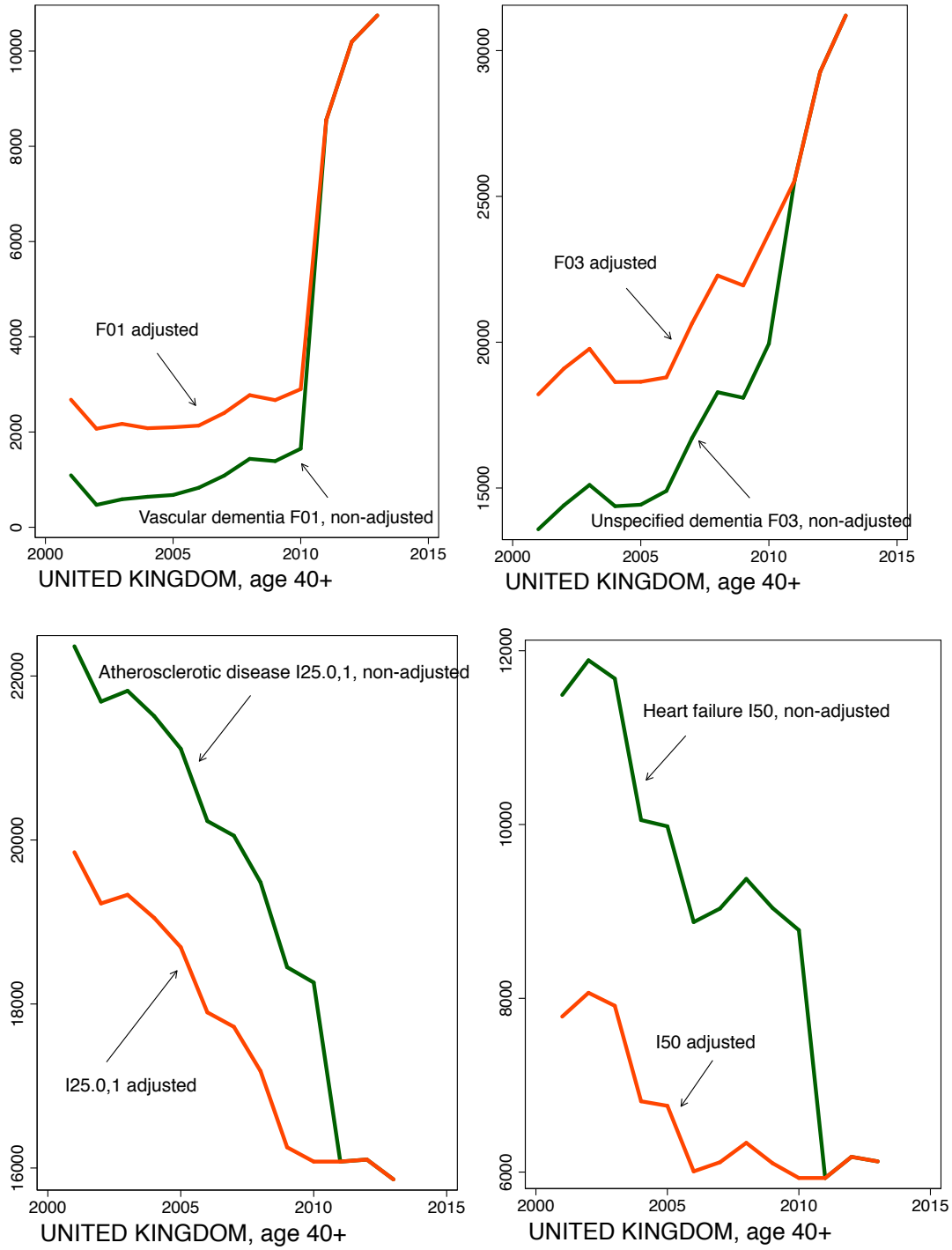


Figure 9A-D. Number of deaths at the age 40 and above for vascular dementia (F01), unspecified dementia (F03), atherosclerotic disease (I25.0, .1) and heart failure (I50) non-adjusted and adjusted, the United Kingdom 2001–2013

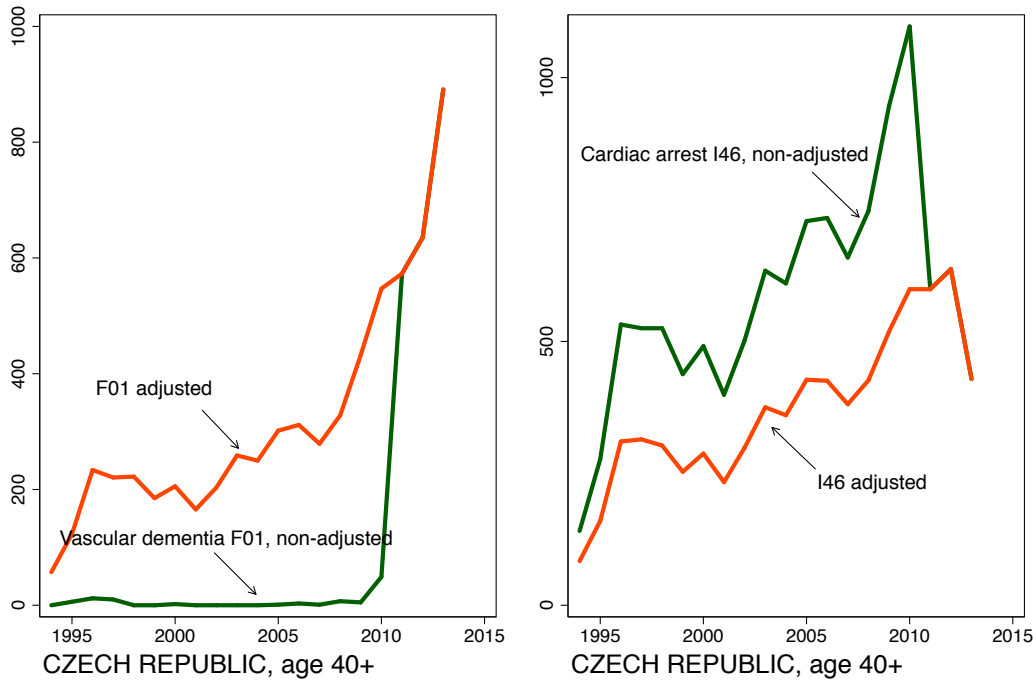


Figure 10A-C. Number of deaths at the age 40 and above for vascular dementia (F01) and cardiac arrest (I46) non-adjusted and adjusted, the Czech Republic 1994–2013

3. Searching for regional correlations in Poland

In this study we failed to find statistically significant correlations between cardiovascular *GCs* and well-defined causes of death in Poland. However, other studies proved that mortality due to ill-defined causes has been negatively correlated with such diseases, as pneumonia or diabetes mellitus (Bijak, 2003; Fihel and Pechholdová, 2016). Therefore, we decided to search for correlations in the geographical dimension, that is, between 380 regions of the NUT4 level[8]. In this case we could not use death counts as the number of local population varies significantly between regions; instead, we used shares of cause-specific deaths in the overall numbers of death registered in each region. We used average shares from the most recent period, 2012–2014 and due to lack of more-detailed data, for all age groups.

At the cause-specific level we found negative, statistically significant correlations between each important cause of death and at least one cardiovascular *GC*. Correlations were found even for such causes of death that seem not to be linked in a pathophysiologic way with cardiovascular *GCs* and that can be univocally diagnosed and documented during medical treatment, such as malignant neoplasm of stomach (C16), colon (C18) or bronchus and lung (C34). At the level of *ICD* chapters we found negative, statistically significant correlations between 13 (out of 19) chapters and four (aggregated) cardiovascular *GCs* (Table A2 in the Annex). Following the

[8] So-called *poviats* or, keeping the Polish spelling, *powiat*. We used data from the National Statistical Office of Poland website, http://demografia.stat.gov.pl/bazademografia/ .

method proposed by Ledermann, we decided to redistribute deaths assigned to the four cardiovascular *GCs* across *ICD* chapters for which negative, statistically significant correlations were found. To each *ICD* chapter we attributed the proportion of *GC* deaths according to the slope in linear regression between the *GCs* time series and each *ICD* chapter time series. We did it for each age group separately using one, common for all age groups (as more specific data were not available) slope of linear regression. The total of slopes exceeded -1 (was equal to -1.017) and therefore, we rescaled the slopes so that their sum was -1. In the next step we redistributed the *GC* deaths inside each *ICD* chapter proportionately to the importance of each cause of death.

Not surprisingly, the highest slope of linear regression was found for the group of cardiovascular diseases: -0.459, which means than an increase in percentage of deaths due to well-defined cardiovascular diseases from one Polish region to another is associated with a decrease in share of deaths due to cardiovascular *GCs* by 0.459%. This slope after rescaling (so that the total of slopes equaled to -1) was -0.45, so we attributed 45% of deaths due to cardiovascular *GCs* to the chapter of cardiovascular diseases, which over 17 years (1997–2013) gave 577,5 thousand of deaths. We present below the results of standardized death rates for two important cardiovascular diseases: acute myocardial infarction (I21) and stroke, not specified as haemorrhage or infarction (I64). Obviously, the SDRs computed with the use of regional coefficient (0.45) are situated between the SDRs without any redistribution of *GCs* and with entire redistribution of (all deaths due to) four cardiovascular *GCs* across the chapter of cardiovascular diseases (Figure 11).

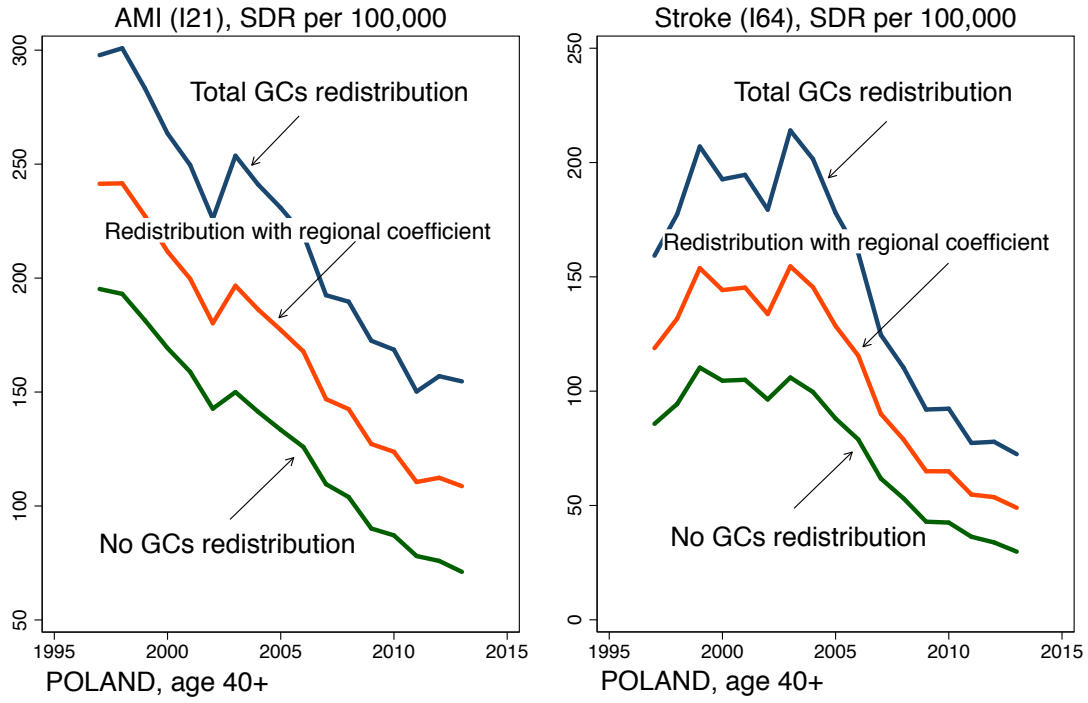


Figure 11 A-B. Standardized death rate (per 100,000) for acute myocardial infarction (I21) and stroke (I64), without *GCs* deaths redistributed, with all *GCs* deaths redistributed across the *I* chapter and with partial redistribution with the use of regional coefficient, Poland, 1997–2013

4. Towards comparability of cardiovascular mortality

The aim of developing non-standard, more specific than usual methods of redistribution of deaths due to *GCs* is to improve comparability of cause-specific time series between countries. In this study we presented two methods of *GCs* redistribution: based on negative correlations between *GCs* and well-defined causes in the temporal dimension (the case of Czech Republic and United Kingdom) and in the regional dimension (the case of Poland). In order to check whether presented methods contribute to a better comparability of mortality data, we compared:

- raw cause-specific time series, that is, without any adjustment,
- time series with cardiovascular *GCs* redistributed proportionately across cardiovascular diseases,
- time series with corrected cardiovascular *GCs* (for the Czech Republic and the United Kingdom) or regionally adjusted *GCs* (for Poland) redistributed proportionately across cardiovascular diseases.

We present the results only for the acute myocardial infarction and stroke, not specified as haemorrhage or infarction (Figure 12). Obviously, the redistribution of *GCs* increased importantly the level of mortality due to well-defined cardiovascular diseases, for instance for the AMI on average by 21% in the United Kingdom and by 36% in the Czech Republic, for the stroke by 18% in the former and by 37% in the latter. Our correction of cardiovascular *GCs* for the Czech Republic and the United Kingdom did not affect in a significant way the level of mortality due to well-defined cardiovascular diseases: the lines representing mortality with adjusted and non-adjusted *GCs* redistributed for Czech Republic overlap, whereas the lines for the United Kingdom are very close to each other. Although our correction of *GCs* improved the coherence and comparability of time series of mortality due to vascular and other dementia, it was too subtle to have an impact on large entities among cardiovascular diseases.

On the contrary, the regional analysis for Poland improved the international comparability of the cause-specific data; the level of mortality due to AMI and stroke without any *GC* redistributed is in Poland close to that registered in the United Kingdom, which is incredible given the epidemiologic trends and the quality of health care in the Central Europe. After redistribution of 45% deaths assigned to cardiovascular *GCs* (in line with the Ledermann's method) the level of cardiovascular mortality approached that of the Czech Republic and after redistribution of all *GCs* deaths it exceeded in the most recent period that of the Czech Republic. In the near future we shall expand the scope of this research and redistribute also deaths assigned to other *GCs*, including ill-defined causes of death (*R* chapter).

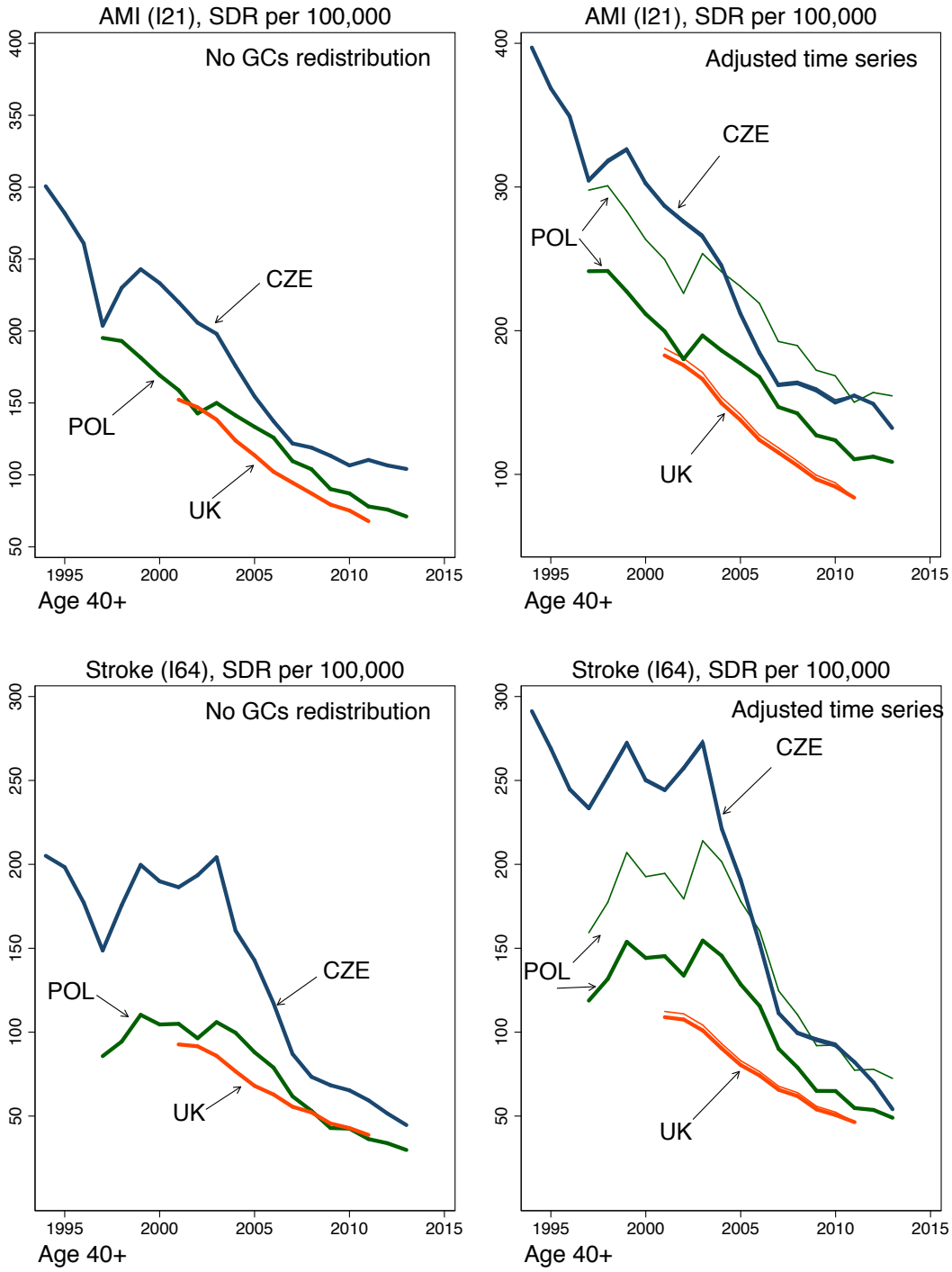


Figure 12 A-D. Standardized death rate (per 100,000) for acute myocardial infarction (I21) and stroke (I64), without *GCs* deaths redistributed, with all *GCs* deaths redistributed across the *I* chapter proportionately (thin lines) and with adjusted *GCs* deaths redistributed (bold lines)

Summary
In this study we aimed at testing a more accurate – than generally adopted – method of redistribution of *garbage codes (GCs)* across well-defined causes of deaths. We selected four important cardiovascular *GCs*: atherosclerotic cardiovascular or heart disease, heart failure, cardiac arrest and atherosclerosis.

The first method consisted of two steps; first assumed finding negative, statistically significant correlations between changes (first differences) in mortality due to a single *GC* and well-defined causes of death. We checked many correlations between the *GCs* and selected well-defined causes of death, consisting of single entities or aggregated in groups of causes. We tested various age- and cause grouping but few statistically significant correlations were found, and even at the age-specific level and for countries with large populations (and high numbers of deaths) we did not manage to disentangle the effects that can be attributed to long-lasting epidemiology of single causes of death from those relating to specific coding practices. Yet, for the correlations that we found, in the second step, we calculated 'redundant' deaths that should be moved from *GCs* to well-defined causes of death. In this way we managed to adjust cause-specific mortality: reduce mortality due to the *GCs* and get rid of disruptions in time trends of well-defined causes of death. This method captured only statistical disruptions resulting from sudden and important changes in coding, such as introduction of automatic system of coding in the Czech Republic or updating it in the United Kingdom. The examples of those countries prove that automatic coding accompanied by a bridge coding study leads to a reduction in the *GC* mortality and to an increase in mortality from well-defined causes of death.

The first method turned out to be unsuccessful for Poland and Russia, where the frequency of *GCs* registration is relatively high. Spatial differences in registration of cause-specific mortality, that remain important in Poland (i.e. Wojtyniak et al., 2012) and Russia (i.e. Danilova et al., 2016) may hinder capturing common coding practices consisting of exchanging well-defined causes of death with the use of *GCs*. Therefore, we tested the second method that was proposed by Ledermann in the 1950s for regional dimension. In the study for 380 regions of Poland we found some negative, statistically significant correlations between shares of deaths assigned to the cardiovascular *GCs* and shares of deaths assigned to entire *ICD* chapters. On the basis of slopes of linear regressions between those time series we redistributed deaths assigned to *GCs* across *ICD* chapters for which statistically significant correlations were found and, inside each chapter, across all causes proportionately to their importance. As a result, the level of mortality from well-defined cardiovascular diseases increased in Poland significantly and became comparable to that registered in other countries of the region of Central Europe (in this exercise: in the Czech Republic).

Obviously, when dealing with *GCs* there is no single general solution for all countries and all causes of death. This is why we seek to develop and test more detailed

methods, adjusted to each country's specificity concerning coding practices and epidemiologic reality. Presented study shows that there are some specific solutions that restore the coherence and international comparability of cause-specific time trends, such as adjustment of mortality from dementia in the Czech Republic and the United Kingdom or from well-defined cardiovascular diseases in Poland. Nevertheless, these solutions require the most detailed data by specific causes of death, small age groups and regions, which is hard to obtain. We plan to continue our study by developing regional methods of *GCs* redistribution for those countries for which the most detailed data are accessible.

Bibliography

Bijak, J., 2003. Międzynarodowa porównywalność danych o zgonach według przyczyn w badaniu regionalnych różnic umieralności na przykładzie Czech, Holandii i Polski w latach 1994-1996. Studia Demograficzne 144, 3–53.

Cierniak, M., 2014. Na co umarł pacjent - czyli co jest wpisywane na kartach zgonu? (what the patient died from - that is, what is written in death certificates? Główny Urząd Statystyczny, Warszawa.

Danilova, I., Shkolnikov, V., Jdanov, D., Meslé, F., Vallin, J., 2016. Identifying potential differences in cause-of-death coding practices across Russian regions. Population Health Metrics 14.

Fihel, A., Muszyńska, M., Wróblewska, W., 2014. Umieralność z przyczyn nieznanych i niedokładnie określonych oraz jej trwałe zróżnicowanie terytorialne w Polsce. Studia Demograficzne 165, 83–102.

Fihel, A., Pechholdová, M., 2016. Analysis of reconstructed series – Poland and Czech Republic. Presented at the Patterns of mortality and causes of death across developed countries, Rostock.

Ledermann, S., 1955. La répartition de decés de causa indeterminée. Review of the International Statistical Institute 23, 47–57.

Mathers, C., Fat, D.M., Inoue, M., Rao, C., Lopez, A.D., 2005. Counting the dead and what they died from: an assessment of the global status of cause of death data. Bulletin of the World Health Organization 83, 171–180.

Meslé, F., Vallin, J., 2012. Mortality and causes of death in 20th-century Ukraine, Demographic Research Monographs. Springer, Dordrecht.

Meslé, F., Vallin, J., 2003. Mortalité et causes de décès en Ukraine au XXe siècle, Les cahiers de l'INED. INED, Paris.

Murray, C., Lopez, A.D., 1996. The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020. Harvard University Press, Cambridge.

Naghavi, M., Makela, S., Foreman, K., O'Brien, J., Pourmalek, F., Lozano, R., 2010. Algorithms for enhancing public health utility of national causes-of-death data. Population Health Metrics 8, 9.

Office for National Statistics, 2011. Results from the ICD–10 v2010 bridge coding study, Statistical Bulletin. Office for National Statistics.

Vallin, J., Meslé, F., 1988. Les causes de décès en France de 1925 à 1978, Travaux et Documents. INED, Paris.

WHO, 2015a. World Health Statistics 2015. WHO, Geneva.

WHO, 2015b. WHO Mortality Database.

WHO, 2013. WHO methods and data sources for global causes of death 2000−2011. WHO, Geneva.

WHO, 2008. International Statistical Classification of Diseases and Related Health Problems, Tenth Revision. WHO, Geneva.

Wojtyniak, B., Rabczenko, D., Pokarowski, P., Poznańska, A., Stokwiszewski, J., 2012. Atlas umieralności ludności Polski w latach 1999–2001 i 2008–2010 [Mortality atlas for the population of Poland, 1999-2001 and 2008-2010].

Annex

Cause of death	ICD-10 code
Diabetes mellitus	E10-E14
Vascular dementia	F01
Parkinson disease	G20
Chronic rheumatic heart diseases	I05-I09
Hypertensive diseases	I10-I15
Ischaemic heart diseases	I20-I24
Pulmonary heart disease and diseases of pulmonary circulation	I26-I28
Cerebrovascular diseases	I60-I69
Diseases of arteries, arterioles and capillaries	I71-I79
Acute upper respiratory infections	J00-J06
Influenza	J09-J11
Pneumonia	J12-J18
Other acute lower respiratory infections	J20-J22
Chronic lower respiratory diseases	J40-J47
Shock, not elsewhere classified	R57
Other sudden death, cause unknown	R96
Unattended death	R98
Unattended death	R99
Falls	W00-W19
Surgical operation and other procedures as the cause of abnormal reaction of the patient, or of later complication, without mention of misadventure at the time of the procedure	Y83-Y84

Table A1. Causes of death correlated with four analyzed *GCs*

ICD chapter	r-Pearson coefficient	Linear slope	Linear slope rescaled to 1
1. Certain infectious and parasitic diseases	-0.204***	-0.01	-0.099
2. Neoplasms	-0.394***	-0.185	-0.182
3. Diseases of blood and blood-forming organs and certain immune disorders	-0.145***	-0.001	-0.001
4. Endocrine, nutritional and metabolic diseases	-0.357***	-0.053	-0.052
5. Mental and behavioral disorders	-0.445***	-0.038	-0.037
6. Diseases of nervous system	-0.409***	-0.034	-0.033
9. Diseases of circulatory system	-0.637***	-0.459	-0.450
10. Diseases of respiratory system	-0.442***	-0.117	-0.115
11. Diseases of digestive system	-0.353***	-0.047	-0.046
13. Diseases of the musculoskeletal system and connective tissue	-0.142***	-0.002	-0.002
14. Diseases of the genitourinary system	-0.251***	-0.023	-0.023
20. External causes of morbidity and mortality	-0.235***	-0.047	-0.046

Table A2. r-Pearson coefficients and slopes of linear regression between shares in deaths due to a selected *ICD* chapter and due to four cardiovascular *GCs* in 380 regions in Poland, 2012–14 (only negative, statistically significant coefficients; *** 0.01 level)