About Japan Data on Causes of Death

by Yui Ohtsu, Magali Barbieri and Futoshi Ishii Last Updated: 26/06/2024

Part I – Vital statistics and population censuses

1. Death Count Data

Source of data

The successive organizations in charge of processing and distributing cause-of-death data in Japan have been the Cabinet Bureau of Statistics (from 1899 to 1943), the Department of Health Statistics at the Ministry of Welfare (1946-1949), the Division of Health Welfare Statistics at the Welfare Ministry (1950-1978), and the Statistics and Information Department at the Ministry of Health, Labour and Welfare (since 1979).

A single electronic file of aggregated death counts by calendar year of occurrence, sex, age, and medical cause for all years from 1950 to 2013 was provided to the HMD by the Statistics and Information Department via the National Institute of Population and Social Security Research. Updated data for years 1995-2013 and new data for 2014-2021 were downloaded from the World Health Organization Mortality Database (https://www.who.int/data/data-collection-tools/who-mortality-database).

It should be noted that, though the HMD all-cause series for Japan is based on statistics which include late-registered deaths, this is not the case for the cause-of-death series, hence some discrepancies between the original death counts in the two series. More specifically, deaths considered unexpected, accidental or suspicious are referred to a coroner who may order a post mortem or carry out a full inquest to ascertain the reasons for the death and rule out foul play. In most countries, the coroner can only register the death once an investigation has concluded as to the exact cause and circumstances of the death. It is not uncommon, at least for a small proportion of all deaths, for such registration delays to extend into several years. For Japan, however, there is a deficit of deaths in the all-cause datafile compared to the World Health Organization data files for all years 2001-2007 and an excess thereafter (see Appendix II for the exact numbers). For the sake of consistency, the death counts and death rates by cause are provided after adjustment for these late-registered deaths so that the series published on the website are in exact agreement.

Part II –Information on CoD coding

Though vital statistics have been compiled since 1872 and vital registration has been required by law since 1898, the first standard death certificate was introduced in Japan in 1900. It included questions about the cause(s) of death and immediately followed the International Classification of Disease (ICD) scheme in its First Revision. The ICD was further retrospectively applied to the classification of deaths from 1899. Throughout the 20th century, Japan has implemented each ICD revision in a timely manner, as indicated in Table 1. However, no mortality statistics (whether by cause or for all causes combined) were published in 1944-1945 because of the administrative disruption resulting from World War II and, for the same reason, data for 1946, though available, are incomplete and unreliable.

Death statistics are published according to the underlying cause of death, defined as the disease that directly leads to death. The exact definition adopted by the Japanese vital statistic system follows international recommandations. It is « (a) the disease or injury which initiated the train of events leading directly to death, or (b) the circumstances of the accident or violence which produced the fatal injury » (World Health Organization, 1975).

Table 1. Periods of implementation of each ICD revision in Japan

ICD Revision	Years Covered
1st	1899-1908
2nd	1909-1922
3rd	1923-1932
4th	1933-1943
5th	1946-1949
6th	1950-1957
7th	1958-1967
8th	1968-1978
9th	1979-1994
10th	1995-present

Implementation of the ICD was a major advance in the comparison of mortality trends by cause across countries. However, the ICD is periodically revised to account for the identification of new diseases (HIV/AIDS, SIDS, or Alzheimer to mention just a few recent examples), improved diagnosis and scientific accuracy of disease classification, and progress in medical knowledge. Though undoubtedly necessary, these revisions introduce disruptions in the series of mortality rates by cause, which are at times very significant. The most accurate method to overcome the issue of classification change is to adjust the series using the results of bridge-coding studies, also called dual-coding studies.

Bridge-coding studies

Bridge-coding studies are regarded as the gold standard for the construction of consistent cause-of-death series. These studies process, code, and classify independently all, or only a representative sample, of the deaths for a given year in both the new revision of an ICD and the previous one. Their main purpose is to quantitatively assess the effects of ICD revisions on time series of mortality statistics.

Japan conducted bridge-coding studies to document the transition from the 7th to the 8th Revision of the ICD, from the 8th to the 9th Revision, and again from the 9th to the 10th Revision. However, because these studies were based on the Statistics Office condensed list of cause-of-death categories, the results are too rudimentary to be used to adjust the HMD series. Furthermore, the fairly substantial adaptation of the 10th Revision introduced in Japan in 2006 has not been documented by bridge-coding.

Special case of the Japanese cause-of-death classification

As previously mentioned, Japan has implemented the International Classification of Diseases (ICD) from the very first Revision published in 1899. The classification system was automated starting in 1989 and further improved in 1995, when a new death certificate was introduced. These changes led to some disruption in cardiovascular mortality trends as they came with a strong warning to discourage the physicians filling out death certificates to use non-specific cardiovascular condition as causes of death. More specifically, a footnote on the new death certificate bear the following mention: “Please do not write heart failure, respiratory failure, etc… as patients’ terminal condition.” Because this new request was widely publicized in advance of the 1995 change, it also affected death counts from heart disease in 1994.

As regards Japan-specific cause of death categories, it should be noted that, to serve its own purpose, Japan has added sub-categories to existing ICD-10 categories by introducing a 5th digit.

Raw data treatment

It is important to carefully consider the medical content of the list. The lists published on the site include only items that can serve as principal, underlying causes of death (UCD). They exclude codes that can not act as mortality causes (such as asterisk codes or postprocedural disorders). Since the classification system is constantly evolving, some obsolete designations may appear. The reference classification for the published data is the 2016 revision of ICD-10. Additionally, certain items are restricted to one sex only or specific age ranges (e.g., senility or infant mortality codes).

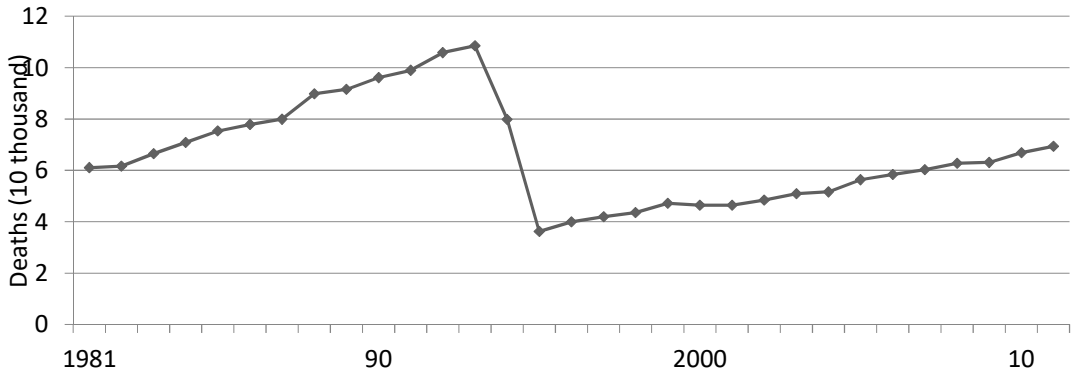
Several corrections and adjustments were applied to the original data to ensure high-quality consistency and comparability:

- Redistribution of deaths with unspecified age: Appendix 3 presents a table listing cases where the column total included a single death with no specification of the age group. These deaths were assigned randomly to any age interval.
- Age-specific cause modification for 1998: In 1998, one death recorded in the 15–20 age group was suspiciously attributed to Parkinson's disease (G20). This case was reassigned to Secondary Parkinsonism (G219) to ensure consistency with medical plausibility.
- Non-UCD deaths reassigned to target causes: the number of deaths attributed to non-UCD causes, which cannot serve as the principal cause of death, were reassigned to target causes. The details of these adjustments are provided in Appendix 4.

Part III – Reconstruction information

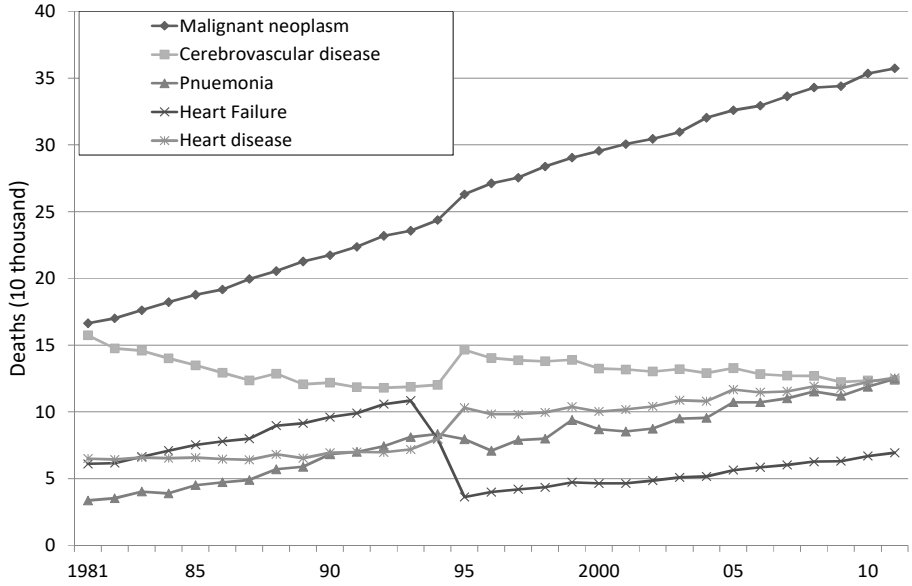
1. Specific treatment of the raw data

As described above, a new standard death certificate was used in Japan in 1995, which was also the first year when ICD-10 was implemented. This led to serious disruptions in cardiovascular mortality trends from 1993 to 1995. For instance, deaths caused by heart failure, which had been gradually increasing until 1993, drastically decreased from 1993 to 1995 (Fig.1). Additionally, a sharp decline in the death count from renal failure was observed between 1994 and 1995. Conversely, deaths in many other causes increased abruptly (Fig.2).



Source: Vital Statistics of Japan, the Ministry of Health, Labour and Welfare (MHLW), Japan.
Note: While ICD-9 code 428 is aggregated in 1994 and earlier, ICD-10 code I50 is aggregated in 1995 and later.

Figure 1: Deaths due to heart failure



Source: Vital Statistics of Japan, the Ministry of Health, Labour and Welfare (MHLW), Japan. Note: Heart disease does not include heart failure and high blood pressure. While ICD-9 code 428 is aggregated in 1994 and earlier, ICD-10 code I50 is aggregated in 1995 and later.

Figure.2: Changes in the mortality rates of major causes of death (by simple classification)

In order to resolve this issue, we redistributed deaths from heart failure in 1993 and 1994 and from renal failure in 1994 and 1995 as described below, following a procedure recommended by Ohtsy et al. (2018).

Stage 1: Adjustment procedure for deaths from heart failure in 1993 and 1994

Between 1993 and 1994, there was a change in the definition of "heart failure, unspecified" (ICD-10 code 428.9) in Japan which translated into a sudden drop in the number of deaths from this cause. We carried out a two-step process to redistribute deaths from heart failures for the period before 1994 to account for the disruption so that the series of death counts for years 1981-1993 is consistent with that for years 1994 and following. In the first step, we identified the causes exhibiting an unexplained symmetric sudden increase in the number of deaths between 1993 and 1994 that could have resulted from the aforementioned change in definition in heart failures and we redistributed deaths from heart failure into these cause-of-death categories. We thus created uniform series of death counts for these other causes with the period before 1994 being consistent with the the period 1994 and thereafter. After this redistribution, there was still a number of excess deaths from heart failure before 1993 compared to 1994. These excess deaths were further redistributed over all other causes of death proportionately over each sex and age category, excluding age 0 for which a slightly modified protocole was implemented as further described later. These two steps are describd in more details below.

Step 1: Modify the deaths of causes with discontinuous increase

(1) Determine the numbers of deaths to be redistributed

At first, we identified causes of death at the ICD-9 3-digit level that experienced a jump between 1993 and 1994. Then, we assessed the number of deaths needed to correct the discontinuity for these causes (a^k in Table 1 below). These causes are listed in Tables 1 and 2.

Table 1: Causes of death to which some deaths from "Heart failure, unspecified" were allocated (3-digit level)

ICD-9	Cause of death	a^k
038	Septicemia	1,000
185	Malignant neoplasm of prostate	200
188	Malignant neoplasm of bladder	200
202	Other malignant neoplasms of lymphoid and histiocytic tissue	200
203	Multiple myeloma and immunoproliferative neoplasms	100
250	Diabetes mellitus	300
263	Other and unspecified protein-calorie malnutrition	20
276	Disorders of fluid, electrolyte, and acid-base balance	500
279	Disorders involving the immune mechanism	40
286	Coagulation defects	100
287	Purpura and other hemorrhagic conditions	40
290	Dementias	200
296	Episodic mood disorders	20
307	Special symptoms or syndromes, not elsewhere classified	30
332	Parkinson's disease	170
345	Epilepsy	70
403	Hypertensive chronic disease	120
410	Acute myocardial infarction	9,300
426	Conduction disorders	100
427	Cardiac dysrhythmias	500
429	Ill-defined descriptions and complications of heart disease	300
438	Late effects of cerebrovascular disease	1,500
444	Arterial embolism and thrombosis	100
446	Polyarteritis nodosa and allied conditions	50
482	Other bacterial pneumonia	200
507	Pneumonitis due to solids and liquids	400
530	Diseases of esophagus	60
590	Infections of kidney	200
595	Cystitis	60
599	Other disorders of urethra and urinary tract	70
710	Diffuse diseases of connective tissue	70

714	Rheumatoid arthritis and other inflammatory polyarthropathies	300
797	Senility without mention of psychosis	1,600
798	Sudden death, cause unknown	150

Note: a^k represents the number of deaths from "heart failure, unspecified" re-allocated to the corresponding cause in 1993.

Table 2: Cause of death to which some deaths from “Heart failure, unspecified” were allocated (4-digit level)

ICD-9	Cause of death	b^i
183.0	Malignant neoplasm of ovary	300

Note: b^i represents the number of deaths from "heart failure, unspecified" re-allocated to Malignant neoplasm of ovary in 1993.

(2) Share of deaths from "Heart failures, unspecified" to be redistributed to the causes listed in Table 1 above

a^k represents the number of deaths from "Heart failure, unspecificed" added to cause k listed in Table 1 in 1993. $d^k_{n,x}(t)$ represents deaths in year t $(1981 \leq t \leq 1993)$, age group n, sex x and cause k with k one of the causes listed at the 3 digit level in Table 1 (e.g., ICD-9 code 038, septicemia). $d^i_{n,x}(t)$ represents deaths in year t $(1981 \leq t \leq 1993)$, age group n, sex x and cause i with i being a sub-category (4 digit level) of cause k (e.g., 038.0, streptococcal septicemia).

We computed $d^i_{n,x}(t)$ for 1993 and earlier as follows:

$$\dot{d}^i_{n,x}(t) = d^i_{n,x}(t) + a^k \cdot \frac{d^i_{n,x}(t)}{\sum_{n,x} d^k_{n,x}(1993)}$$

$$\text{where } d^k_{n,x}(1993) = \sum_{i \in k} d^i_{n,x}(1993).$$

(3) Share of deaths from "Heart failures, unspecified" to be redistributed to the causes listed in Table 2 above

b^i represents the number of deaths from "Heart failure, unspecificed" added to cause i listed in Table 2 in 1993 . $d^i_{n,x}$ for 1993 and earlier was similarly calculated as follows:

$$\dot{d}^i_{n,x}(t) = d^i_{n,x}(t) + b^i \cdot \frac{d^i_{n,x}(t)}{\sum_{n,x} d^i_{n,x}(1993)}$$

(4) Deaths remaining in the “Heart failure, unspecified” category

The total number of deaths added in (2) and (3) were deducted from deaths attributed to "Heart failure, unspecified" by age group, sex and year.

$$\dot{d}_{n,x}^{428.9} = d_{n,x}^{428.9} - \sum_{i} (\dot{d}_{n,x}^{i} - d_{n,x}^{i})$$

Step 2: Redistribution of the remaining excess deaths from heart failure

A second step had to be carried out because the redistribution implemented in step 1 was insufficient to smooth out the drop in the number of deaths from "Heart failure, unspecified" from 1993 to 1994,.

(1) Modifying the disruption in "Heart failure, unspecified" completely

We modified the number of deaths from "Heart failure, unspecified" in 1993 as follows:

$$\hat{d}_{n,x}^{428.9}(1993) = \begin{cases} \dot{d}_{n,x}^{428.9}(1993), & \dot{d}_{n,x}^{428.9}(1993) \leq d_{n,x}^{428.9}(1994) \\ \dot{d}_{n,x}^{428.9}(1994), & \dot{d}_{n,x}^{428.9}(1993) > d_{n,x}^{428.9}(1994) \end{cases}$$

For all sex and age groups where the number of deaths from "Heart failure, unspecified" was larger in 1993 (after carrying out the steps previously described) than in 1994, we substituted the number of deaths from "Heart failure, unspecified" to the number actually observed in 1993 for each of these sex and age groups. For all other cases (for all sex and age groups where deaths from "Heart failure, unspecificed" was smaller in 1993 than in 1994), we simply preserved that number.

(2) Expanding this modification to years 1992 and earlier

We multiplied the number of deaths from "Heart failure, unspecified" for 1992 and earlier after implementing the correction described in 1st step, i.e., $\dot{d}_{n,x}^{428.9}(t)$, by the ratio of the deaths for 1993 after correction of (1), $\hat{d}_{n,x}^{428.9}(1993)$, to those after correction of 1st step, $\dot{d}_{n,x}^{428.9}(1993)$, for each sex and age group. That is, we assumed that the share of deaths "Heart failure, unspecified" to be redistributed to each of the causes listed in Tables 1 and 2 above from 1993 to 1994 also applied to all other years before 1993.

$$\hat{d}_{n,x}^{428.9}(t) = \dot{d}_{n,x}^{428.9}(t) \cdot \frac{\hat{d}_{n,x}^{428.9}(1993)}{\dot{d}_{n,x}^{428.9}(1993)}$$

(3) Redistributing the remainder of excess deaths from heart failure

The difference of deaths due to heart failure, $\dot{d}_{n,x}^{428.9}(t) - \hat{d}_{n,x}^{428.9}(t)$, for 1993 and earlier was redistributed to all causes of death except for ICD-9 code 428.9 (i.e. "Heart failure, unspecified") and 800-999 (i.e. injury and poisoning).

$$\hat{d}^{i}_{n,x}(t) = \dot{d}^{i}_{n,x}(t) + \frac{\dot{d}^{i}_{n,x}(t)}{\sum_{i \in I} d^{i}_{n,x}(t)} \cdot \{\dot{d}^{428.9}_{n,x}(t) - \hat{d}^{428.9}_{n,x}(t)\}$$

where i stands for any causes other than ICD-9 code 428.9 and 800-999, and $\dot{d}^{i}_{n,x}(t) = d^{i}_{n,x}(t)$ for causes which had not been modified in the step 1.

Modify the number of deaths under 1 year of age

Deaths under age 1 were modified in the same way as those for other age groups in step 1. However, the list of causes exhibiting a jump from 1993 to 1994 are different than those for other ages. These causes are listed in Table 3.

Table 3: Causes of death allocated from "Heart failure, unspecified" under 1 year of age

ICD-9	Cause of death	a^k
768	Intrauterine hypoxia and birth asphyxia	40
183	Malignant neoplasm of ovary and other uterine adnexa	300
775	Endocrine and metabolic disturbances specific to the fetus and newborn	5

Stage 2: Adjust the discontinuity related to heart failure and renal failure from 1994 to 1995

We modified the discontinuous decrease of heart failure and renal failure and the increasing deaths of accompanied causes by age group and sex according to the following procedure, very similar to that described in Stage 1 above.

Step 1: Redistribution of heart failure

Step 2: Redistribution of renal failure

Step 1: Redistribution of heart failure

In this step, we redistributed deaths classified in "Heart failure, unspecified" (ICD-9 code 428.9) in 1994 and earlier to other related causes so that the number of deaths from these causes was redistributed from 1994 to 1995 according to the following method (1)-(5).

(1) Determine the numbers of deaths to be redistributed

At first, we identified causes of death for which an upward disruption appeared from 1994 to 1995 corresponding to the sharp decline in the number of deaths attributed to "Heart failure, unspecified." Then, we assessed the number of deaths needed to correct the discontinuity of these causes. We show the list in Table 4.

Table 4: Causes and numbers of deaths allocated from "Heart failure, unspecified" in 1994

ICD-9	Cause of death	α^i
151.9	Malignant neoplasm of stomach, unspecified	2,200
153.9	Malignant neoplasm of colon, unspecified	350
154.0	Malignant neoplasm of rectosigmoid junction	150
154.1	Malignant neoplasm of rectum	500
155.0	Malignant neoplasm of liver, primary	2500
156.1	Malignant neoplasm of gallbladder and extrahepatic bile ducts, Extrahepatic bile ducts	500
157.9	Malignant neoplasm of pancreas, part unspecified	400
159.9	Malignant neoplasm of other and ill-defined sites within the digestive organs and peritoneum, Ill-defined	500
162.9	Malignant neoplasm of bronchus and lung, unspecified	1000
188.9	Malignant neoplasm of bladder, part unspecified	200
202.8	Other malignant neoplasm of lymphoid and histiocytic tissue, Other lymphomas	120
203.0	Multiple myeloma and immunoproliferative neoplasms, Multiple myeloma	150
250.1	Diabetes with ketoacidosis	40
250.2	Diabetes with coma	70
250.7	Diabetes with other specified manifestations	1000
269.9	Unspecified nutritional deficiency	150
273.8	Protein metabolism	70
278.0	Other disorders of plasma protein metabolism	20
307.1	Anorexia nervosa	40
394.9	Other and unspecified mitral valve diseases	70
414.0	Coronary atherosclerosis	500
414.9	Chronic ischaemic heart disease, unspecified	1000
422.9	Other and unspecified acute myocarditis	40
424.0	Mitral valve disorders	100
424.9	Endocarditis, valve unspecified	300
425.4	Other primary cardiomyopathies	300
427.1	Paroxysmal ventricular tachycardia	50
427.5	Cardiac arrest	700
429.9	Heart disease, unspecified	450
428.1	Left heart failure	150
434.9	Cerebral artery occlusion, unspecified	6500
459.9	Unspecified circulatory system disorder	150
410	Acute myocardial infarction	9000
411	Other acute and subacute forms of ischaemic heart disease	2500
412	Old myocardial infarction	400
438	Late effects of cerebrovascular disease	14000
174.9	Malignant neoplasm of female breast, unspecified (only for female)	200
180.9	Malignant neoplasm of cervix uteri, unspecified (only for female)	300

185	Malignant neoplasm of prostate (only for male)	350

(2) Allocate the deaths for redistribution to each age group and sex

Let α^i be the number of deaths from "Heart failure, unspecified" for 1994 to be added to deaths from cause i listed in Table 3. In addition, let $d^i_{n,x}(t)$ be deaths for year t $(1981 \leq t \leq 1994)$, age group n, sex x and cause i with cause i in Table 4. We allocated α^i to each age group and sex proportionally to the disruption in the time series for deaths from cause i in 1994. The allocated deaths for age group n, sex x, cause i is as follows:

$$\alpha^i_{n,x}(1994) = \alpha^i \cdot \frac{d^i_{n,x}(1994)}{\sum_{n,x} d^i_{n,x}(1994)}$$

For three specific ICD-9 code, the redistribution was specific to one sex: i.e., we redistributed α^i to males only for ICD-9 code 185 and females only for ICD-9 code 174.9 and 180.9.

(3) Expand correction to 1993 and earlier

We defined $\alpha^i(t)$, the number of deaths to be added to cause i for the year 1993 and earlier, as follows:

$$\alpha^i_{n,x}(t) = \alpha^i_{n,x}(1994) \cdot \frac{d^{428.9}_{n,x}(t)}{d^{428.9}_{n,x}(1994)}$$

(4) Computing new numbers of deaths for each cause

We added $\alpha^i_{n,x}(t)$ to deaths of each cause in Table 4 within each age group, sex and year:

$$\hat{d}^i_{n,x}(t) = d^i_{n,x}(t) + \alpha^i_{n,x}(t)$$

(5) Correcting the deaths of "Heart failure, unspecified"

At last, we removed the sum of deaths allocated to each cause from deaths of "Heart failure, unspecified" in each age group, sex and year:

$$\hat{d}^{428.9}_{n,x}(t) = d^{428.9}_{n,x}(t) - \sum_i \alpha^i_{n,x}(t)$$

Step 2: Redistribution of renal failure

In this step, we redistributed deaths of "Renal failure, unspecified" (ICD-9 code 586) in 1994 and earlier to other related causes so that the number of deaths from these causes from 1994 to 1995 was consistent, in the same way as the redistribution of deaths from "heart failure, unspecified". However, the redistribution was carried out only for deaths at age 30 years and older. We show the list of causes i and the redistributed numbers of deaths, β^i, in Table 5.

Table 5: Causes of death allocated from “Renal failure, unspecified“ in 1994

ICD-9	Cause of death	β^i
403.9	Hypertensive renal disease, Unspecified	500
580.9	Acute glomerulonephritis with unspecified pathological lesion in kidney	20
583.4	Nephritis and nephropathy, not specified as acute or chronic, With lesion of rapidly progressive glomerulonephritis	20
582.9	Chronic glomerulonephritis with unspecified pathological lesion in kidney	250
583.9	Nephritis and nephropathy, not specified as acute or chronic, With unspecified pathological lesion in kidney	50
588.8	Other specified disorders resulting from impaired renal function	40
590.1	Acute pyelonephritis and acute pyonephrosis	100
584.5	Acute renal failure, With lesion of tubular necrosis	10
584.9	Acute renal failure, unspecified	500
600	Hyperplasia of prostate	70

2. Reconstruction information

Table 6 shows the number and share of all fundamental associations (as described in Pechholdova et al, 2017) and deaths by type in 1995, the year of transition from ICD-9 to ICD-10 in Japan (Pechholdova et al, 2017).

Table 6: Distribution of fundamental associations of items by type and death counts between revisions of 1994 & 1995

Association type	1995 Revision			
	Associations		Deaths (in 1995)	
	Number	Proportion, %	Number	Proportion, %
type 1:1	2 274	62.0	130 993	14.2
type 1:n	877	23.9	143 125	15.5
type n:1	121	3.3	3 294	0.4
type n:n	395	10.8	644 726	69.9
Total	**3 667**	**100.0**	**922 138**	**100.0**

Results of the reconstruction

Figure 3 presents an example of the reconstruction for causes in a complex association (type n:n). The causes of death involved in this association overlap multiple chapters: chapter V (Mental and behavioural disorders), VI (Diseases of the nervous system), X (Diseases of the respiratory system), XVIII (Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified), and XX (External causes of morbidity and mortality).

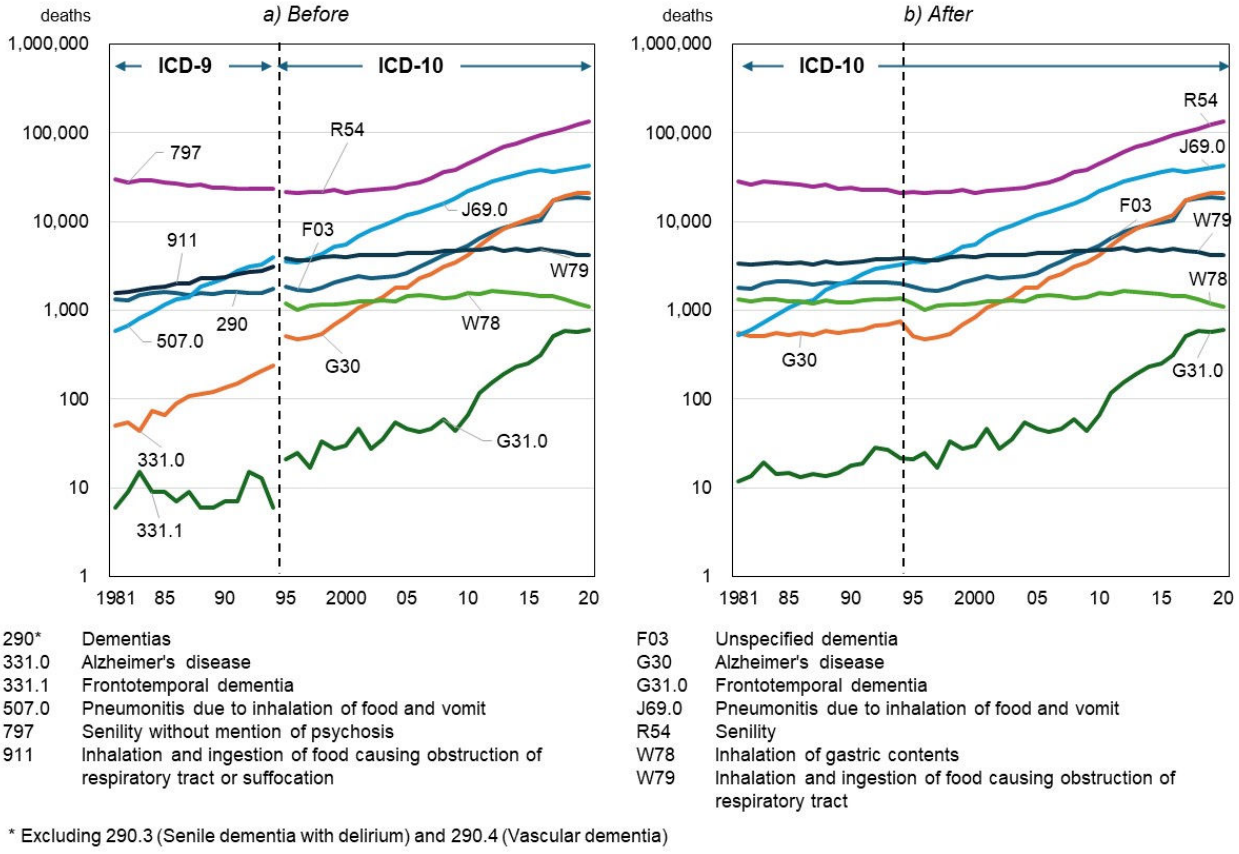


Figure 3. Example of reconstruction. Japan, old-age related diseases, 1981-2020 (from ICD-9 to ICD-10)

3. A posteriori corrections

Note that the corrections described below exclude the disruption observed for a number of causes when, in 2016, Japan started applying the WHO revised version of ICD-10 that had been published in 2013. The ICD-10 codes for which the largest disruptions were identified because of this change are listed in Table 7. However, since very few years have elapsed since then and no stable trend has been observed after the implementation of the new version, we are not modifying them at this time. Other corrections have however been implemented as described in the rest of this section.

Table 7: Summary of a posteriori correction

	From	To	Year of Ref.	Year of Correction
(1)	B16.2, B16.9, B17.1, B17.8	B18.1, B18.2, B18.8	2006	1981-2005
(2)	C41.0, C41.1	C03.0	2003, 3006	2004-2005
(3)	E10, E11	E14	1996	1981-1995
(4)	D62	D64.9	1995, 2001	1996-2000
(5)	I62.0	I63.9	1992	1981-1991

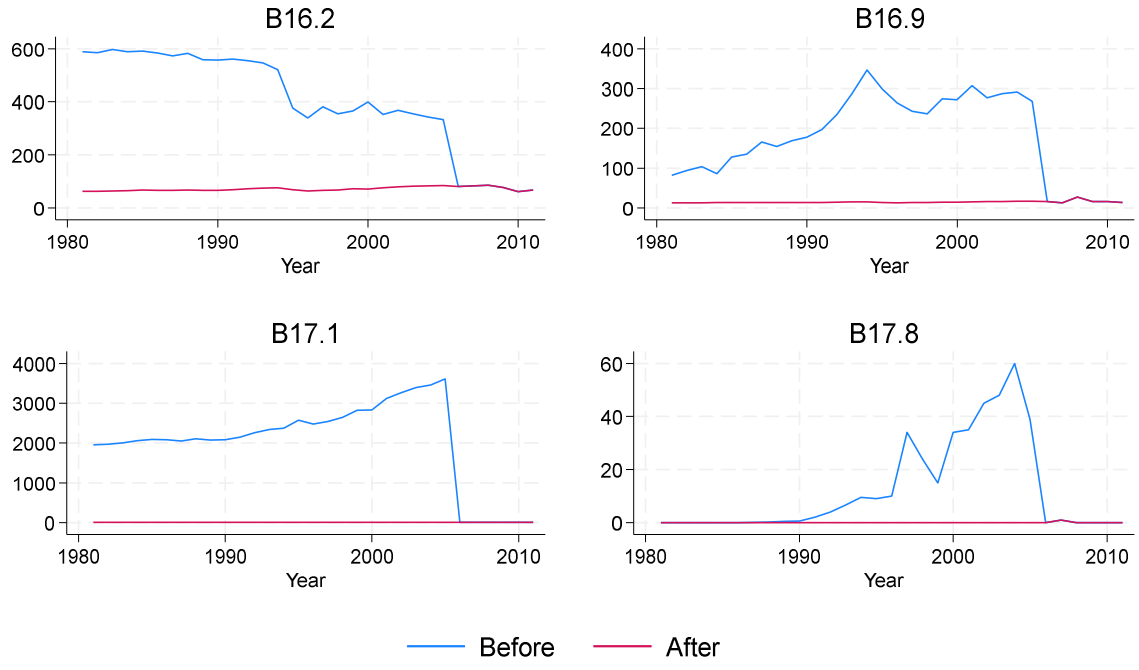
(6)	M05.3	M06.8	1997	1981-1996
(7)	X59	I50.9	1991	1981-1992

(1) Viral hepatitis

Disruptions were observed in the raw data for deaths with several codes corresponding to viral hepatitis (B16.2, B16.9, B17.1, B17.8, B18.1, B18.2, and B18.8) between 2005 and 2006 with a compensation between the codes within the larger Viral hepatitis category. We adjusted the proportion of each item in the total for each year before 2006 to match their proportions in 2006. Specifically, the following calculations were carried out:

$$\hat{d}^i_{n,x}(t) = \sum_i d^i_{n,x}(t) \cdot \frac{\sum_n d^i_{n,x}(2006)}{\sum_{n,i} d^i_{n,x}(2006)}$$

where $d^i_{n,x}(t)$ denotes deaths for year $t(\leq 2006)$, age group n, sex x and cause i.



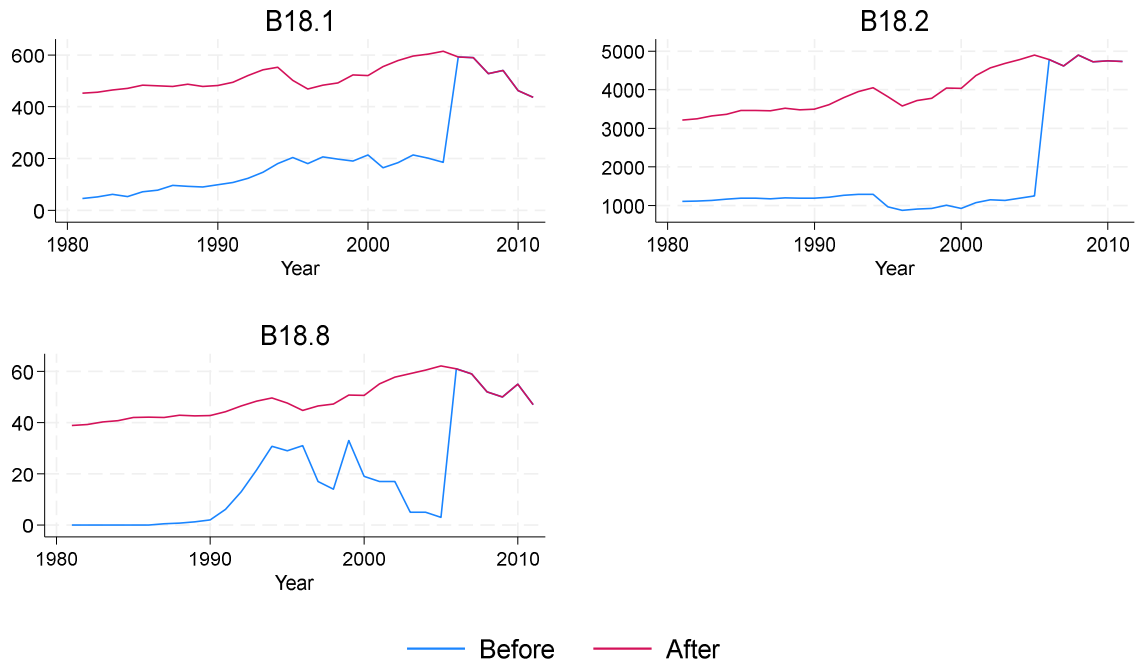


Figure 5: Results of a posteriori correction, Viral hepatitis

(2) Malignant neoplasm of the upper gum, bones of the skull and face, and mandible

Similarly, we observed a disruption for malignant neoplasms of the bones of the skull and face (C41.0), and of mandibles (C41.1) in 2004 and 2005 that is compensated by a symmetric disruption in deaths from malignant neoplasms of the upper gum (C03.0). To deal with this problem, we adjusted these values as follows:

$$\hat{d}^i_{n,x}(2004) = \sum_i d^i_{n,x}(2004) \cdot \left\{ \frac{2}{3} \cdot \frac{\sum_n d^i_{n,x}(2003)}{\sum_{n,i} d^i_{n,x}(2003)} + \frac{1}{3} \cdot \frac{\sum_n d^i_{n,x}(2006)}{\sum_{n,i} d^i_{n,x}(2006)} \right\}$$

$$\hat{d}^i_{n,x}(2005) = \sum_i d^i_{n,x}(2005) \cdot \left\{ \frac{1}{3} \cdot \frac{\sum_n d^i_{n,x}(2003)}{\sum_{n,i} d^i_{n,x}(2003)} + \frac{2}{3} \cdot \frac{\sum_n d^i_{n,x}(2006)}{\sum_{n,i} d^i_{n,x}(2006)} \right\}$$

where $d^i_{n,x}(t)$ denotes deaths for year t, age group n, sex x and cause i.

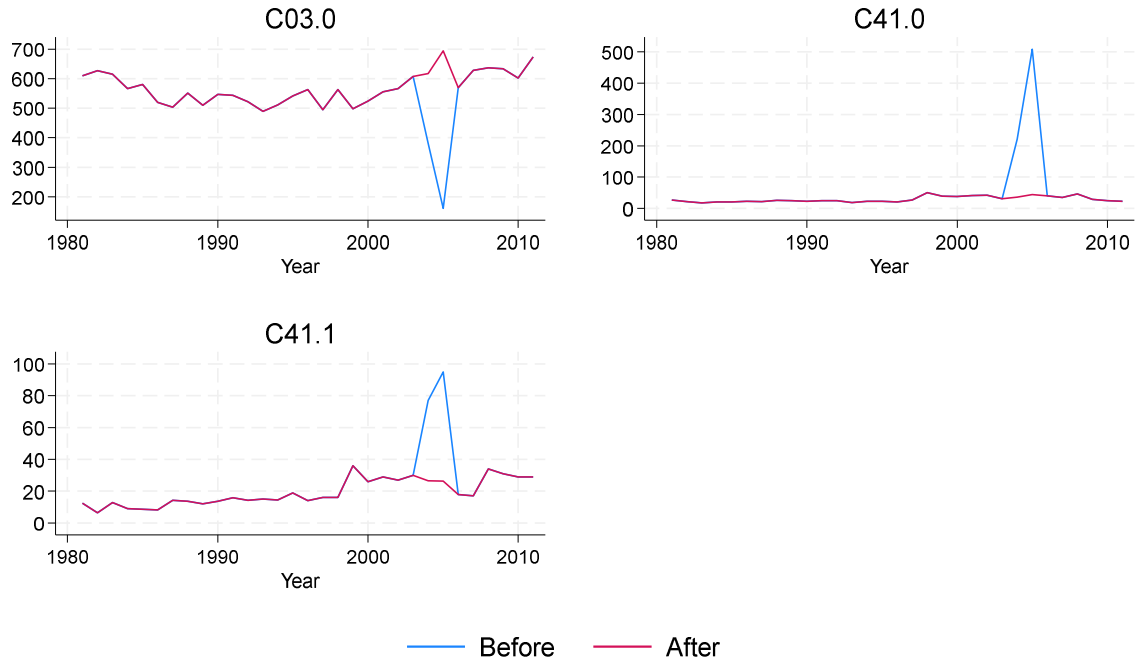
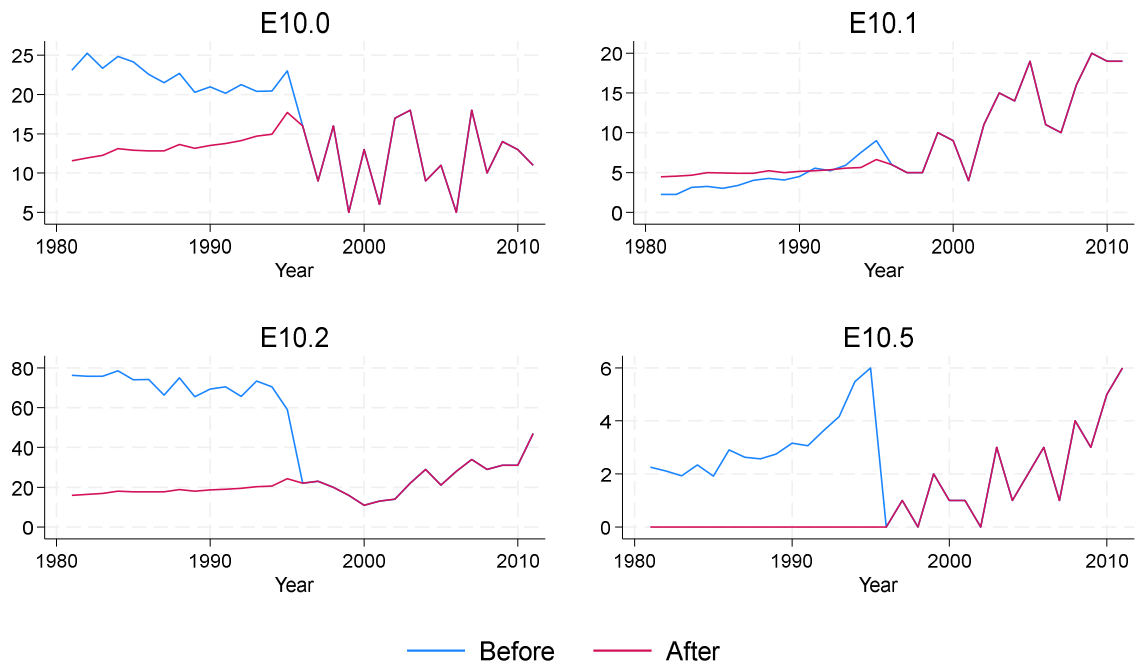
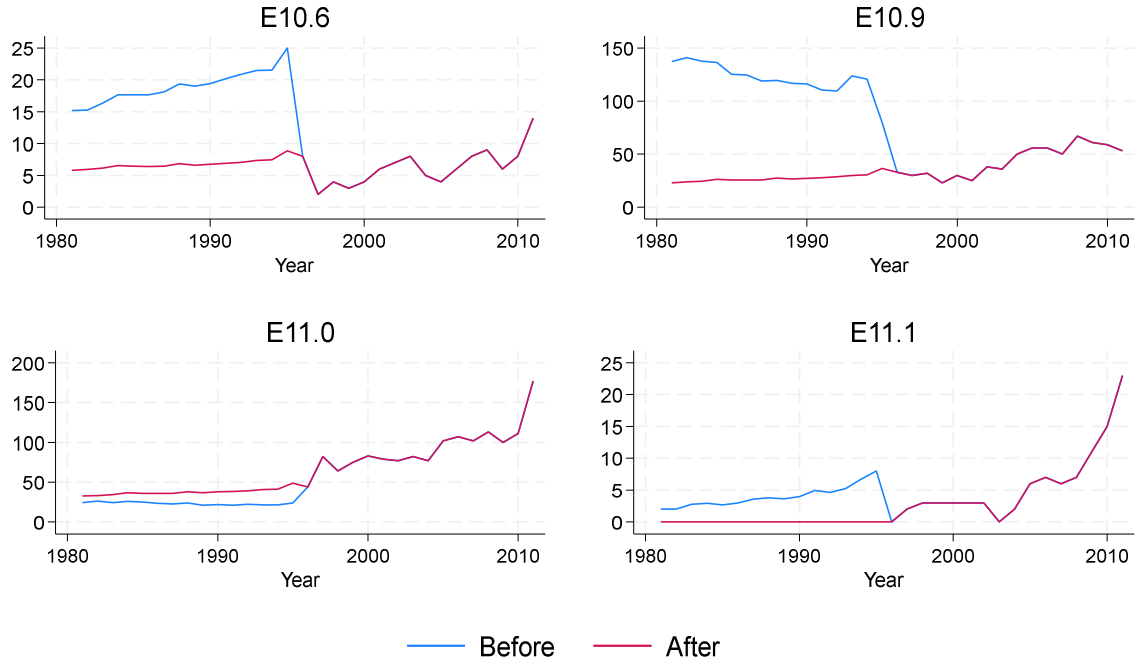


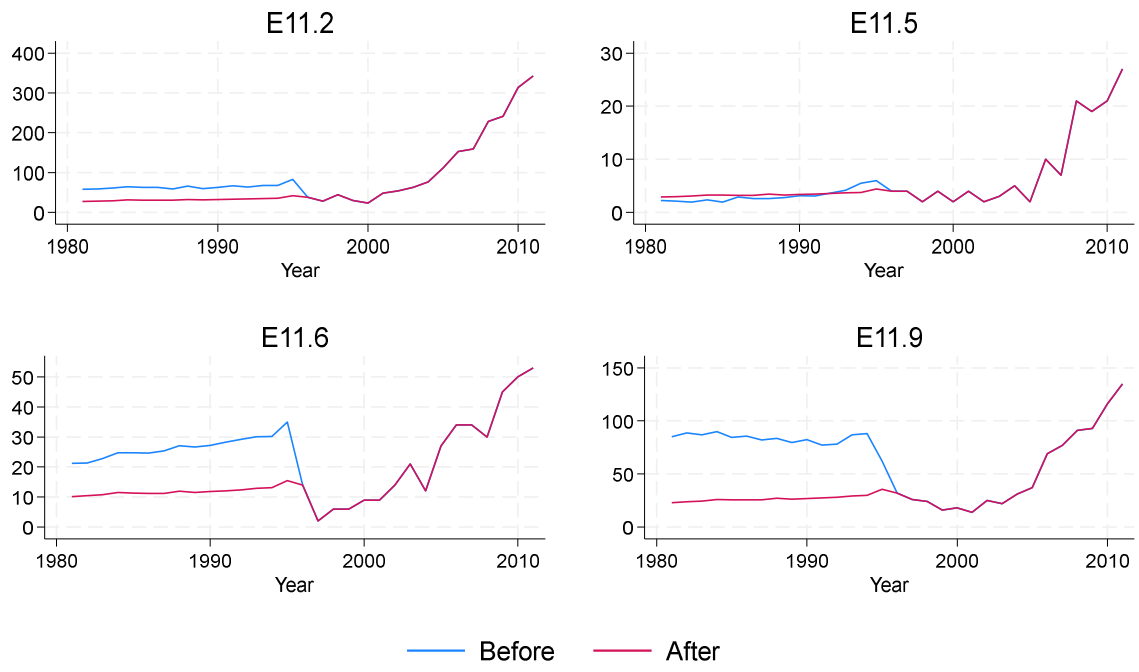
Figure 4: Results of a posteriori correction, Malignant neoplasm of the upper gum, bones of the skull and face, and mandible

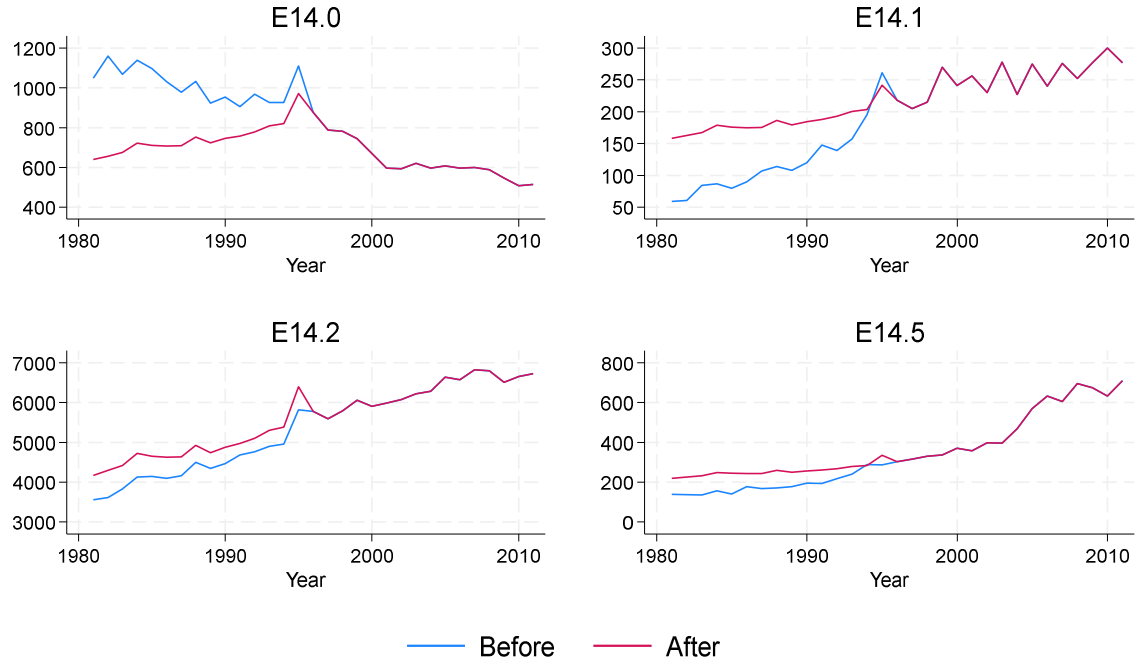
(3) Diabetes Mellitus

Compensating disruptions were also observed for deaths from Diabetes Mellitus (codes in E10, E11, and E14) between 1995 and 1996. Again, we adjusted the proportion of each item in the total of these items for each year before 1995 to match the proportion in 1996 in the same way as (1).


E10.0
E10.1
E10.2
E10.5
Year
1980
1990
2000
2010
Before
After


E10.6
E10.9
E11.0
E11.1
Year
1980
1990
2000
2010
Before
After


E11.2
E11.5
E11.6
E11.9
Year
Before
After


E14.0
E14.1
E14.2
E14.5
Year
Before
After

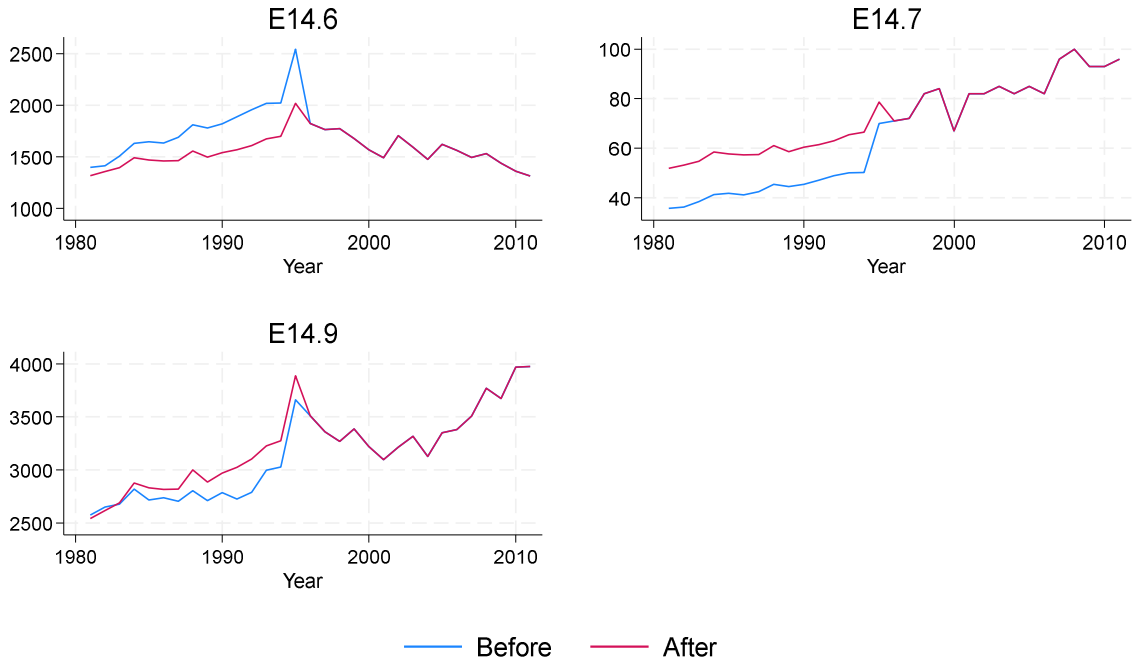


Figure 6: Results of a posteriori correction, Diabetes Mellitus

(4) Acute posthemorrhagic anemia

To account for the unplausibly high number of deaths from acute posthemorrhagic anemia (D62) in 1996-2000, we reallocated some of them to "Anemia, unspecified" (D64.9), by the following method:

$$\hat{d}^i_{n,x}(1996) = \sum_i d^i_{n,x}(1996) \cdot \left\{ \frac{5}{6} \cdot \frac{\sum_n d^i_{n,x}(1995)}{\sum_{n,i} d^i_{n,x}(1995)} + \frac{1}{6} \cdot \frac{\sum_n d^i_{n,x}(2001)}{\sum_{n,i} d^i_{n,x}(2001)} \right\}$$

$$\hat{d}^i_{n,x}(1997) = \sum_i d^i_{n,x}(1997) \cdot \left\{ \frac{4}{6} \cdot \frac{\sum_n d^i_{n,x}(1995)}{\sum_{n,i} d^i_{n,x}(1995)} + \frac{2}{6} \cdot \frac{\sum_n d^i_{n,x}(2001)}{\sum_{n,i} d^i_{n,x}(2001)} \right\}$$

$$\hat{d}^i_{n,x}(1998) = \sum_i d^i_{n,x}(1998) \cdot \left\{ \frac{3}{6} \cdot \frac{\sum_n d^i_{n,x}(1995)}{\sum_{n,i} d^i_{n,x}(1995)} + \frac{3}{6} \cdot \frac{\sum_n d^i_{n,x}(2001)}{\sum_{n,i} d^i_{n,x}(2001)} \right\}$$

$$\hat{d}^i_{n,x}(1999) = \sum_i d^i_{n,x}(1999) \cdot \left\{ \frac{2}{6} \cdot \frac{\sum_n d^i_{n,x}(1995)}{\sum_{n,i} d^i_{n,x}(1995)} + \frac{4}{6} \cdot \frac{\sum_n d^i_{n,x}(2001)}{\sum_{n,i} d^i_{n,x}(2001)} \right\}$$

$$\hat{d}^i_{n,x}(2000) = \sum_i d^i_{n,x}(2000) \cdot \left\{ \frac{1}{6} \cdot \frac{\sum_n d^i_{n,x}(1995)}{\sum_{n,i} d^i_{n,x}(1995)} + \frac{5}{6} \cdot \frac{\sum_n d^i_{n,x}(2001)}{\sum_{n,i} d^i_{n,x}(2001)} \right\}$$

where $d^i_{n,x}(t)$ denotes deaths for year t, age group n, sex x and cause i.

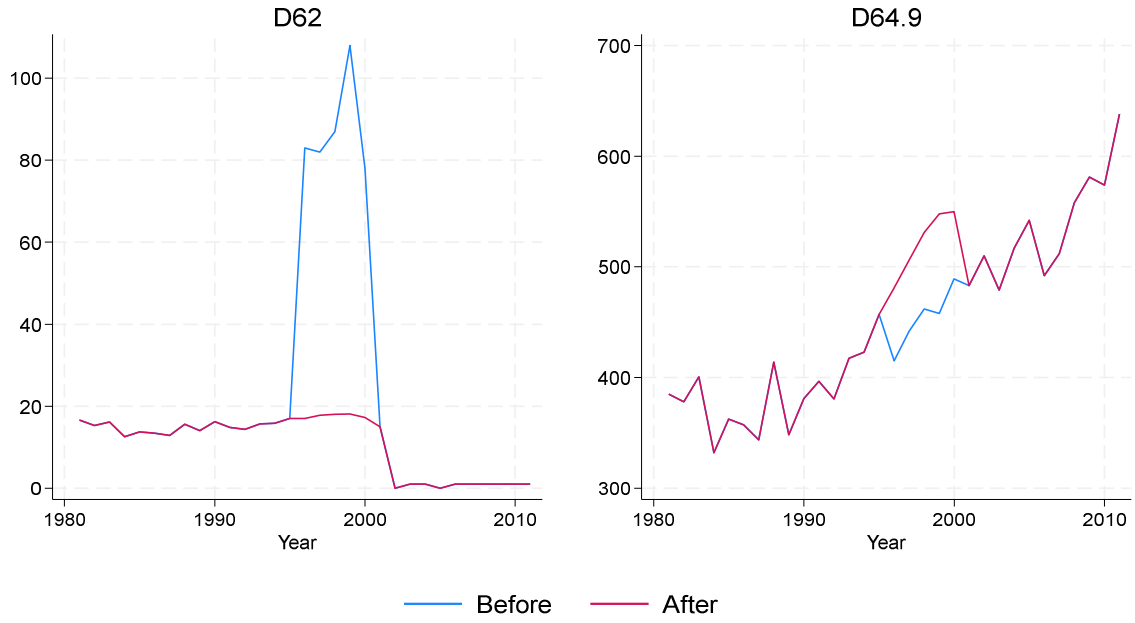


Figure 7: Results of a posteriori correction, Acute posthemorrhagic anemia

(5) Nontraumatic subdural hemorrhage and cerebral infarction

Another very large disruption was observed from 1991 to 1992 for deaths due to nontraumatic subdural haemorrhage (I62.0), which can be linked to "cerebral infarction, unspecified" (I63.9). We adjusted the proportion of each item in the total of both items for each year before 1992 to match the proportion in 1992 in the same way as (1).

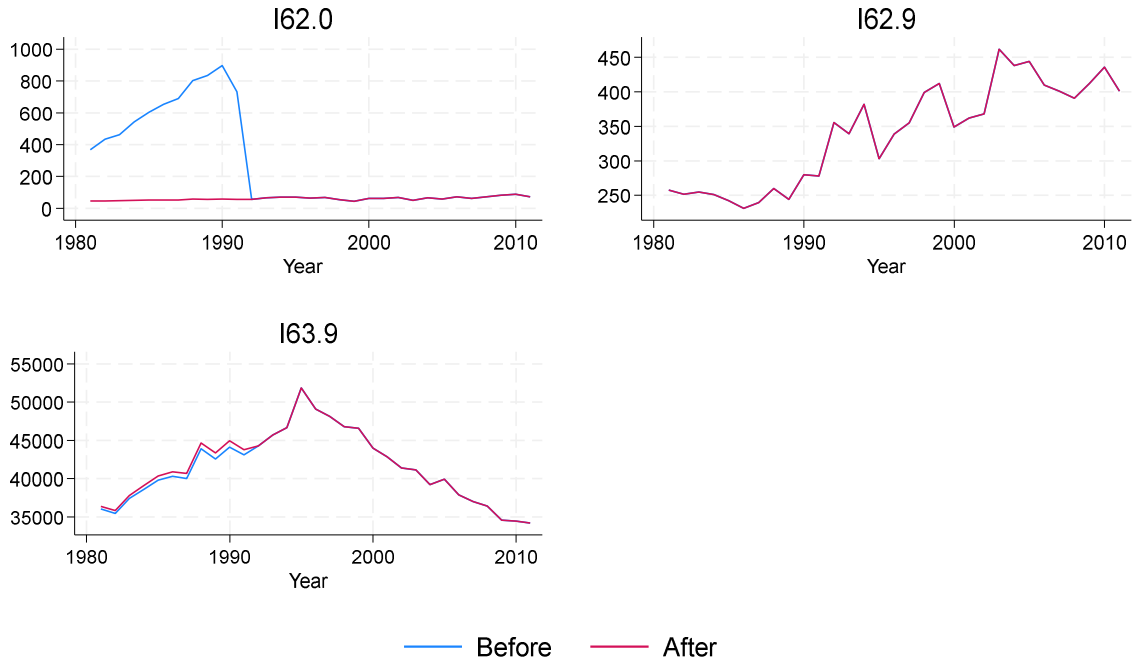


Figure 8: Results of a posteriori correction, Nontraumatic subdural haemorrhage and cerebral infarction

(6) Rheumatoid arthritis

To account for the disruption in the death counts from rheumatoid arthritis (M05.3 and M06.8) between 1996 and 1997, we adjusted the proportion of each item in the total of both items for each year before 1997 to match the proportion in 1997 in the same way as (1).

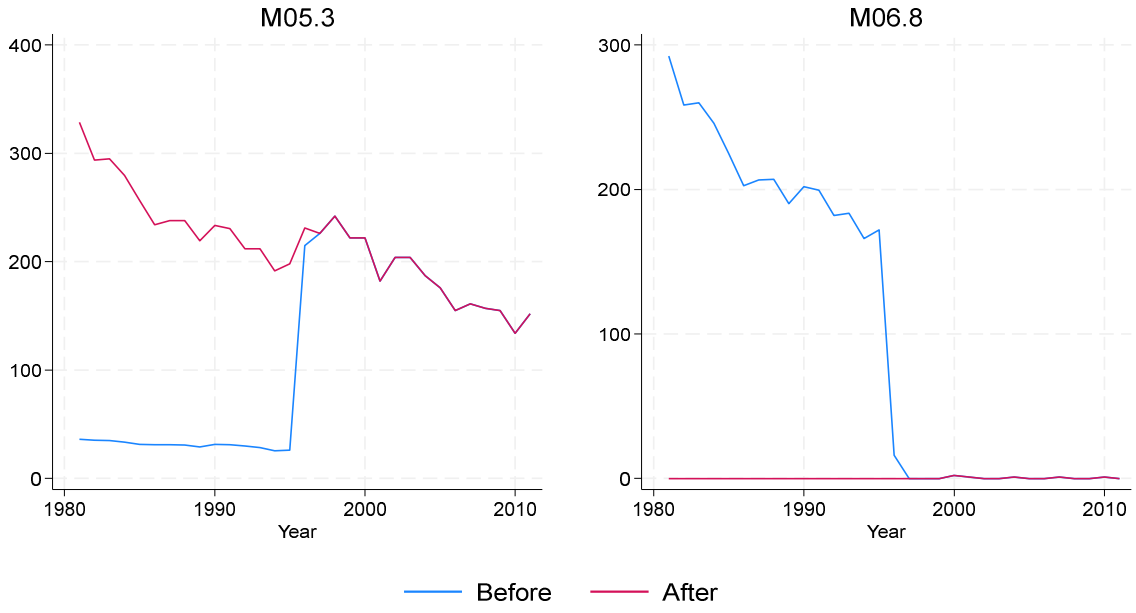
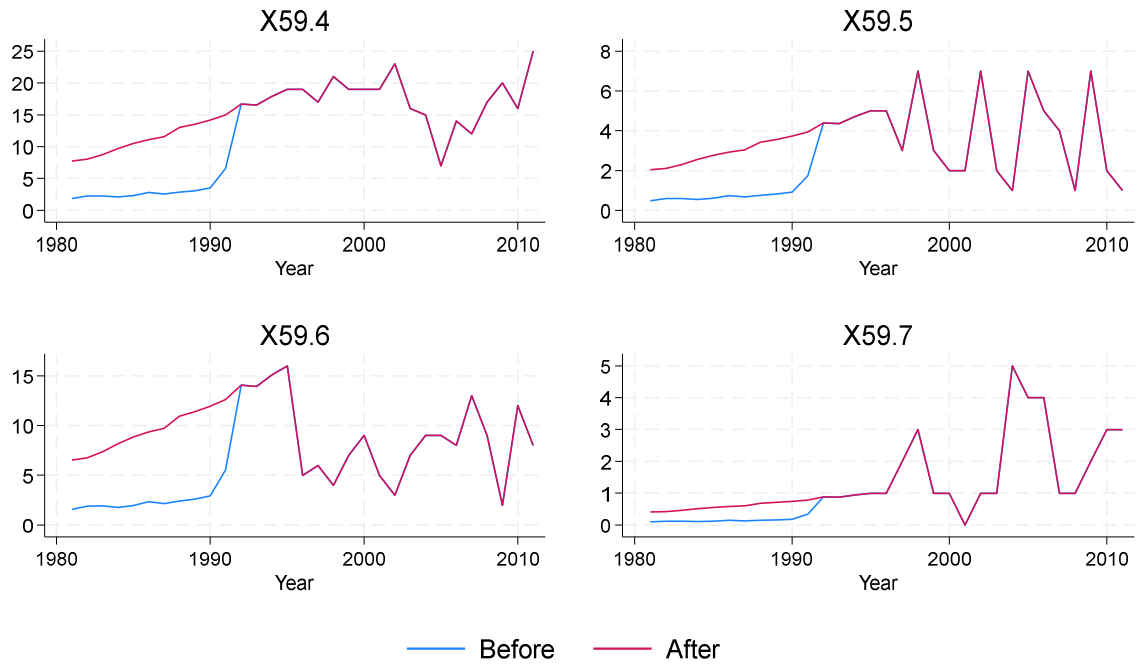
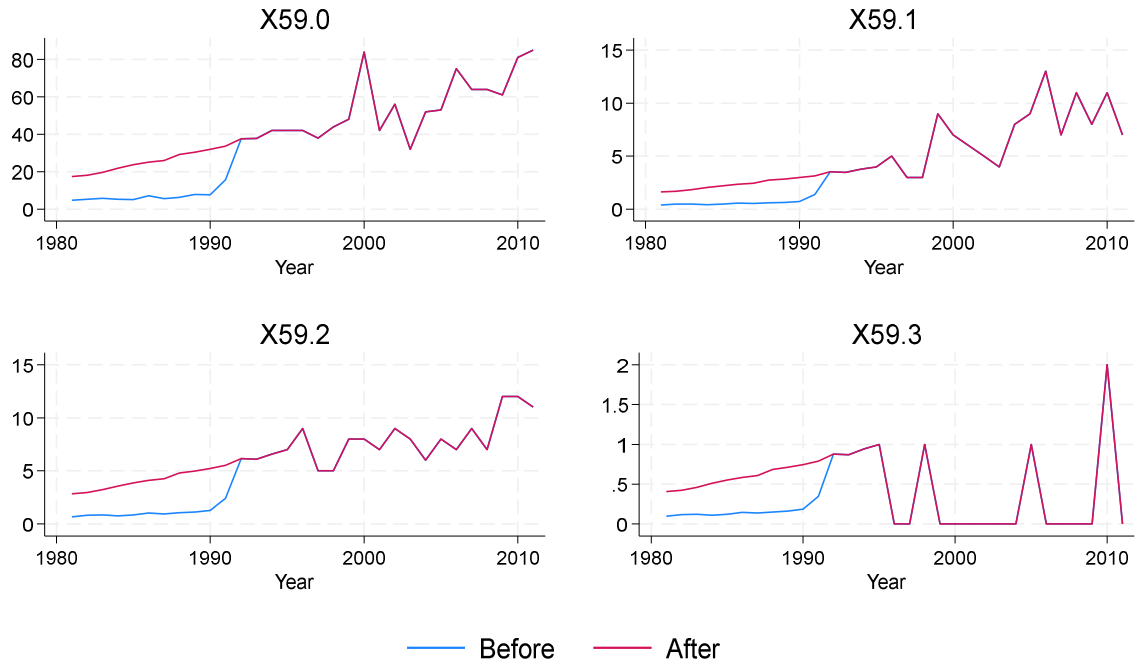


Figure 9: Results of a posteriori correction, Rheumatoid arthritis

(7) Accidental exposure to unspecified factor

A disruption was found for deaths from a series of codes for "accidental exposure to unspecified factor" (X59.0, X59.1, X59.2, X59.3, X59.4, X59.5, X59.6, X59.7, X59.8) between 1991 and 1992, which can be linked to "Heart failure, unspecified" (I59.0). We adjusted the proportion of each item in the total of these items for each year before 1992 to match the proportion in 1992 in the same way as (1).





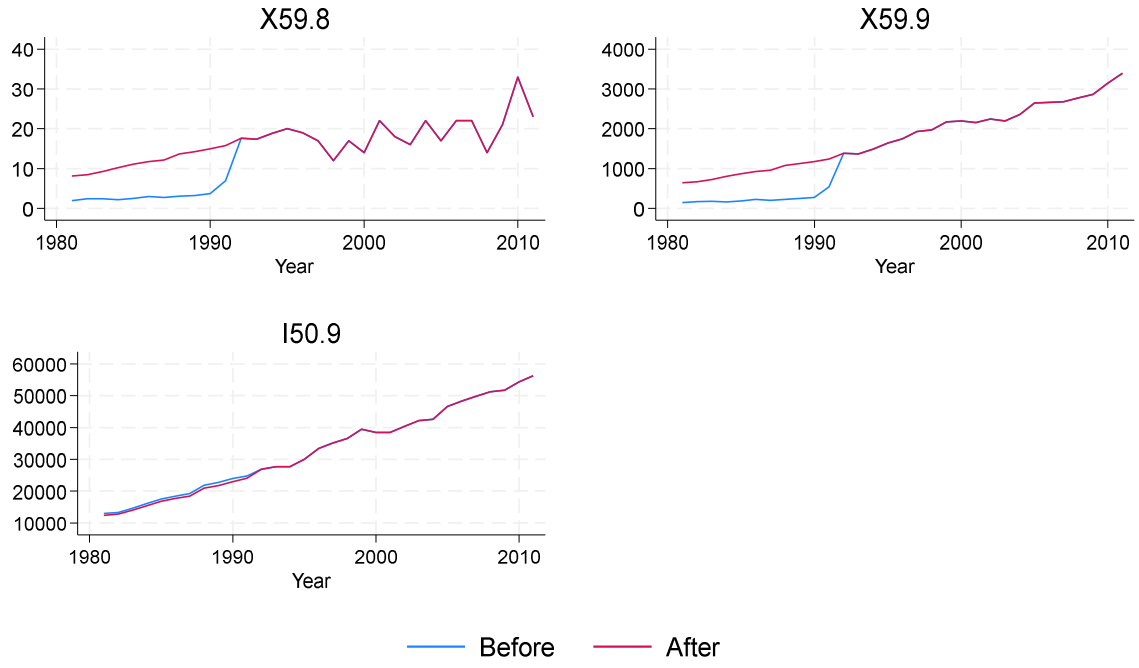


Figure 10: Results of a posteriori correction, Accidental exposure to unspecified factor

4. Redistribution of ill-defined causes

Deaths of ill-defined causes, i.e., those coded R00-R99 excluding sudden infant death syndrome (R95), were redistributed proportionally among other causes of death and accidents for the whole period for each sex and each age group. Finally, the series of them were reconstructed as shown in Figure 4.

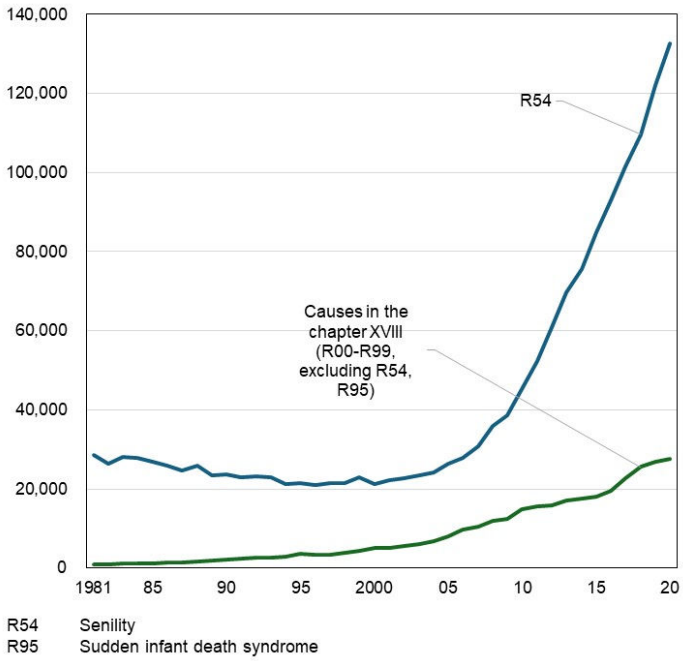


Figure 4: Results of reconstruction. Japan, ill defined-causes (Chapter XVIII, R00-R99)

References

1. Ohtsu Y., Korekawa Y., Ishii F., Pechholdová M., Meslé F., & Vallin J. (2018) *Towards the Reconstruction of Long-term Time Series Statistics on Causes of Death in Japan: How to Remove the Influence of a New Death Certificate in 1995.* Journal of population problems, 74(2), pp. 99-117. [in Japanese]
2. Pechholdova M., C., Meslé F., and Vallin J. (2017), Reconstructing Long-Term Coherent Cause-of-Death Series, a Necessary Step for Analyzing Trends, European Journal of Population, 33(5), 629-650.

Appendixes

This background information is relative to cause-of-death data for the period 1950-2016. For general information about mortality statistics in Japan, the reader should refer to the general HMD Background and Documentation file [https://mortality.org/File/GetDocument/hmd.v6/JPN/Public/InputDB/JPNcom.pdf].

Appendix 1. Data used for the cause-of-death database

Period	Type of Data	Age Grouping	Comments	RefCode(s)†
1950-1994	Annual number of deaths, by sex, five-year age group and medical cause-of-death.	0, 1, 2, 3, 4, 5-9, …, 100+, UNK		30
1995-2018	Annual number of deaths, by sex, age group, and medical cause-of-death coded to the 4th digit of the ICD.	0, 1-4, 5-9, ...,90-94, 95+,unknown		50

† The reference code is used in the raw data files (Input Database) to link data with sources.

Appendix 2. Differences between the all-cause HMD series and the cause-specific series in the raw data

Year	HMD	COD	HMD-COD
1950	904876	904876	0
1951	838998	838998	0
1952	765068	765068	0
1953	772547	772547	0
1954	721491	721491	0
1955	693523	693523	0
1956	724460	724460	0
1957	752445	752445	0
1958	684189	684189	0
1959	689932	689932	0
1960	706599	706599	0
1961	695644	695701	-57
1962	710265	710265	0
1963	670770	670770	0
1964	673067	673067	0
1965	700438	700437	1

Year	HMD	COD	HMD-COD
1985	752283	752283	0
1986	750620	750620	0
1987	751172	751172	0
1988	793014	793014	0
1989	788594	788594	0
1990	820305	820305	0
1991	829797	829797	0
1992	856643	856643	0
1993	878532	878532	0
1994	875933	875933	0
1995	922139	922139	0
1996	896211	896211	0
1997	913402	913402	0
1998	936484	936484	0
1999	982031	982031	0
2000	961653	961653	0

1966	670342	670342	0
1967	675006	675006	0
1968	686555	686555	0
1969	693787	693787	0
1970	712962	712962	0
1971	684521	684521	0
1972	683751	683751	0
1973	709416	709416	0
1974	710510	710510	0
1975	702275	702275	0
1976	703270	703270	0
1977	690074	690074	0
1978	695821	695821	0
1979	689664	689664	0
1980	722801	722801	0
1981	720262	720262	0
1982	711883	711883	0
1983	740038	740038	0
1984	740247	740247	0

2001	969663	970331	-668
2002	981744	982379	-635
2003	1014296	1014951	-655
2004	1028019	1028602	-583
2005	1083236	1083796	-560
2006	1084100	1084450	-350
2007	1108089	1108334	-245
2008	1142831	1142407	424
2009	1142871	1141865	1006
2010	1198014	1197012	1002
2011	1254350	1253066	1284
2012	1257406	1256359	1047
2013	1269391	1268436	955
2014	1274017	1273004	1013
2015	1291457	1290444	1013
2016	1308880	1307748	1132
2017	1341543	1340397	1146
2018	1363485	1362470	1015

Appendix 3. Detected total cases of death with no age specification

Year	Cause	Sex	Deaths
1995	W656	1	1
1995	X005	2	1
1995	X998	2	1
1995	Y114	2	1
1995	Y335	1	1
1996	Y266	2	1
1997	X806	2	1
1997	X845	1	1
1998	Y298	1	1
1999	X532	1	1
1999	X925	2	1
2000	W128	1	1
2000	X312	2	1
2000	X532	1	1
2000	Y339	2	1

Appendix 4. Recoding of non-UCD codes

Original cuse	Target cause	Type of original cause	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
A91_	A979	obsolete											1									
B485	B488	obsolete																				
C832	C839	obsolete	3	5	7	3	2	5	2	2	2	1	1									
C834	C839	obsolete	7	13	9	3	11	9	4	10	12	13	8	17	16	16	21	44	41			
C836	C839	obsolete						1			2			2	1			2				
C850	C859	obsolete	15	17	11	10	9	7	4	4	2	7	1	1	1	6		2	1			
C881	C887	obsolete										1										
C912	C919	obsolete		1	1	3		3	1		1	3	4	2	2	2	1	1				
C945	C947	obsolete	6	7	6	12	7	9	5	1	3	6	4	4	4	5	3	3	1			
C961	C969	obsolete	104	89	70	80	59	51	44	45	41	37	33	39	38	27	31	34	32			
C963	C969	obsolete		1																		
D463	D469	obsolete				1	1	1	2	1				1								
D752	D759	obsolete	14	15	10	9	20	12	9	21	11	11	12	19	20	21	22	15	18			
D760	D763	obsolete	15	7	13	12	4	5	10	6	10	6	6	6	10	5	8	7	7			
F100	X45_	non-UCD	246	196	196	187	169	154	135	166	160	143	138									
F130	X41_	non-UCD										1										
F150	X41_	non-UCD						1														
F180	X46_	non-UCD		1																		
F190	X40_	non-UCD		1																		
G903	G909	obsolete	80	91	104	84	100	75	118	76	90	101	104	92	81	72	66	58	48	35	40	35
H547	H549	obsolete					2										1					
I150	I139	non-UCD																				
I220	I212	non-UCD	14	8	7	9	8	11	13	14	5	16	12	9	5	4	8	7	4	5	4	3
I221	I212	non-UCD	5	3	1	7	1		6	5	2	1		2	1			1	1	3	1	
I228	I212	non-UCD	4	3	3	6	8	11	10	10	11	8	8	6	7	12	10	6	8	6	5	3

I229	I212	non-UCD	596	541	535	557	472	472	466	431	415	383	378	391	333	363	345	280	253	209	179	161
I252	I258	non-UCD	3232	3376	3561	3723	3904	4090	4129	4199	4200	4298	4330	4338	4195	4336	4255	4452	4384			
K350	K358	obsolete	80	50	49	64	79	51	53	53	51	54	66	40	54	37	39	37	51			
K351	K358	obsolete	1	5	7	4	10	8	9	12	6	4	9	7	10	10	7	13	8			
K359	K358	obsolete	8	9	9	11	7	9	1	9	11	8	13	13	14	21	26	31	19			
K588	K580	obsolete																				
M725	M729	obsolete	10	7	14	16	35	46	58	67	76	93	91									
O60_	O600	obsolete										1										
O96_	O969	obsolete	3	4	2		2	2		2		3		1	2							
P704	P969	non-UCD	2	1		1			1	1	1		1									
P708	P969	non-UCD			1																	
P711	P969	non-UCD				1																
P728	P969	non-UCD									1											
P740	P969	non-UCD	5		1	1	1	3	1	4	5	4	2									
P741	P969	non-UCD	1		1			1														
P742	P969	non-UCD	1		1	1		1														
P743	P969	non-UCD	4	7	13	12	4	6	12	9	10	4	6									
P748	P969	non-UCD	3									1										
P749	P969	non-UCD	3				2															
Q314	Q319	obsolete	1	1				1				1										
R179	R17_	error																				
R501	R509	obsolete						1			1											
R659	A419	non-UCD																				
Y85_	Y850	obsolete	127	103	93	100	81	79	94	103	82	71	83	102	106	101	107	90	79			