

Reconstruction of continuous time series of deaths by cause



Markéta Pechholdová

INED / University of Economics in Prague

Lithuanian Social Research Centre

Teaching seminar „Causes of death: data and methods of analysis“

February 23-24, 2015 Vilnius

pechholdova@gmail.com



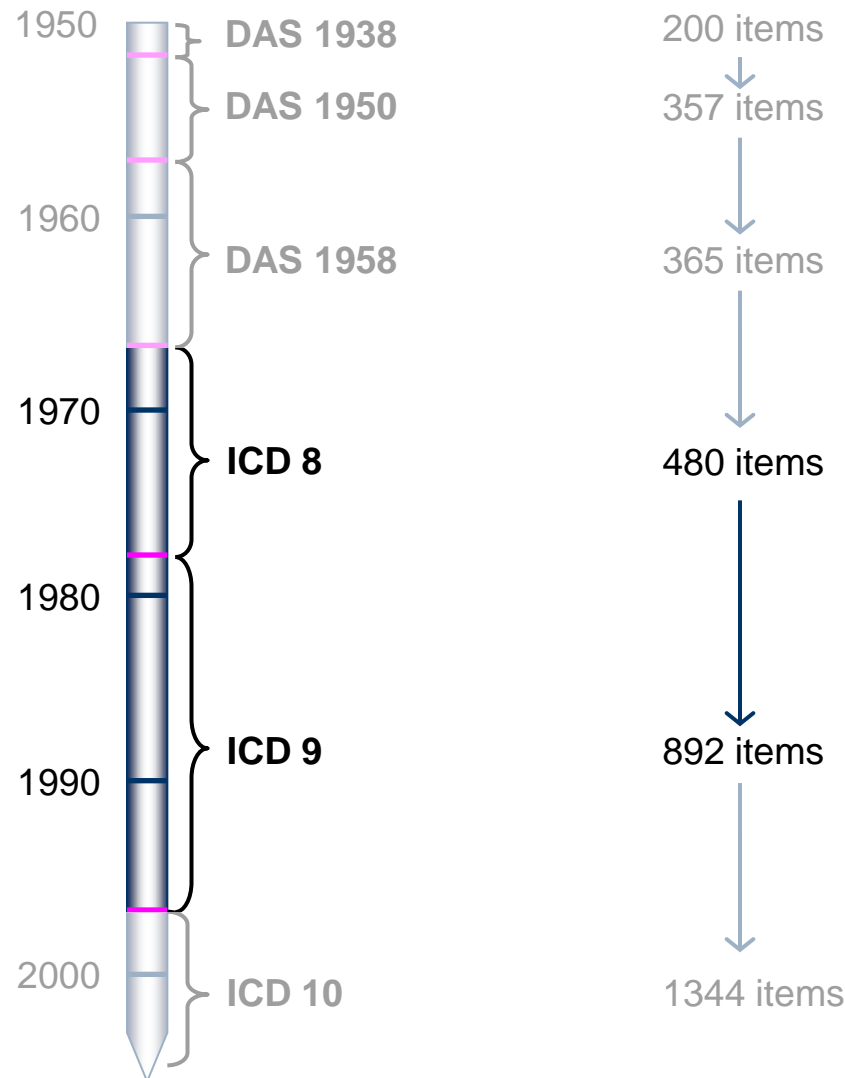
The research is funded by EU structural assistance (European Social Fund) to Lithuania under the measure VP-1-3.1-ŠMM-07-K “Support to Research Activities of Scientists and Other Researchers (Global Grant)” project Nr. VP-1-3.1-ŠMM-07-K-02-067

RESEARCH CONTEXT

- **Periodical changes in International Classification of Diseases (ICD) disrupt the statistical trends in mortality by cause of death.**
- **For a long-term trends analysis, data reconstruction is required.**
- **MPIDR in collaboration with INED conducted a reconstruction project based upon the methodology developed by F. Meslé and J. Vallin**
- **Reconstruction already done / in progress in:**
 - ➔ **France, Russia, Ukraine, Baltic countries, Caucasian countries, Czech Republic, East Germany, Poland, Moldova, Belarus, Spain, Japan, US, UK**
 - ➔ **Netherlands**

CLASSIFICATION HISTORY IN FRG

Since 1950 6 classifications were used in West Germany:



West Germany didn't publish the full detail of 3-digit ICD8, some causes were grouped.

Some ICD8 4-digit items were published.

ICD9 was published with 3-digit detail.

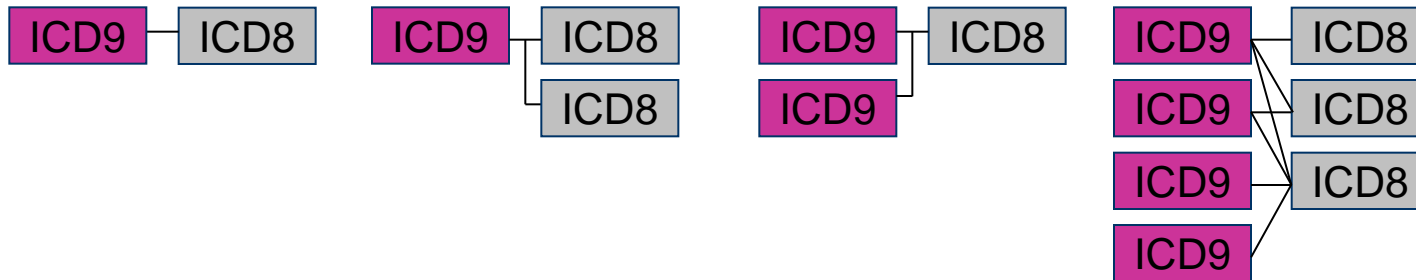
5-year age groups up to 90+

BASIC PRINCIPLES

- **Double classification of deaths by cause at the year of ICD change is not available for West Germany → we produce such double classification *a posteriori***
- **Method is based on the most possibly detailed comparison of the medical content of the items of the two successive revisions**
- **OUTPUT: transition coefficients to redistribute the deaths according to another revision**

DETAILED STEPS

- **Correspondence tables – mutual correspondences between two ICD revisions**
- **Elementary associations – the smallest possible clusters of deaths of the same medical content within two successive ICD revisions**



- **Transition coefficients – derived from each elementary association**
- **Redistribution**
- **Corrections – by sex and age, corrections *a posteriori***

THE TRANSITION FROM ICD9 TO ICD8 IN WEST GERMANY

CORRESPONDENCE TABLE ICD9-ICD8

ICD9		corresponds to	ICD8	
Code	Diagnosis		Code	Diagnosis
402	Hypertensive heart disease	→	402	Hypertensive heart disease
		→	4001	Malignant hypertension with heart involvement
		→	412	Chronic ischemic heart disease
403	Hypertensive renal disease	→	403	Hypertensive renal disease
		→	4003	Malignant hypertension with renal involvement
410	Acute myocardial infarction	→	410	Acute myocardial infarction
412	Old myocardial infarction	→	412	Chronic ischemic heart disease
		→	414	Asymptomatic ischemic heart disease
		→	429	Ill-defined heart disease
413	Angina pectoris	→	413	Angina pectoris
414	Other forms of chronic ischemic heart disease	→	412	Chronic ischemic heart disease
		→	425	Cardiomyopathy
425	Cardiomyopathy	→	425	Cardiomyopathy
		→	429	Ill-defined heart disease
429	Ill-defined descriptions and complications of heart disease	→	412	Chronic ischemic heart disease
		→	429	Ill-defined heart disease
442	Other aneurysm	→	44a	Remaining diseases of arteries, arterioles and capillaries
443	Other peripheral vascular disease	→	44a	Remaining diseases of arteries, arterioles and capillaries
446	Polyarteritis nodosa and allied conditions	→	44a	Remaining diseases of arteries, arterioles and capillaries
447	Other disorders of arteries and arterioles	→	44a	Remaining diseases of arteries, arterioles and capillaries
448	Disease of capillaries	→	44a	Remaining diseases of arteries, arterioles and capillaries



CORRESPONDENCE TABLE ICD8-ICD9

ICD8		corresponds to	ICD9	
Code	Diagnosis		Code	Diagnosis
4001	Malignant hypertension with heart involvement	→	402	Hypertensive heart disease
4003	Malignant hypertension with renal involvement	→	403	Hypertensive renal disease
402	Hypertensive heart disease	→	402	Hypertensive heart disease
403	Hypertensive renal disease	→	403	Hypertensive renal disease
410	Acute myocardial infarction	→	410	Acute myocardial infarction
412	Chronic ischemic heart disease	→	402	Hypertensive heart disease
		→	412	Old myocardial infarction
		→	414	Other forms of chronic ischemic heart disease
		→	429	Ill-defined descriptions and complications of heart disease
413	Angina pectoris	→	413	Angina pectoris
414	Asymptomatic ischemic heart disease	→	412	Old myocardial infarction
425	Cardiomyopathy	→	414	Other forms of chronic ischemic heart disease
		→	425	Cardiomyopathy
429	Ill-defined heart disease	→	412	Old myocardial infarction
		→	425	Cardiomyopathy
		→	429	Ill-defined descriptions and complications of heart disease
44a	Remaining diseases of arteries, arterioles and capillaries	→	442	Other aneurysm
		→	443	Other peripheral vascular disease
		→	446	Polyarteritis nodosa and allied conditions
		→	447	Other disorders of arteries and arterioles
		→	448	Disease of capillaries

1:1 ASSOCIATION

1. Elementary associations

ICD 9	Diagnosis	1979	1978	ICD8	Portion	Diagnosis
410	Acute myocardial infarction	81121	79347	410	T	Acute myocardial infarction
	Sum	81121	79347			

ICD 9	Diagnosis	1979	1978	ICD8	Portion	Diagnosis
413	Angina pectoris	1081	1024	413	T	Angina pectoris
	Sum	1081	1024			

2. Transition coefficients

ICD9	ICD8	Coefficient
410	410	1.000

ICD9	ICD8	Coefficient
413	413	1.000

1:N ASSOCIATION

1. Elementary associations

ICD 9	Diagnosis	1979	1978	ICD8	Portion	Diagnosis
403	Hypertensive renal disease	1222	1044	403	T	Hypertensive renal disease
			96	4003	T	Malignant hypertension with renal involvement
	Sum	1222	1140			

2. Transition coefficients

ICD9	ICD8	Coefficient
403	403	1.000
403	4003	1.000

N:1 ASSOCIATION

1. Elementary associations

ICD 9	Diagnosis	1979	1978	ICD8	Portion	Diagnosis
442	Other aneurysm	97	1100	44a	P	Remaining diseases of arteries, arterioles and capillaries
443	Other peripheral vascular disease	935		44a	P	(442, 443, 446-448)
446	Polyarteritis nodosa and allied condition	78		44a	P	
447	Other disorders of arteries and arteriole	60		44a	P	
448	Disease of capillaries	18		44a	P	
	Sum	1188	1100			

2. Transition coefficients

ICD9	ICD8	Coefficient
442	44a	0.082
443	44a	0.787
446	44a	0.066
447	44a	0.050
448	44a	0.015

} $\Sigma = 1$

COMPLEX ASSOCIATION (N:N)

1. Elementary associations

ICD 9	Diagnosis	1979	1978	ICD8	Portion	Diagnosis
402	Hypertensive heart disease	10693	9042 121	402 4001	T T	Hypertensive heart disease Malignant hypertension with heart involvement
			55126	412	P	Chronic ischemic heart disease
412	Old myocardial infarction	1301	2	412 414	P T	Asymptomatic ischemic heart disease
			1507	429	P	Ill-defined heart disease
414	Other forms of chronic ischemic heart disease	37473		412	P	
			378	425	P	Cardiomyopathy
429	Ill-defined descriptions and complications of heart disease	16861		412	P	
				429	P	
425	Cardiomyopathy	1036		425 429	P P	
	Sum	67364	66176			

COMPLEX ASSOCIATION (N:N)

1. Double classification (cross-table)

Items of ICD9	Items of ICD8						Deaths in 1979
	402	4001	412	414	425	429	
402							
412							
414							
425							
429							
1979 (estimate)							
1978 (real)							

COMPLEX ASSOCIATION (N:N)

2. Transition coefficients

Items of ICD9	Items of ICD8					
	402	4001	412	414	425	429
402						
412						
414						
425						
429						
Sum	1.000	1.000	1.000	1.000	1.000	1.000

ICD9	ICD8	Coefficient
402	402	1.000
402	412	0.024
412	412	0.020
414	412	0.663
429	412	0.292
412	414	1.000
414	425	0.663
425	425	0.337
412	429	0.117
425	429	0.591
429	429	0.292
402	4001	1.000

$\Sigma = 1$

$\Sigma = 1$

$\Sigma = 1$

HOW THE COEFFICIENTS WORK

Redistribution equations: ICD8 into ICD9

ICD9	ICD8	Coefficient
402	402	1.000
402	412	0.024
402	4001	1.000
403	403	1.000
403	4003	1.000
410	410	1.000
412	412	0.020
412	414	1.000
412	429	0.117
413	413	1.000
414	412	0.663
414	425	0.663
425	425	0.337
425	429	0.591
429	412	0.292
429	429	0.292
442	44a	0.082
443	44a	0.787
446	44a	0.066
447	44a	0.050
448	44a	0.015

$$D(410)^{ICD9} = 1.000 * D(410)^{ICD8}$$

$$D(413)^{ICD9} = 1.000 * D(413)^{ICD8}$$

$$D(403)^{ICD9} = 1.000 * D(403)^{ICD8} + D(4003)^{ICD9}$$

$$D(442)^{ICD9} = 0.082 * D(44a)^{ICD8}$$

$$D(443)^{ICD9} = 0.787 * D(44a)^{ICD8}$$

$$D(446)^{ICD9} = 0.066 * D(44a)^{ICD8}$$

$$D(447)^{ICD9} = 0.050 * D(44a)^{ICD8}$$

$$D(448)^{ICD9} = 0.015 * D(44a)^{ICD8}$$

$$D(402)^{ICD9} = 1.000 * D(402)^{ICD8} + 0.024 * D(412)^{ICD8} + 1.000 * D(4001)^{ICD8}$$

$$D(412)^{ICD9} = 0.020 * D(412)^{ICD8} + 1.000 * D(414)^{ICD8} + 0.117 * D(429)^{ICD8}$$

$$D(414)^{ICD9} = 0.663 * D(412)^{ICD8} + 0.663 * D(425)^{ICD8}$$

$$D(425)^{ICD9} = 0.337 * D(425)^{ICD8} + 0.591 * D(429)^{ICD8}$$

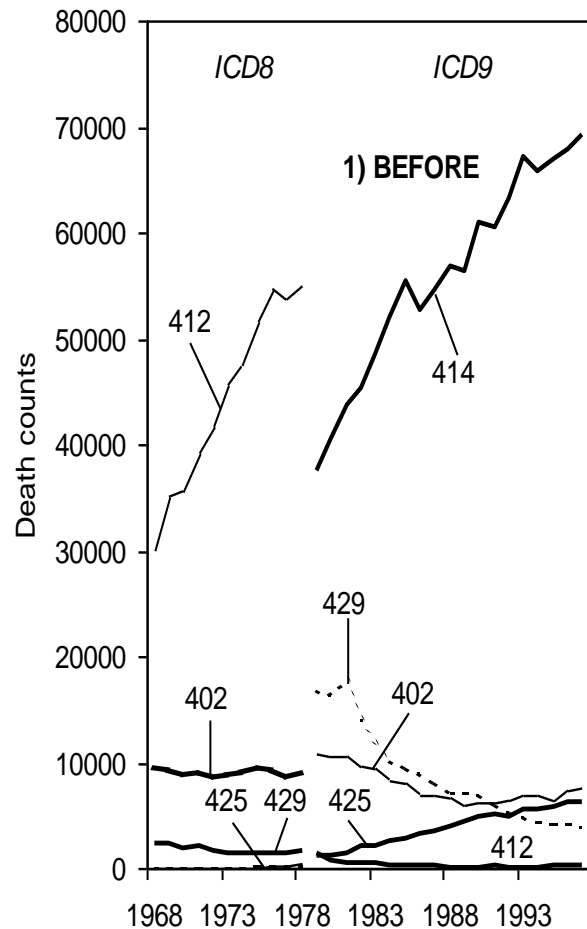
$$D(429)^{ICD9} = 0.292 * D(412)^{ICD8} + 0.292 * D(429)^{ICD8}$$

Redistribution result (year 1978)

ICD9 code	Deaths in 1979	Deaths of 1978 redistributed into ICD8	ICD8 code	Deaths in 1978
402	10693	?	402	9042
412	1301		4001	121
414	37473		412	55126
425	1036		414	2
429	16861		429	1507
			425	378
Sum	67364	66176		66176

COMPLEX ASSOCIATION (N:N)

Before and after reconstruction



ICD8:

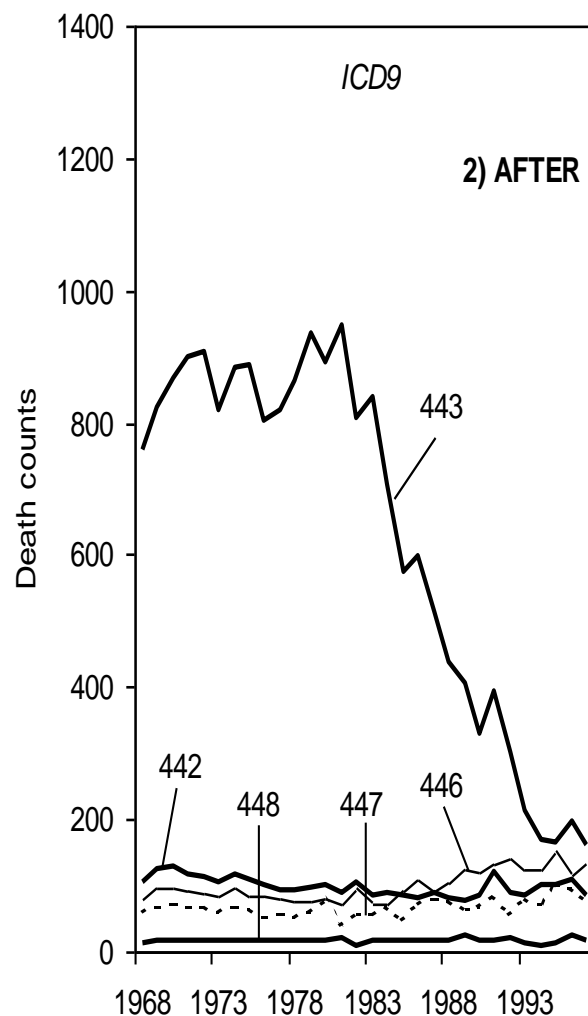
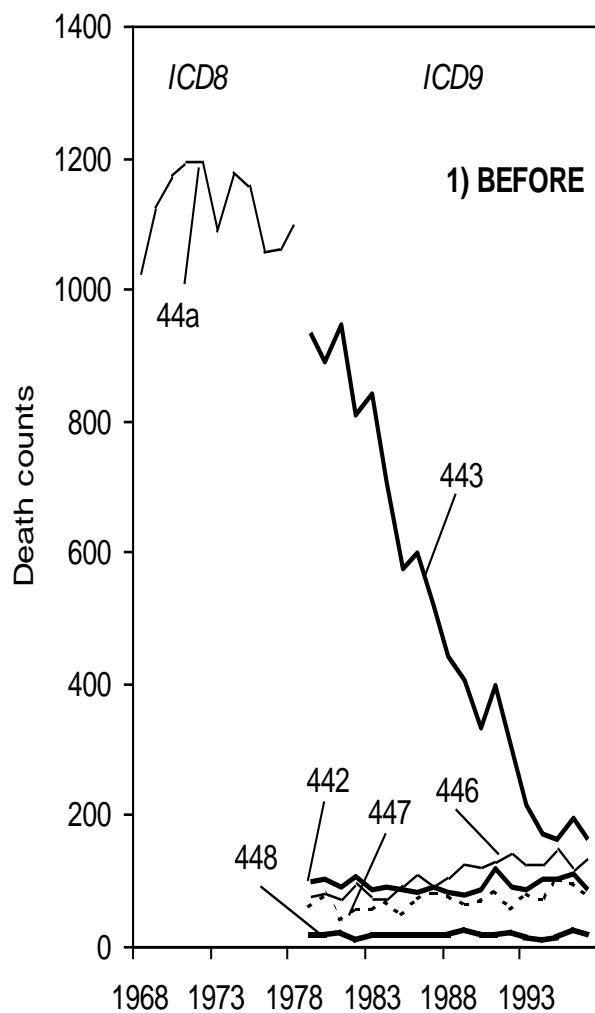
- 402 Hypertensive heart disease
- 4001 Malignant hypertension with heart involvement
- 412 Chronic ischemic heart disease
- 414 Asymptomatic ischemic heart disease
- 425 Cardiomyopathy
- 429 Ill-defined heart disease

ICD9:

- 414 Other forms of ischemic heart disease
- 429 Ill-defined heart disease
- 402 Hypertensive heart disease
- 425 Cardiomyopathy
- 412 Old myocardial infarction

SPLITTING (1:N)

Before and after reconstruction



ICD8:

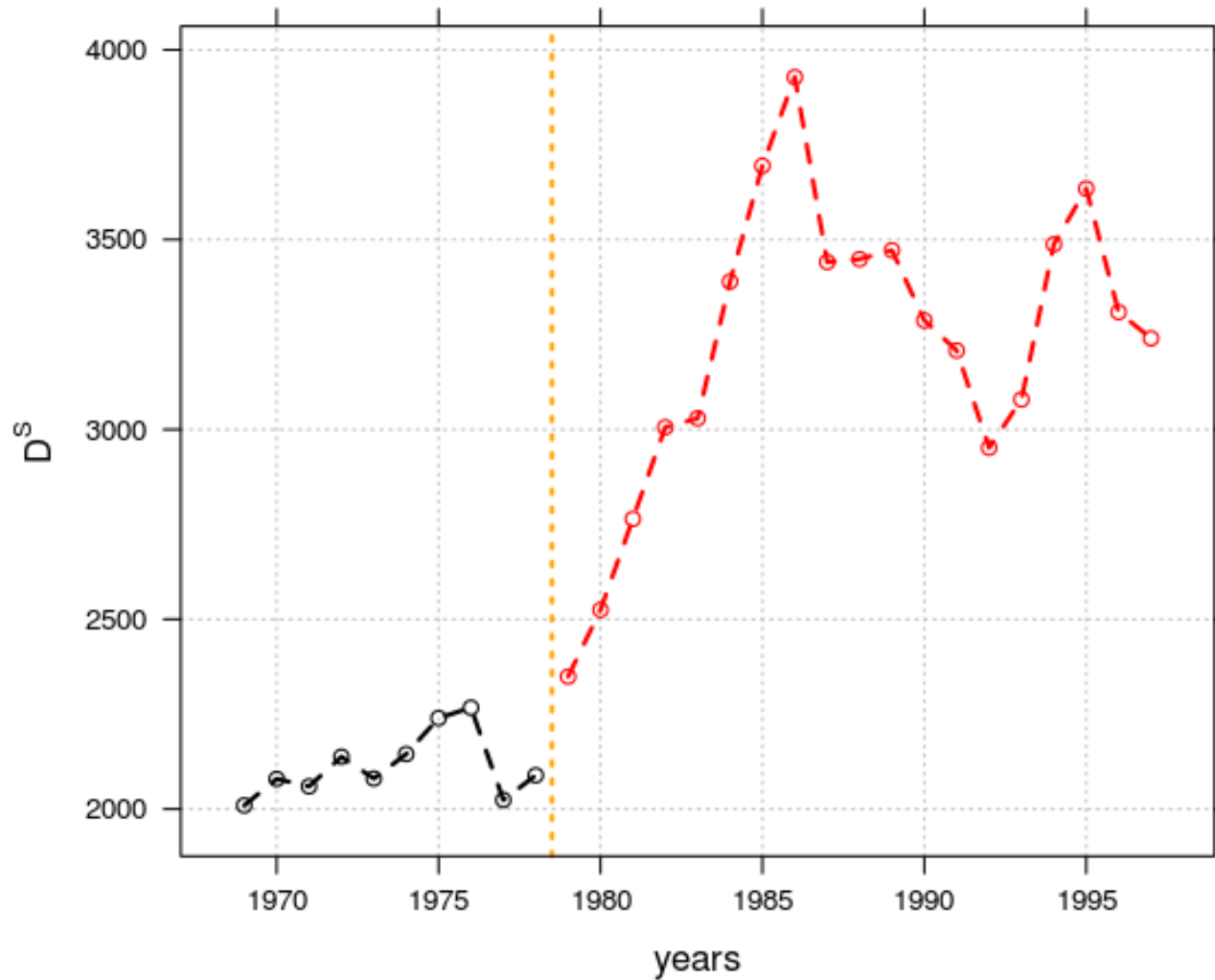
44a Remaining diseases of arteries, arterioles and capillaries

ICD9:

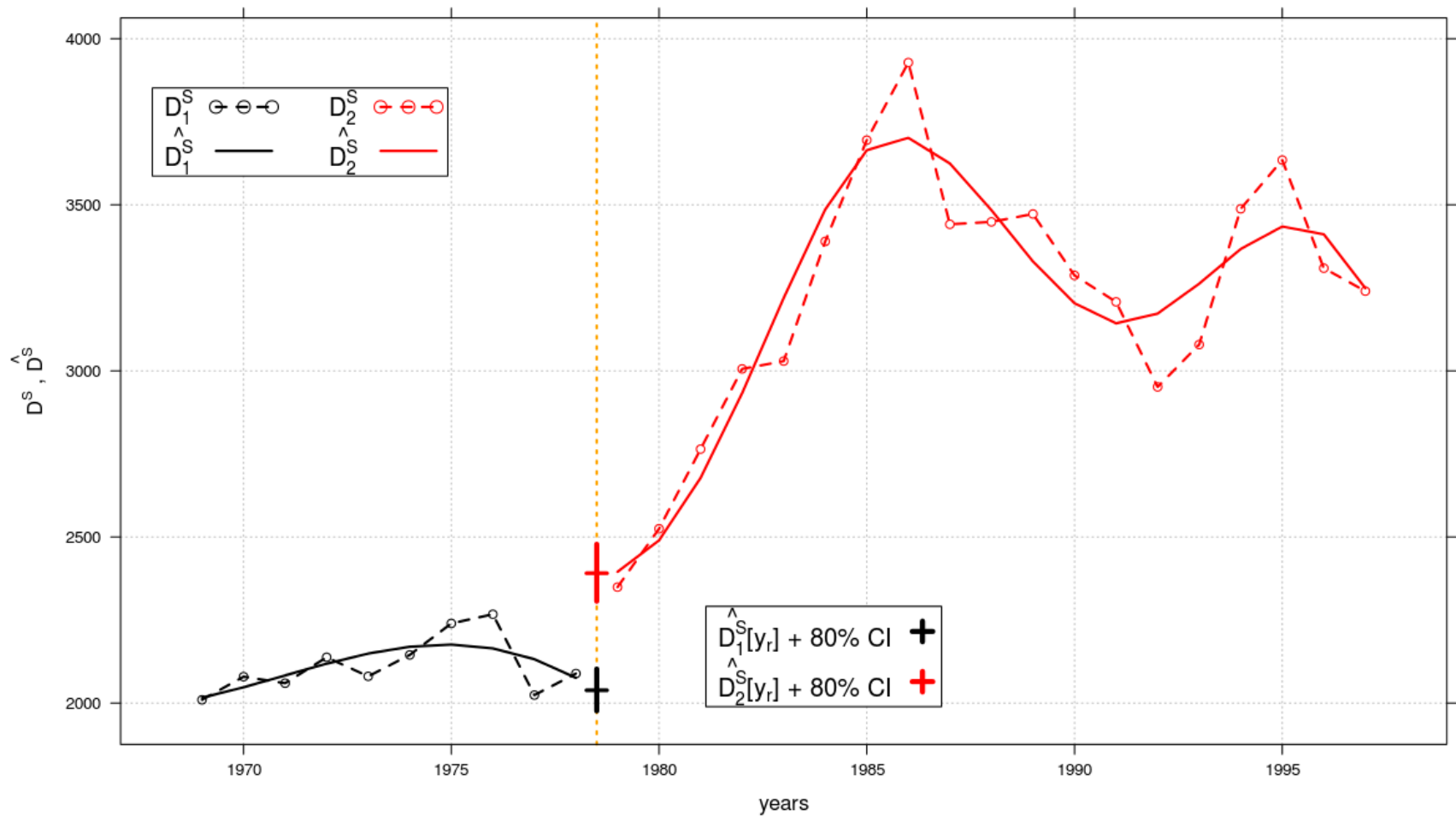
442 Other aneurysm
443 Other peripheral vascular disease
446 Polyarteritis nodosa and allied conditions
447 Other disorders of arteries and arterioles
448 Disease of capillaries

CHECKING STATISTICAL COHERENCE

CHECKING THE STATISTICAL COHERENCE OF ELEMENTARY ASSOCIATIONS



CHECKING THE STATISTICAL COHERENCE



Source: Camarda and Pechholdová, 2014

AFTER THE REDISTRIBUTION ...

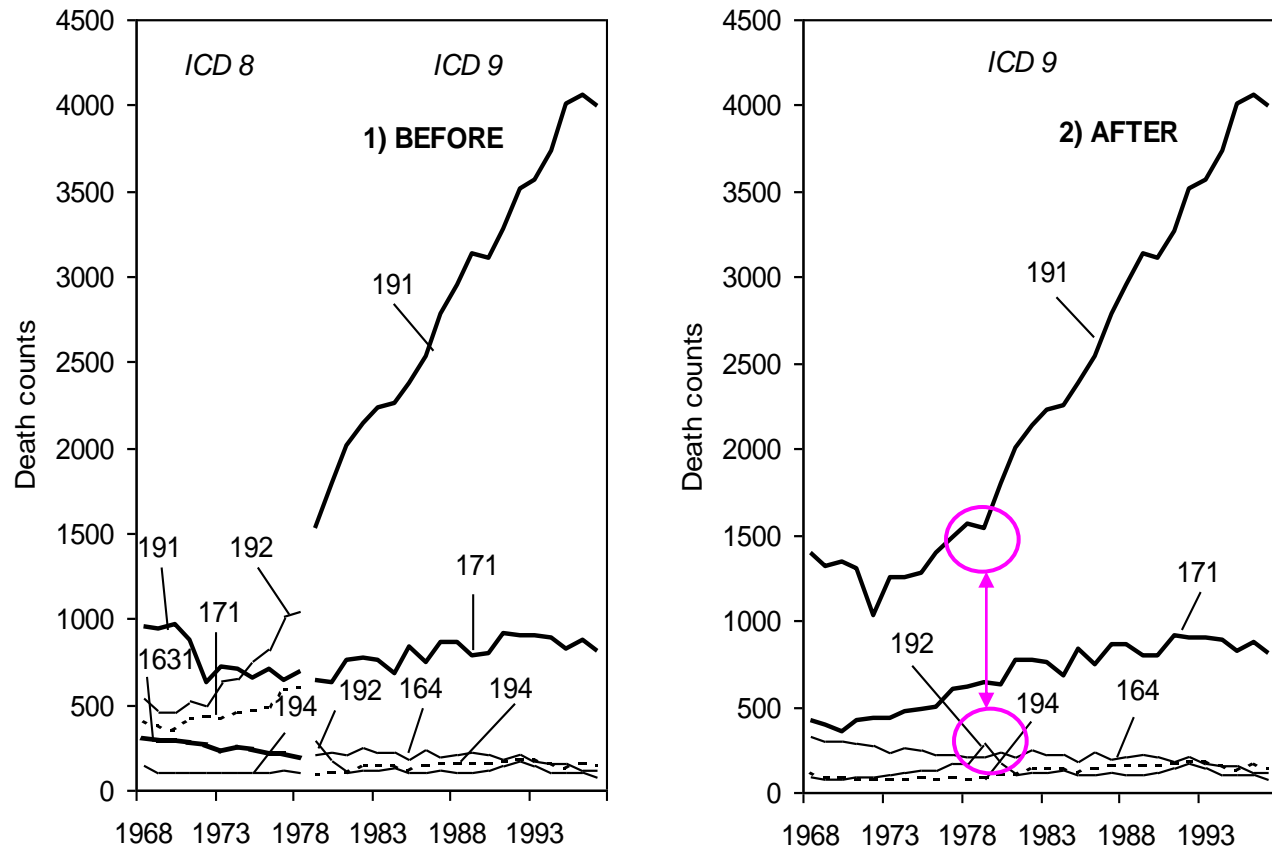
COMPLEX ASSOCIATION (2)

1. Elementary association

ICD 9	Diagnosis	1979	1978	ICD8	Portion	Diagnosis
164	Malignant neoplasm of thymus, heart, and mediastinum	211	189	1631	T	Malignant neoplasm of mediastinum
			600	171	P	Malignant neoplasm of connective and other soft tissue
			100	194	P	Malignant neoplasm of other endocrine glands
			neglect	(197a)	P	Malignant neoplasm of mediastinum, pleura and other respiratory organs, specified as secondary
			neglect	(1989)	P	Malignant neoplasm of other sites, specified as secondary
171	Malignant neoplasm of connective and other soft tissue	640		171	P	
			1041	192	P	Malignant neoplasm of other part of nervous system
			194	194	P	
			neglect	(1989)	P	
194	Malignant neoplasm of other endocrine glands and related structures	88		194	P	
			neglect	(1989)	P	
191	Malignant neoplasm of brain	1529		192	P	
			695	191	T	Malignant neoplasm of brain
192	Malignant neoplasm of other and unspecified parts of nervous system	288		192	P	
(190)			neglect	(1984)	P	
	Sum	2756	2625			

CORRECTIONS A POSTERIORI

2. Elementary association before and after reconstruction, including corrections *a posteriori*



ICD9 items in this association:

191 CA of brain

171 CA of connective and other soft tissue

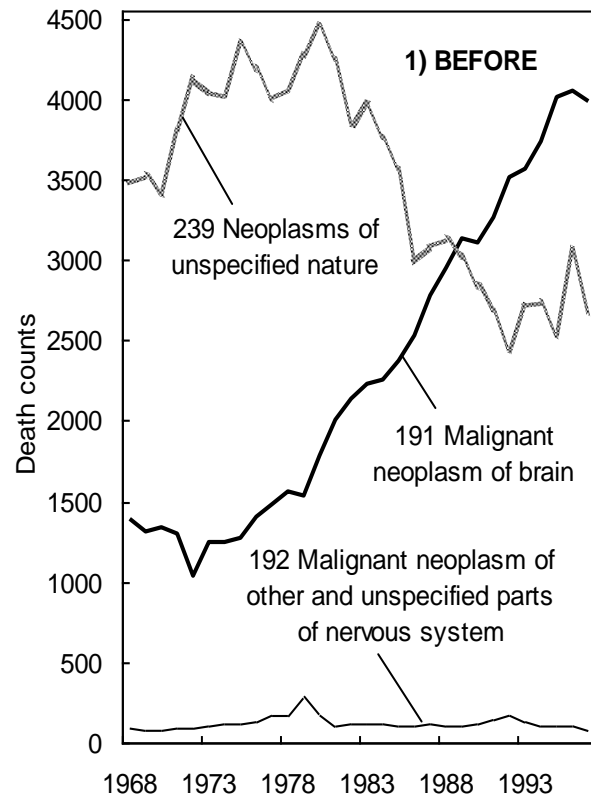
164 CA of thymus, heart and mediastinum

192 CA of other parts of nervous system

194 CA of endocrine glands

CORRECTIONS A POSTERIORI

2. Elementary association before and after reconstruction, including corrections *a posteriori*



ICD9 items in this association:

191 CA of brain

171 CA of connective and other soft tissue

164 CA of thymus, heart and mediastinum

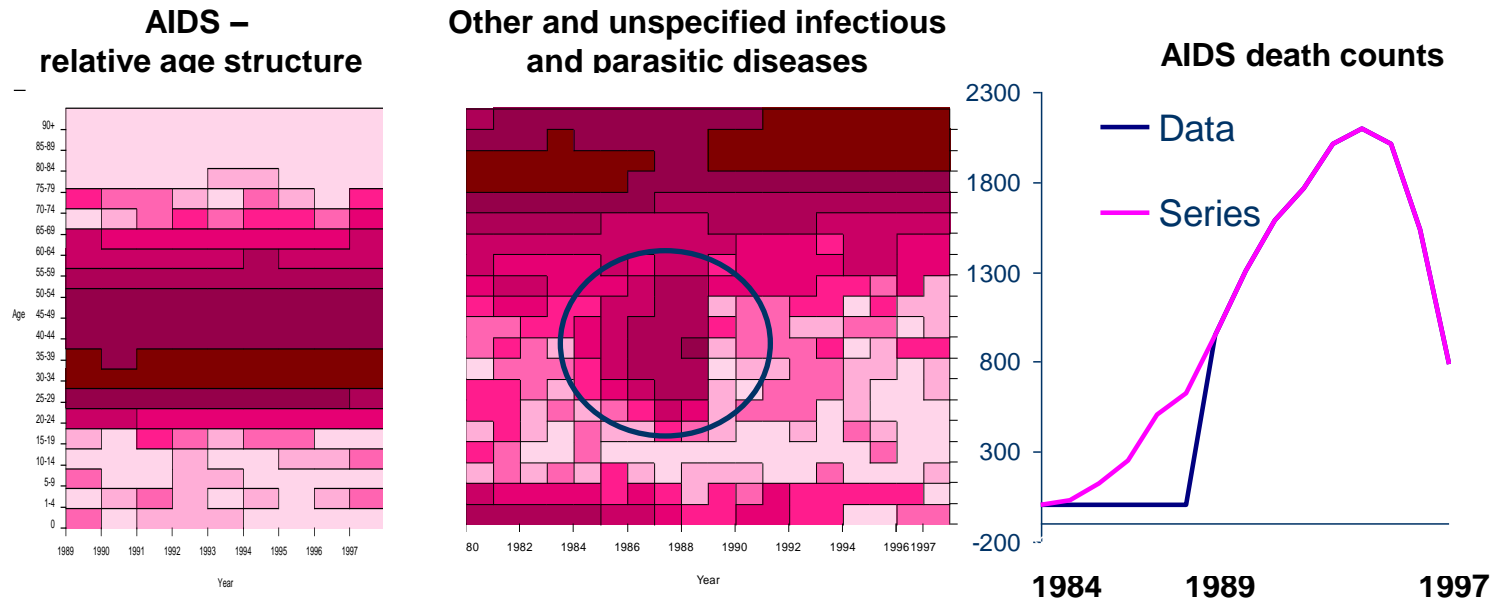
192 CA of other parts of nervous system

194 CA of endocrine glands

SPECIAL CASE OF AIDS

- 1986 – AIDS was implemented in ICD9
- 1989 – adopted in West Germany

Where the AIDS deaths were classified before 1989?



NOTE: In France, AIDS deaths were classified under “Deficiency of cell-mediated immunity”

COMPUTATIONAL TIPS

1. Transition coefficients = transition vectors

Items of ICD9	Items of ICD8					
	402	4001	412	414	425	429
402	1.000	1.000	0.024	0.000	0.000	0.000
412	0.000	0.000	0.020	1.000	0.000	0.117
414	0.000	0.000	0.663	0.000	0.663	0.000
425	0.000	0.000	0.000	0.000	0.337	0.591
429	0.000	0.000	0.292	0.000	0.000	0.292
Sum	1.000	1.000	1.000	1.000	1.000	1.000

COMPUTATIONAL TIPS

- Redistribution equation for **ICD9 402**:

$$D(402)^{ICD9} = 1.000 * D(402)^{ICD8} + 1.000 * D(4001)^{ICD8} + 0.024 * D(412)^{ICD8}$$

- Redistributed deaths are **dot (scalar) products** of two vectors: **TC** (transition coefficients) and **D** (deaths in ICD8)

Scalar product of vectors A and B: $A \cdot B = \sum_{i=1}^n A_i B_i = A_1 B_1 + A_2 B_2 + \dots + A_n B_n$

ICD9: 402

	402	4001	412	414	425	429
TC	1.000	1.000	0.024	0.000	0.000	0.000
D	9042	121	55126	2	378	1507

COMPUTATIONAL TIPS

1. Think about transition coefficients table as a transition matrix **TM** and ICD8 cause-specific deaths as vector **D**:

$$\text{TM} = \begin{bmatrix} 1.000 & 1.000 & 0.024 & 0.000 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.020 & 1.000 & 0.000 & 0.117 \\ 0.000 & 0.000 & 0.663 & 0.000 & 0.663 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.000 & 0.337 & 0.591 \\ 0.000 & 0.000 & 0.292 & 0.000 & 0.000 & 0.292 \end{bmatrix} \quad \text{D} = \begin{bmatrix} 9042 \\ 121 \\ 55126 \\ 2 \\ 1507 \\ 378 \end{bmatrix}$$

Matrix-vector product

multiply

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nj} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ \boxed{y_i} \\ \vdots \\ y_n \end{bmatrix}$$

The diagram illustrates the matrix-vector product. A matrix with elements a_{ij} and a vector x are multiplied to produce a vector y . The i -th row of the matrix and the corresponding element x_j are highlighted, and their product is shown as the i -th element of the resulting vector y_i .

Scalar product of i -th row and vector x

COMPUTATIONAL TIPS

Redistribution task can be performed as matrix-vector product:

$$D^{redistributed} = TM \cdot D$$

$$D^{redistributed} = \begin{bmatrix} 1.000 & 1.000 & 0.024 & 0.000 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.020 & 1.000 & 0.000 & 0.117 \\ 0.000 & 0.000 & 0.663 & 0.000 & 0.663 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.000 & 0.337 & 0.591 \\ 0.000 & 0.000 & 0.292 & 0.000 & 0.000 & 0.292 \end{bmatrix} \times \begin{bmatrix} 9042 \\ 121 \\ 55126 \\ 2 \\ 1507 \\ 378 \end{bmatrix}$$

SUMMARY

- **The method yields continuous series classified according to the recent ICD revision and recent coding practices.**
- **Due to different coding practices, the resulting transition coefficients are not applicable to other countries.**
- **The reconstructed detailed time series allow for an in-depth mortality research.**
- **In combination with morbidity and disease fatality measures, the detailed cause of death series permit to better assess the impact of medical progress on the mortality over time.**

The research is funded by EU structural assistance (European Social Fund) to Lithuania under the measure VP-1-3.1-ŠMM-07-K “Support to Research Activities of Scientists and Other Researchers (Global Grant)” project Nr. VP-1-3.1-ŠMM-07-K-02-067

