

'Modern Data & Health' - EHPS-net - 17-18 October 2016
Det kongelige videnskabsakademi, København

The Historical Population Register for Norway



Gunnar Thorvaldsen
Norwegian Historical Data Centre
www.rhd.uit.no



HBR – Historical Population Register

Summary

- Use: Research, genealogy, teaching
- Periods: 1801-1920, 1920-1964, 1964 -> Central Pop Register
- Sources: Censuses, church records, emigration lists (etc)
- Geographic areas: Nation, regions, municipalities
- Problems: (Scanning), **transcription**, record linkage, judicial

Project host: The University of Tromsø / Norwegian Historical Data Centre

Partners:

- The National and Regional Archives of Norway (Arkivverket)
 - Statistics Norway
 - The Norwegian Computing Center
 - The Norwegian Institute of Public Health
 - The Universities of Oslo, Bergen and Stavanger
 - The Norwegian Institute of Local History
 - BSS-database (Busetnadsoge : genealogy books)
 - Local historians and genealogists
-
- The Demographic Database at Umeå University
 - The Historical Sample of the Netherland
 - The Minnesota Population Center

Characteristics of a population registry

- It is longitudinal in the sense that it maintains a continuously updated overview of the population in an administrative geographic area.
- Records about migrants are linked together when they move inside the administrative area, ideally also when they cross its administrative borders.
- The population registry is often based on census records and reports about vital events, linking these together at the individual level and ideally in addition to record information about migration
- The composition of the population and the whereabouts of the individuals are documented more or less continuously.
- Modern, contemporary population registries are updated in real time so that migration and vital events are mirrored in the database as they are reported.
- The continuous nature of the updating of historical population registers must not be interpreted literally,
- There may be underreporting of events, especially of co-habitation, illegal immigration or short-distance mobility.

27	Tussøen						
First	Last name	Sex	Family	Marital stat	Occupation	Year	Place of birth
Hans	Hagerup Brox	M	Hf	G	Farmer Fisher	1827	
Sara	Brox	K	Hf	G	Farmer wife	1845	
Albert	Brox	M	S	U	Son	1874	
Edevard	Brox	M	S	U	Son	1877	
Nekolai	Brox	M	S	U	Son	1885	
Anne	Gunders.	K	Tj	U	Servant girl	1860	Lenvik
27	Tussøen						
Johanne	Brox	K	Hm	E	Grocer's widow, farmer	1836	Tromsø
Regine	Brox	K	D	U	Daughter	1875	
Else	Anders.	K	Tj	U	Servant girl	1876	
7	Tussøen J						
Hans A.	Brox	M	Hf	G	Farmer Fisher	1833	
Margrethe	Brox	K	Hm	G	Farmer's wife	1841	Berg i Senjen Tr
Else M.	Brox	K	D	U	Daughter	1873	
Daniel	Brox	M	S	U	Son	1874	
Haakon	Brox	M	S	U	Son	1884	
Hans M.	Brox	M	S	U	Son	1886	
Johanne	Brox	K	D	U	Daughter	1869	
Magnhild	Ols.	K	Pld	U	Step daughter	1897	
Karoline	Karls.	K	Pld	U	Step daughter	1891	
Mette	Øyen	K	Tj	U	Servant girl	1879	Tromsøsund
Lorentine	Karesius.	K	Forsør.	U	Poor relief	1840	
Edvard	Antons.	M	Hf	G	Fisher. lodger	1873	Tromsøsund

Tussøy 1900

Historical microdata –

Male census 1660s	Censuses 1801	1865, 1875, 1890, 1900. 1907 1910, 1920, 1930, 1946, 1950, 1960, 1970, 1980, 1990. 2001
Church records Andebu 1623	(forms 1812)	Tax lists 1832, 1886

Censuses

Male 1660-, 1701

Numeric 1769, 1815-1855

Nominative, public 1801, 1865, 1875, 1900

Nominative, closed 1920-1950

Nominative, closed, digital 1960-1990

Population Registers

Christiania / Oslo from 1905

Nationally, municipal from 1946

The Central Population Register from 1964

Additional sources

- Censuses (1769), 1801, (1815, 1825, 1835), 1865, 1875, 1890, 1900, . . .
- Almost complete church register from the mid 18th century

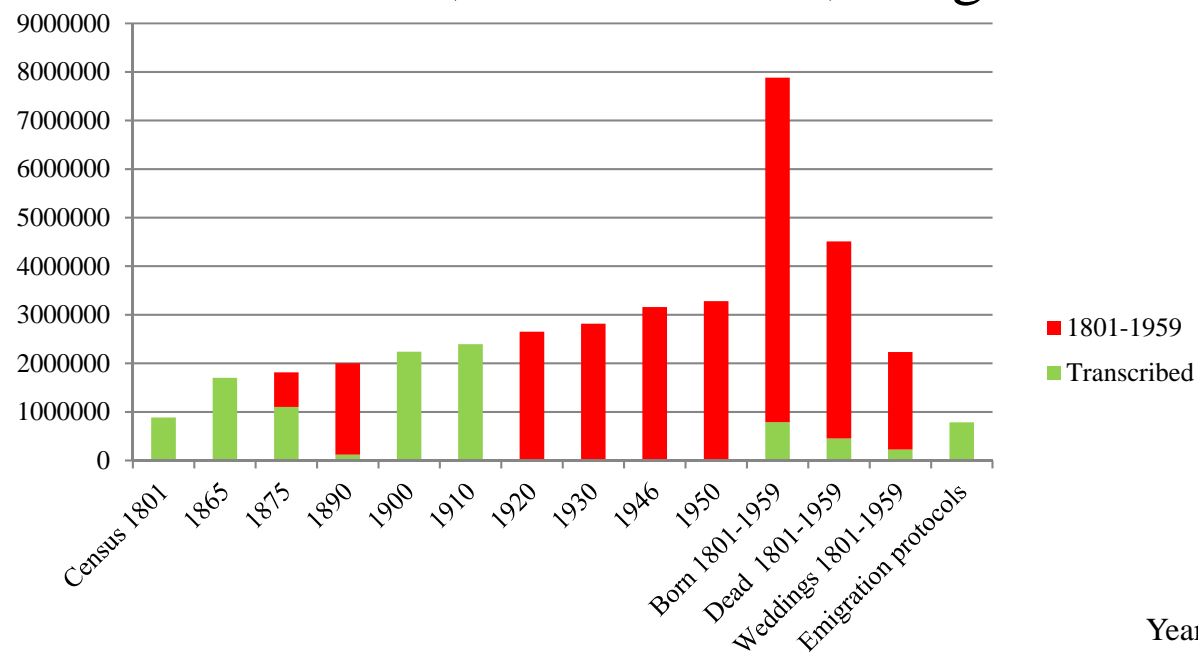
Confirmations

In- and outmigration (from 1812, partial)

- Reports of births, marriages, deaths to Statistics Norway from 1866
- Central birth register from 1916
- Immigration from 1916

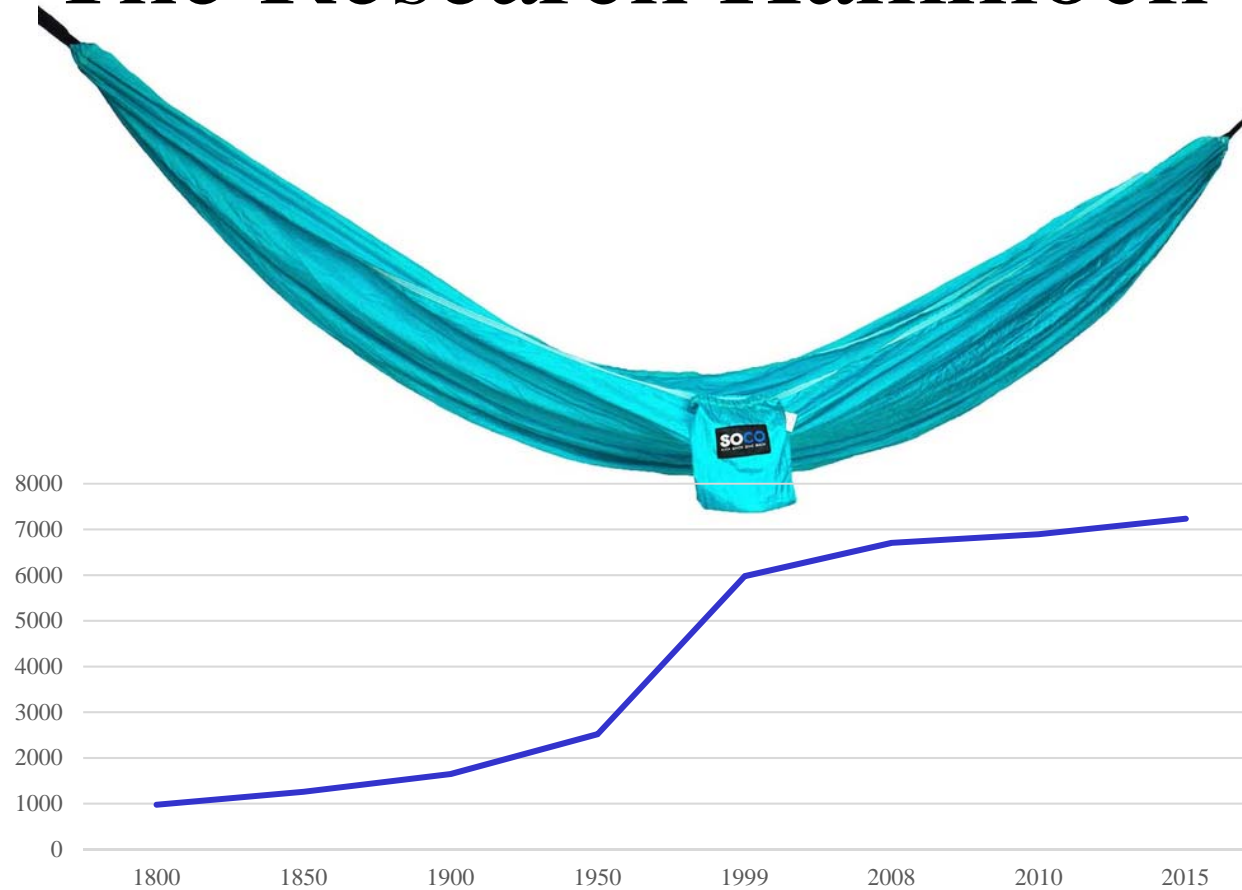
- Local population registers from 1905 / complete 1946
- Central population register from 1964

Total number of records, and remaining transcription in censuses, church books, emigration lists 19th to 20th century

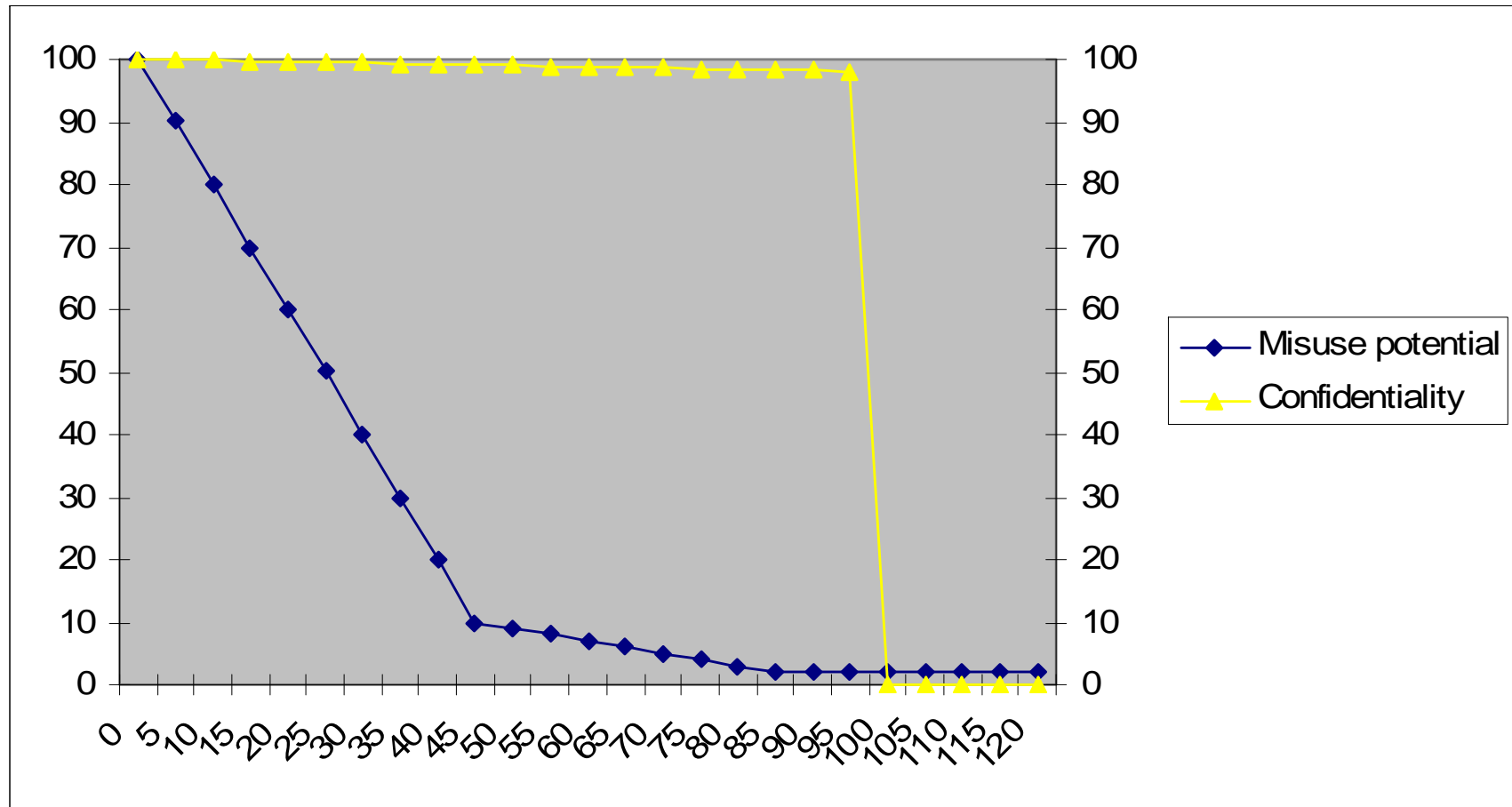


	Year	Records	Transcribed	Remaining	
Census	1801	883603	883603	0	0
	1865	1701756	1701756	0	0
	1875	1813424	1100000	713424	713424
	1890	2000917	120000	1880917	1880917
	1900	2240032	2240032	0	0
	1910	2391782	2391782	0	0
	1920	2649775		2649775	
	1930	2814194		2814194	
	1946	3156950		3156950	
	1950	3278546		3278546	
Born 1801-1959		7881121	788112	7093009	5088464
Dead 1801-1959		4508279	450828	4057451	2981562
Weddings 1801-1959		2231793	223179	2008614	909258
Emigration protocols		786000	786000	0	0
Sum		38338172	10685292	27652880	11573625

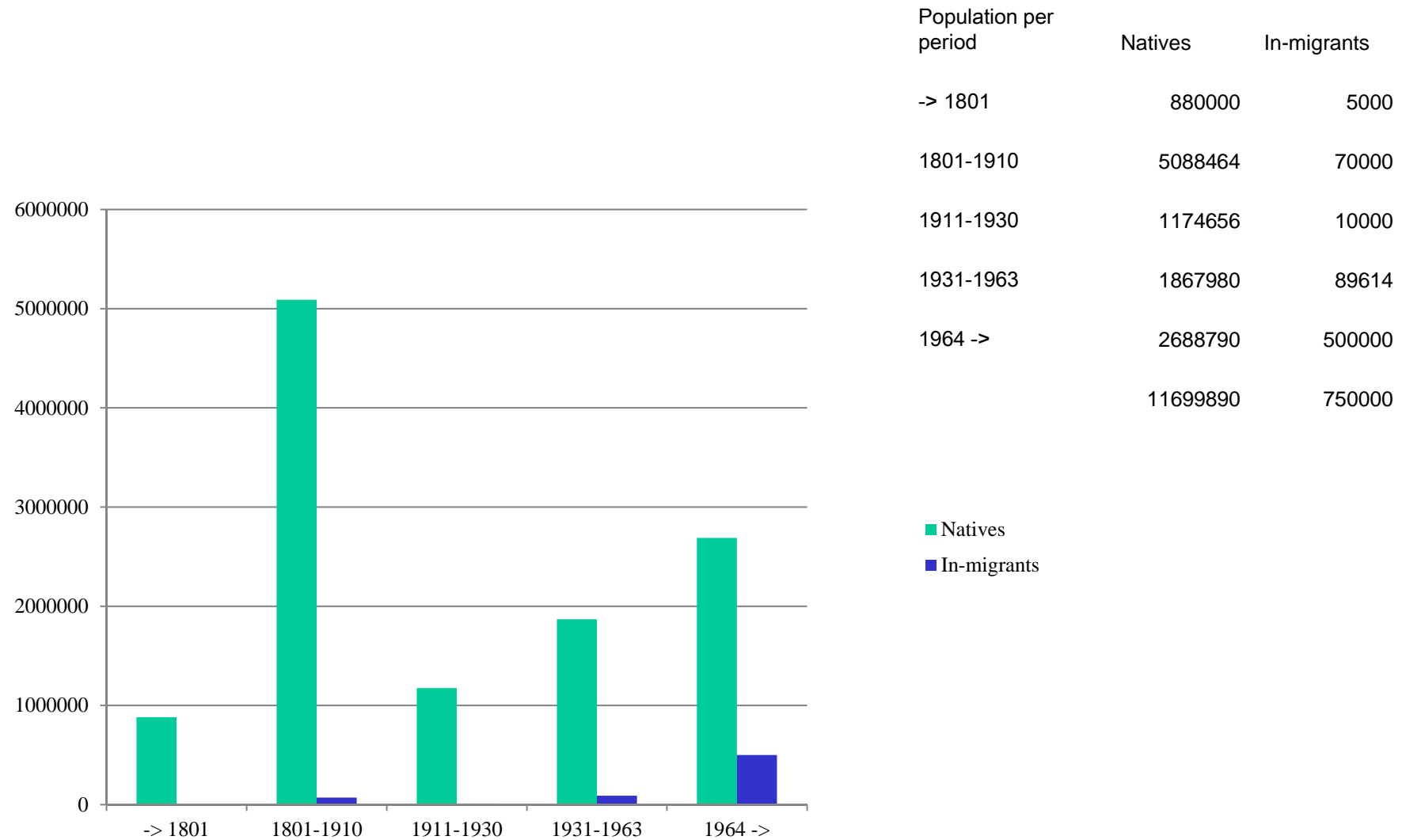
The Research Hammock



Limited access due to confidentiality in 20th century microdata



Population to be contained in the population register by period of privacy protection



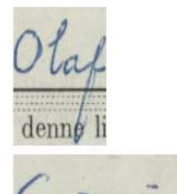
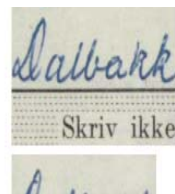
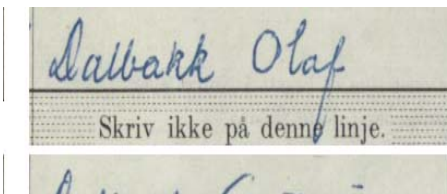
The closed period ca 1920 - 1964

- 1920 census – transcription from ca 2020
- 1930 census – being scanned
- 1946 census – no concrete plans
- 1950 census – scanning completed – semi-automatic transcription
- Mortality register 1951-> in database at National Institute of Health
- Marriages – scanned until ca 1950 – transcription abroad?
- Birth registers – scanned until ca 1950 – transcription abroad until ca 1935?
- Central Population Register 1964 -> - being transferred from Statistics Norway to the National Archives

1950 census – semi-automatic transcription



- Software developed by the Norwegian Computing Center
- The information is cut to virtual pieces fieldwise
- Reference #, sex, marital and family status read «automatically»
- Birthplaces are clustered
- Names and birth dates will be sent abroad for transcription
- What resources needed for proof-reading?
- A 2 % sample being transcribed manually
- Will be linked to the Central Pop Register and the older Open Pop Register (early) next year



Online clustering of birthplaces



Irrelevant pictures are marked for exclusion by clicking before collective transcription

Mortality after the Death of a Spouse in Norway

- Extract from the Central Population Register (CPR) for Norway
- 912,597 women and 888,859 men who were born before 1952
- Left truncated as married or widowed from 1975 until 2006
- Contains birth year, annually updated marital status, updated spousal reference numbers and date of death until 2006
- Widowed persons who remarried were right censored at date of new marriage
- 291 couples who died on the same day analysed as special cases
- Spousal death compared with the expected number of deaths as computed from the life tables
- During the first year after losing the spouse, mortality is 34 % higher among men and 29 % higher among women, compared to the mortality among married persons
- Excess proportion of deaths is significant from age group 55-59 to 90-94, biggest for age group 60-64 for both genders
- Significant for a 10 year period after spousal death, but highest during the first days
- No social status indicators such as occupation or education

Ytterstad, E. and T. Brenn (2015). Epidemiology **26**(3): 289-294.

TABLE 2. Number of Deaths by Time Since Spousal Death in Widowed Individuals Ages 55 to 94 Years in Norway, 1975–2006

Time Since Spousal Death	Widowed Men			Widowed Women		
	Observed	Expected	HR (95% CI)	Observed	Expected	HR (95% CI)
Days						
1–7	294	174.5	1.69 (1.49, 1.88)	300	170.6	1.76 (1.56, 1.96)
8–30	879	570.0	1.55 (1.44, 1.65)	812	557.9	1.46 (1.36, 1.56)
31–100	2,278	1,702.9	1.35 (1.29, 1.40)	2,179	1,678.6	1.30 (1.25, 1.36)
101–365	7,739	6,049.2	1.31 (1.28, 1.34)	7,578	6,103.1	1.25 (1.23, 1.28)
1–365	11,190	8,496.6	1.34 (1.31, 1.36)	10,869	8,510.2	1.29 (1.26, 1.31)
Year						
2	9,612	8,038.0	1.25 (1.22, 1.27)	10,315	8,609.4	1.22 (1.20, 1.25)
3	8,837	7,559.9	1.25 (1.22, 1.27)	10,344	8,660.0	1.23 (1.21, 1.26)
4	8,090	7,082.5	1.25 (1.22, 1.27)	10,091	8,711.8	1.21 (1.19, 1.24)
5	7,130	6,575.3	1.21 (1.18, 1.24)	9,897	8,632.2	1.21 (1.19, 1.24)
6	6,630	6,076.9	1.25 (1.22, 1.28)	9,856	8,510.7	1.24 (1.22, 1.27)
7	6,008	5,438.1	1.26 (1.22, 1.29)	9,524	8,133.0	1.23 (1.21, 1.26)
8 ^a	5,338	4,984.8	1.25 (1.21, 1.28)	9,366	7,936.0	1.25 (1.23, 1.28)
9 ^b	4,802	4,508.1	1.27 (1.23, 1.30)	8,982	7,664.0	1.26 (1.23, 1.29)
10 ^c	4,151	4,103.2	1.24 (1.20, 1.28)	8,744	7,366.4	1.29 (1.26, 1.32)
1–10 ^c	70,162	60,688.9	1.26 (1.25, 1.27)	96,948	81,031.9	1.25 (1.24, 1.25)
^a Age 55 to 93. ^b Age 55 to 92. ^c Age 55 to 91.						

Issues to be discussed for all databases

- If there is a time gap between the contents of the database and present-day population and medical data, how can this gap be covered? In other words: potential methods and strategies for bridging the gap. *The censuses and vital event lists from 1920 are scanned and transcribed abroad or via clustering.*
- What “epidemiological” information is there in the sources and with what reliability – such as causes of death, stature, functional disorders. *Causes of death in databases from 1950, earlier to be transcribed from scanned pictures.*
- What links can be made to health sources – clinical data, medical records, biobanks (existing, planned or potential). *The Institute of Public Health is a partner and has biobanks and medical records.*
- How are issues of privacy protection and data security handled? *The National Archives negotiate legal issues with the Data Protection Inspectorate.*
- Other obstacles and problems.

What is being / to be done?

- digitized and encoded the entire 1801, 1865, 1900, 1910 census
- scanned most church books to 1930
- transcribing 19th and 20th century ministerial records, in China
- 1801 to 1815 are transcribed and being linked
- standardized most names and other variables
- local and regional record linkage progressing
- early 20th century record linkage with birth dates
- 1950 census is scanned, to be transcribed fieldwise and linked to population register
- completing the 1875 and 1891 censuses