The Danish National Diabetes Register (NDR)

Bendix Carstensen
Steno Diabetes Center
& Department of Biostatistics, University of Copenhagen
bxc@steno.dk http://BendixCarstensen.com

Data processing in Health Care
Copenhagen, December 2014

Background for the NDR

Population surveillance
- Monitor and describe:
    - Prevalence (no. and %)
    - Incidence (no. and rates)
    - Mortality and SMR

Health care surveillance
- Keep track of diabetes patients
- Predictions of likely future developments
- Match patients to treatment indicators (GPs)
- . . . improve accuracy of treatment information

Results up to 31.12.2006 reported in:
Carstensen *et al.*: The Danish National Diabetes Register: Trends in incidence, prevalence and mortality, Diabetologia, 2008.

Construction of the register

- Based on existing registers in Denmark:
    - National Patient Register
      — all hospital and outpatient clinic contacts.
    - National Health Insurance Service Registry
      — all services provided in the NHS.
    - Register of Medicinal Product Statistics
      — all prescriptions taken out at pharmacies.
- Linked to mortality and migration data from the Central Person Register.
- All records are CPR-identified, for linkage purposes.
- Inclusion start at 1.1.1990.

Inclusion criteria for the register

- ▸ Diagnosis of DM in NPR.
- ▸ Gestational diabetes excluded. A diagnosis of GDM precludes inclusion for a period of 1 year.
- ▸ Foot-therapy for diabetics recorded in NHISR.
- ▸ 5 blood-glucose measurements within 1 year recorded in NHISR.
- ▸ 2 blood-glucose measurements per year in 5 consecutive years recorded in NHISR.
- ▸ Prescription on insulin or oral antidiabetics in Register of Medicinal Product Statistics.

  Metformin alone in women aged 20–39 excluded (PCOS).

Variables in the NDR

- ▸ D_FODDTO   - date of birth
- ▸ C_SEX   - sex
- ▸ D_INKLDTO   - date of inclusion
- ▸ C_INKLAARSAG - criterion first met
- ▸ D_DODSDTO   - date of death
- ▸ D_LPR   - first DM diagnosis in LPR
- ▸ D_FODT   - first date of chiropody
- ▸ D_BLOD2I5   - first date of 2 BG / 5y
- ▸ D_BLOD5I1   - first date of 5 BG / 1y
- ▸ D_OAD   - date of 2nd OAD purchase
- ▸ D_INS   - date of 2nd insulin purchase
- ▸ V_PID   - person-id

Random sample from NDR

```
   D_FODDTO C_SEX D_INKLDTO  C_INKLAAR D_DODSDTO      D_LPR      D_FODT   D_BLOD2I5

   09NOV1935   K   12OCT2009  oad           .            .          .         .
   11SEP1919   M   19APR1990  lpr       22MAY1992   19APR1990      .         .
   12JUN1923   K   03JUN1998  blod5i1   22FEB2008       .          .         .
   18MAR1936   M   18APR2001  blod5i1       .        06JUN2007  23MAY2007     .
   12AUG1959   K   08OCT2008  blod5i1       .            .          .         .
   24DEC1941   M   16MAR2005  blod5i1   24FEB2007       .          .         .
   03JUL1944   M   09JAN2003  oad           .            .          .         .
   22JAN1964   K   22JAN1997  blod5i1       .            .          .         .
   29MAR1941   K   01OCT2009  lpr           .        01OCT2009      .         .
   01JUN1949   M   06OCT2005  oad           .            .          .         .
   15AUG1962   M   29SEP2009  oad           .            .          .         .
   02APR1949   K   18AUG2004  blod5i1       .        21JAN2009  19MAR2008  23APR2008
   21JUL1931   K   14MAY2003  blod5i1       .            .          .         .
   08OCT1901   K   08AUG1992  lpr       20DEC1993   08AUG1992      .         .
   19APR1913   K   23JAN1991  fodt      29AUG1992       .       23JAN1991     .
   09MAR1913   K   03APR1998  oad       20MAY1999       .          .         .
   15APR1947   M   24APR2001  oad           .        21MAY2001      .         .
   12DEC1940   K   16JUL2002  lpr           .        16JUL2002      .         .
   31DEC1916   M   24MAY1991  lpr       28JUN1991   24MAY1991      .         .
   21JUN1919   M   16FEB1992  lpr       15NOV1993   16FEB1992  20JAN1993     .
   31DEC1944   K   05OCT1993  lpr           .        05OCT1993  21APR2004     .
   30JUN1916   K   01FEB2006  blod5i1   18MAR2009       .          .         .
   16OCT1971   K   08DEC2004  blod5i1       .            .          .         .
   16MAY1965   K   25MAY2005  blod5i1       .        22MAY2006      .         .
   06AUG1923   K   28OCT1998  blod5i1   01APR2004       .          .         .
   26JAN1932   M   20FEB2008  blod5i1       .            .       21MAY2008     .
   16JUN1932   M   25FEB1998  lpr       24APR2006   25FEB1998  18NOV1998  04AUG2004
   15FEB1914   M   22JUL1992  blod5i1   17FEB1993       .          .         .
   05MAR1957   M   11AUG2004  blod5i1       .        27AUG2004      .         .
   17OCT1948   M   14MAY2008  blod5i1       .            .          .         .
```

Random sample from NDR

| D_INKLDTO	C_INKLAAR	D_LPR	D_FODT	D_BLOD2I5	D_BLOD5I1	D_OAD	D_INS
12OCT2009	oad	.	.	.	.	12OCT2009	.
19APR1990	lpr	19APR1990	.	.	04JUL1990	.	.
03JUN1998	blod5i1	.	.	.	03JUN1998	.	.
18APR2001	blod5i1	06JUN2007	23MAY2007	.	18APR2001	01JUN2001	.
08OCT2008	blod5i1	.	.	.	08OCT2008	.	.
16MAR2005	blod5i1	.	.	.	16MAR2005	.	.
09JAN2003	oad	.	.	.	12DEC2007	09JAN2003	.
22JAN1997	blod5i1	.	.	.	22JAN1997	.	.
01OCT2009	lpr	01OCT2009	.	.	.	22OCT2009	.
06OCT2005	oad	.	.	.	11JAN2006	06OCT2005	.
29SEP2009	oad	.	.	.	.	29SEP2009	.
18AUG2004	blod5i1	21JAN2009	19MAR2008	23APR2008	18AUG2004	08SEP2007	.
14MAY2003	blod5i1	.	.	.	14MAY2003	.	.
08AUG1992	lpr	08AUG1992	.	.	.	.	.
23JAN1991	fodt	.	23JAN1991	.	.	.	.
03APR1998	oad	.	.	.	.	03APR1998	.
24APR2001	oad	21MAY2001	.	.	.	24APR2001	.
16JUL2002	lpr	16JUL2002	.	.	.	13JAN2006	17JAN2006
24MAY1991	lpr	24MAY1991	.	.	.	.	.
16FEB1992	lpr	16FEB1992	20JAN1993	.	17JUN1992	.	.
05OCT1993	lpr	05OCT1993	21APR2004	.	.	05NOV1994	.
01FEB2006	blod5i1	.	.	.	01FEB2006	.	.
08DEC2004	blod5i1	.	.	.	08DEC2004	.	.
25MAY2005	blod5i1	22MAY2006	.	.	25MAY2005	23NOV2005	.
28OCT1998	blod5i1	.	.	.	28OCT1998	.	.
20FEB2008	blod5i1	.	21MAY2008	.	20FEB2008	05MAR2008	.
25FEB1998	lpr	25FEB1998	18NOV1998	04AUG2004	27SEP2000	03MAR2000	30JUL2002
22JUL1992	blod5i1	.	.	.	22JUL1992	.	.
11AUG2004	blod5i1	27AUG2004	.	.	11AUG2004	11SEP2004	.
14MAY2008	blod5i1	.	.	.	14MAY2008	.	.
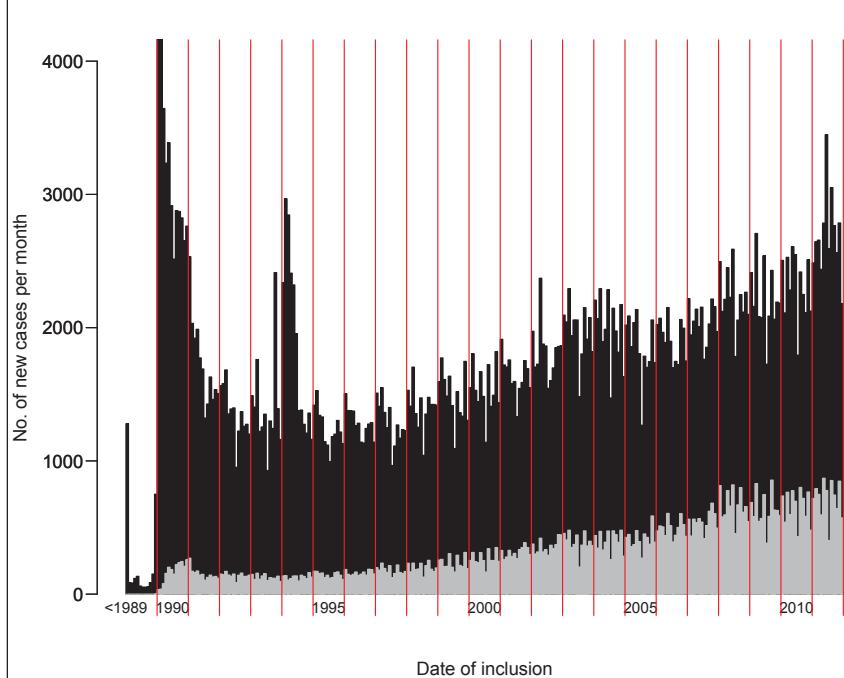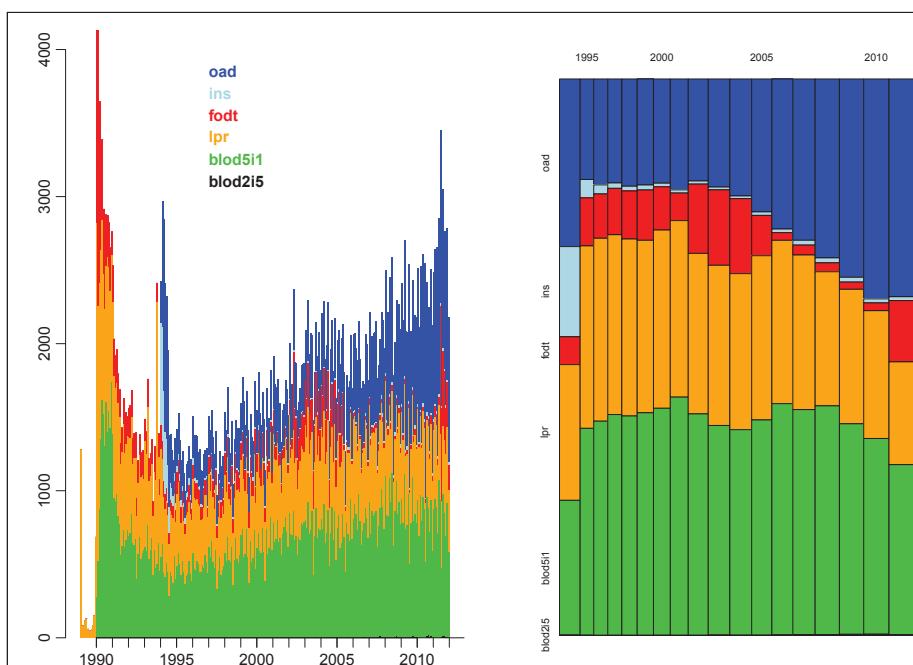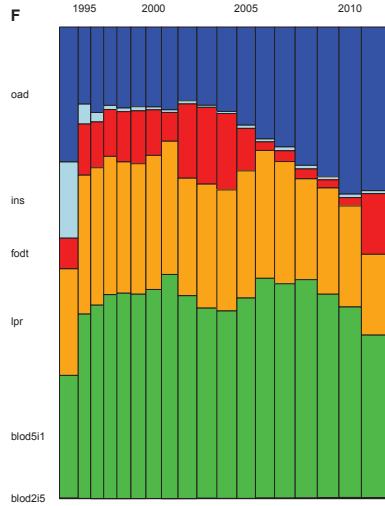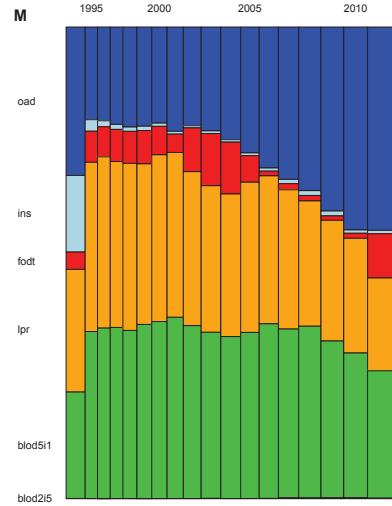
Methods: Incidence

- ▶ New cases tabulated by age and date of diagnosis (1995-2007) and date of birth, in 1-year classes.

- ▶ Corresponding person-years figures from Statistics Denmark — person-years in the diabetes register subtracted.

- ▶ Incidence rates analysed by Poisson-regression with smooth parametric terms in age and date of diagnosis, using log-person-years as offset.
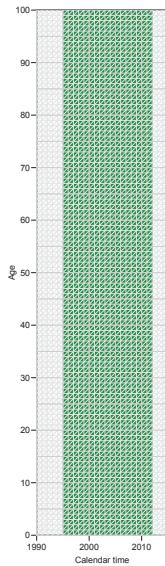
Digression: Lexis diagram

Wilhelm Lexis
(1837–1914)
German demographer,
statistician and
economist.

New cases in NDR	2006			2011		
Year	M	F	All	M	F	All
≤ 1989	1,480	1,310	2,790	1,514	1,330	2,844
1990	21,347	24,738	46,085	21,434	24,775	46,209
1991	10,681	9,987	20,668	10,763	10,055	20,818
1992	8,554	7,855	16,409	8,463	7,786	16,249
1993	9,165	7,639	16,804	9,196	7,642	16,838
1994	12,103	10,733	22,836	11,993	10,688	22,681
1995	7,745	7,148	14,893	7,756	7,150	14,906
1996	8,015	7,388	15,403	8,016	7,388	15,404
1997	7,923	7,528	15,451	7,928	7,533	15,461
1998	8,800	8,039	16,839	8,819	8,048	16,867
1999	9,295	8,537	17,832	9,314	8,565	17,879
2000	9,614	8,881	18,495	9,620	8,883	18,503
2001	10,181	9,468	19,649	10,215	9,481	19,696
2002	11,123	10,745	21,868	11,178	10,790	21,968
2003	12,385	11,378	23,763	12,370	11,361	23,731
2004	12,465	11,465	23,930	12,472	11,477	23,949
2005	11,607	10,535	22,142	11,619	10,582	22,201
2006	12,007	10,865	22,872	12,094	10,920	23,014
2007				12,719	11,783	24,502
2008				14,005	12,663	26,668
2009				14,298	12,360	26,658
2010				14,974	12,818	27,792
2011				17,080	15,314	32,394
1995–2006	121,160	111,977	233,137	121,401	112,178	233,579
1995–2011				194,477	177,116	371,593
Total	184,490	174,239	358,729	257,840	239,392	497,232
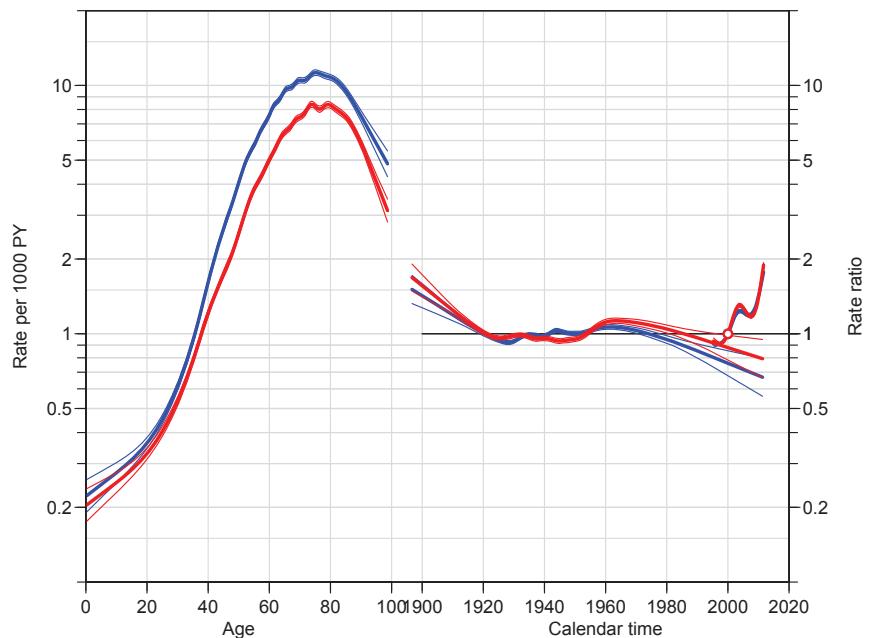
Incidence summary

▶ Women with gestational diabetes are closer monitored (?), hence the increased incidence rates in ages 20–35 ($\mathrm{RR} \approx 1.5$).

▶ Incidence rates peak in ages 70–85: 1.2%/year for men, 1.2%/year for women in 2005.

▶ Annual increase in incidence rates over the period 1995–2004 was 5.6%, after 2004 almost flat

Glucose criteria

- ▶ Women who have a glucose tolerance test triggesrs typically 6 blood glucose measuremnts.
- ▶ Omitting the glucose criteria:
  - ▶ Some are removed from the register
  - ▶ Some have a later diagnosis (meeting a different criterion).
- ▶ The new diagnostic criteria based on HbA1c makes the blood glucose criteria even more uncertaing.
- ▶ No consensus on how to proceed.
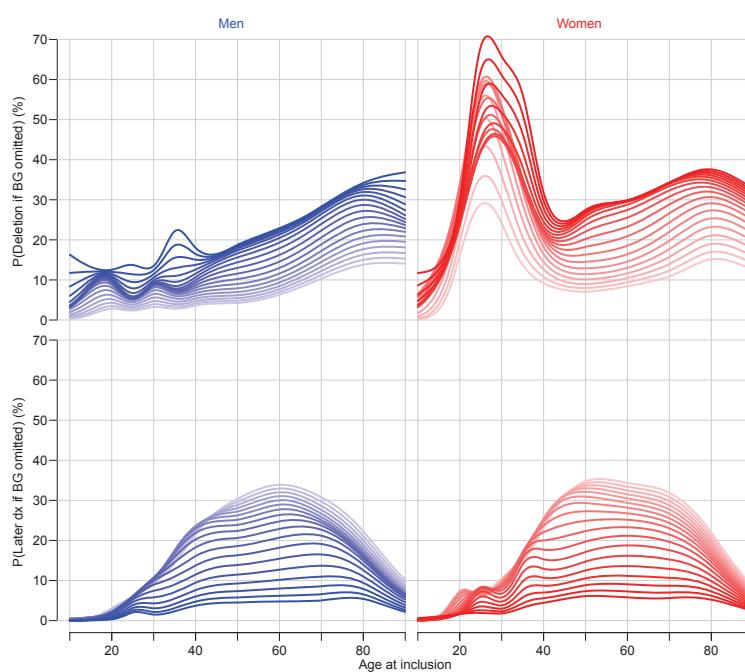
Methods: Prevalence

- ▶ Prevalent cases by 1 Jan 1995,. . . ,2010 tabulated by sex and 1-year age.
- ▶ Corresponding population figures from Statistics Denmark.
- ▶ Prevalence analysed by a binomial model with log-link and the population size as denominator. Separate parametric terms used for each sex and date.

Prevalence of diabetes

1 January	No. patients			Prevalence (%)		
	M	F	All	M	F	All
1995	49,438	49,126	98,564	1.92	1.86	1.89
1996	53,443	52,631	106,074	2.06	1.98	2.02
1997	57,739	56,551	114,290	2.22	2.12	2.17
1998	61,832	60,522	122,354	2.36	2.26	2.31
1999	66,746	64,974	131,720	2.54	2.42	2.48
2000	71,798	69,692	141,490	2.73	2.59	2.65
2001	77,120	74,570	151,690	2.92	2.76	2.84
2002	82,914	79,895	162,809	3.12	2.94	3.03
2003	89,291	86,364	175,655	3.35	3.17	3.26
2004	96,706	93,199	189,905	3.62	3.42	3.52
2005	104,149	100,227	204,376	3.89	3.67	3.78
2006	110,581	106,028	216,609	4.12	3.87	3.99
2007	117,328	112,018	229,346	4.35	4.07	4.21
2008	124,501	118,726	243,227	4.59	4.30	4.44
2009	132,847	126,295	259,142	4.86	4.54	4.70
2010	140,940	133,318	274,258	5.14	4.78	4.96
2011	149,702	140,507	290,209	5.43	5.01	5.22
2012	160,352	150,309	310,661	5.80	5.34	5.57

Prevalence of diabetes — age distribution
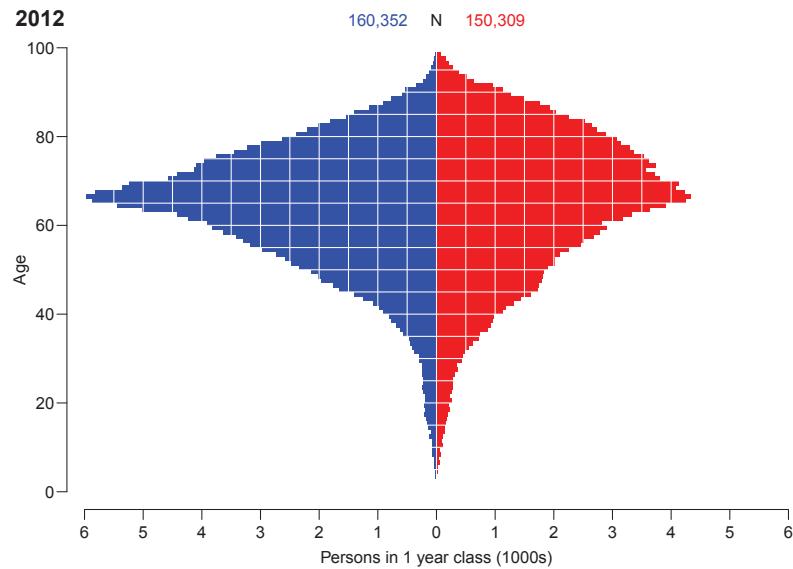
Prevalence of diabetes — % by age
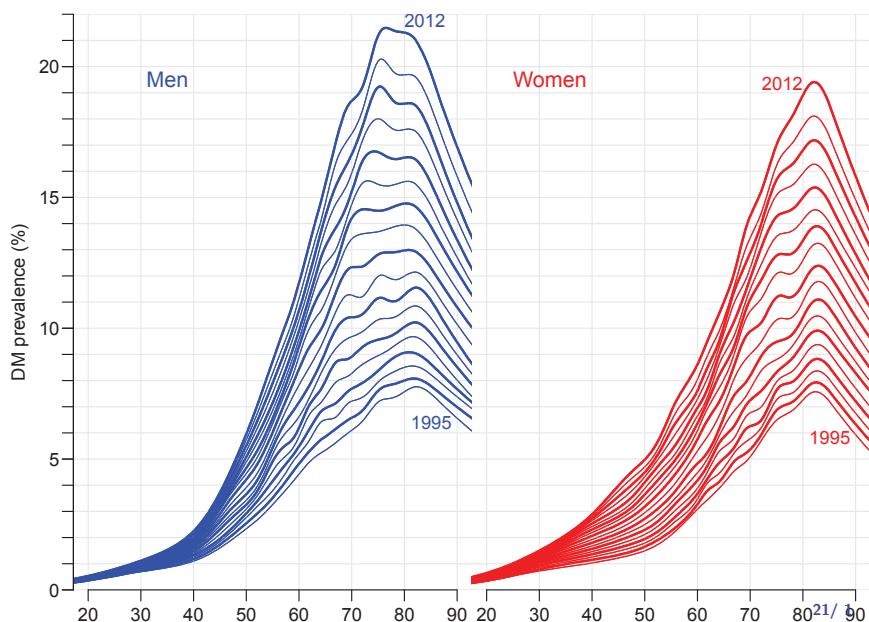
Methods: Mortality and SMR

- ▶ Deaths and person-years of follow-up among diabetics tabulated by age and period at follow-up in 1-year classes.
- ▶ Corresponding mortality figures from Statistics Denmark. Deaths and person-years from the diabetes cohort subtracted.
- ▶ Mortality analysed by Poisson-regression of deaths with smooth parametric terms for current age, current date and current disease duration, using log-person-years as offset.
- ▶ SMR analysed using dataset amended by mortality among non-DM persons, using interaction between DM / non-DM and age / duration.

Number of deaths — imbalance

Year	New cases	Deaths	Surplus
1995	14,874	7,377	7,497
1996	15,385	7,147	8,238
1997	15,438	7,366	8,072
1998	16,842	7,457	9,385
1999	17,853	8,058	9,795
2000	18,480	8,242	10,238
2001	19,675	8,521	11,154
2002	21,948	9,072	12,876
2003	23,712	9,427	14,285
2004	23,927	9,421	14,506
2005	22,186	9,879	12,307
2006	23,001	10,227	12,774
2007	24,477	10,544	13,933
2008	26,648	10,647	16,001
2009	26,639	11,455	15,184
2010	27,770	11,767	16,003
2011	32,374	11,782	20,592

Incident cases

Mortality
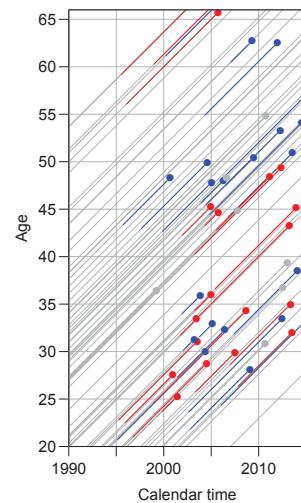
Mortality among DM patients
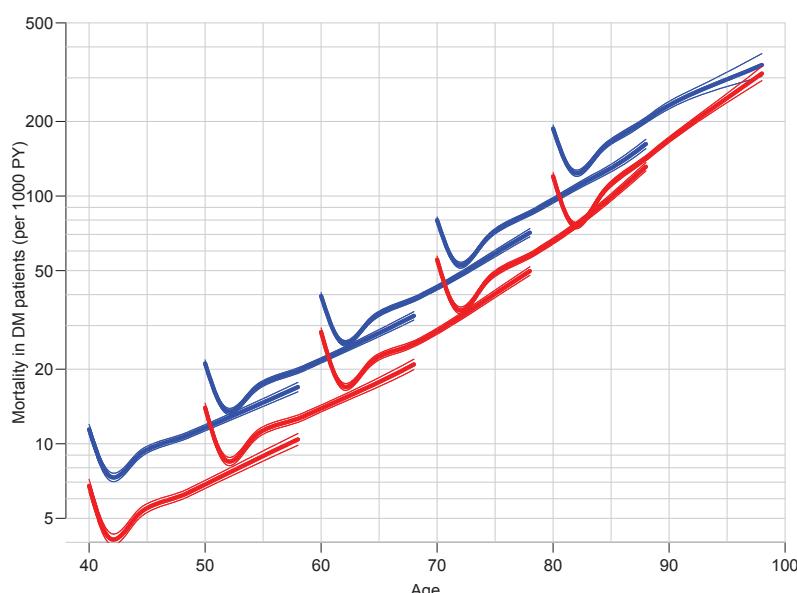
SMR relative to persons without DM
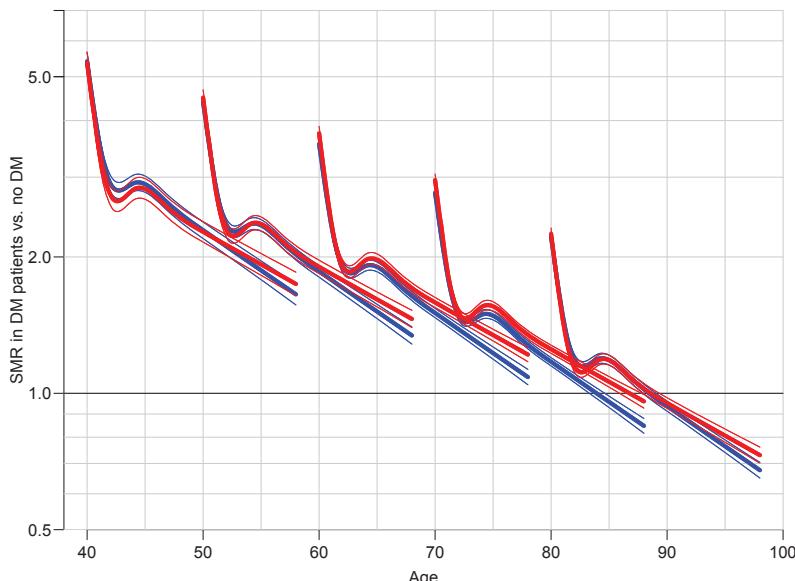
Mortality summary

- ▸ Mortality in DM patients increases exponentially by age.
- ▸ Decreasing by time: 4.0/3.8%/year (non-DM: 2.8/2.4)
- ▸ SMR is the same for men and women.
- ▸ SMR is 3 at age 45, 1 at 85

Summary: Danish Diabetes Register

Note: a register of "administrative" diabetes:

- ▸ Population based 1995–2011 (17 years)
- ▸ Coverage 100%
- ▸ No selection bias at individual level
- ▸ Imprecise diagnostic criteria
- ▸ Incidence rates increasin in general
- ▸ Mortality and SMR decrease
- ▸ Prevalence increases

Cancer among diabetes patients

- ▶ Merge the Diabetes Register with the Danish Cancer Register.
- ▶ Compute the RR of cancer between persons with and without diabetes.
- ▶ 25 cancer sites, 2 sexes, age-interaction, duration.

Aims

- ▶ Describe cancer incidence rates among diabetes patients in Denmark.
- ▶ and how rates vary relative to the non-DM population with:
    - ▶ duration of diabetes
    - ▶ duration of insulin use
- ▶ for all types of cancer
- ▶ and for specific sites of cancer

ARTICLE

Cancer occurrence in Danish diabetic patients: duration and insulin effects

B. Carstensen · D. R. Witte · S. Friis

Abstract

Aims/hypothesis Cancer is more frequent among diabetes patients, but it is unknown how this excess varies with duration of diabetes and insulin use. The aim of this study was to analyse disease data to examine this issue further. *Methods* We linked the Danish National Diabetes Register and Cancer Registry and performed a cohort analysis of the entire Danish population by diabetes status, duration of diabetes and insulin use, comparing cancer incidence rates in diabetic patients with the non-diabetic population for the population decreased from over 2 at diagnosis to 1.15 after 2 years of diabetes duration. The cancer incidence rate ratio was higher among patients using insulin, decreasing from 5 at the start of insulin treatment to about 1.25 after 5 years of insulin use. Among non-insulin users, cancers of the stomach, colorectum, liver, pancreas, lung, corpus uteri, kidney and brain, and lymphomas were elevated. Among insulin users the rate ratio of prostate cancer was decreasing by duration whereas we found higher risk of cancer of the stomach, lung, liver, pancreas and kidney. Breast cancer

Follow-up of the Danish population

Well 75,637,670 → 346,138 → Ca

Well 75,637,670 → DM 1,279,524

DM 1,279,524 → 20,032 → DM + Ca

DM 1,279,524 → DM/Ins 239,516

DM/Ins 239,516 → 2,794 → DM/Ins + Ca

---
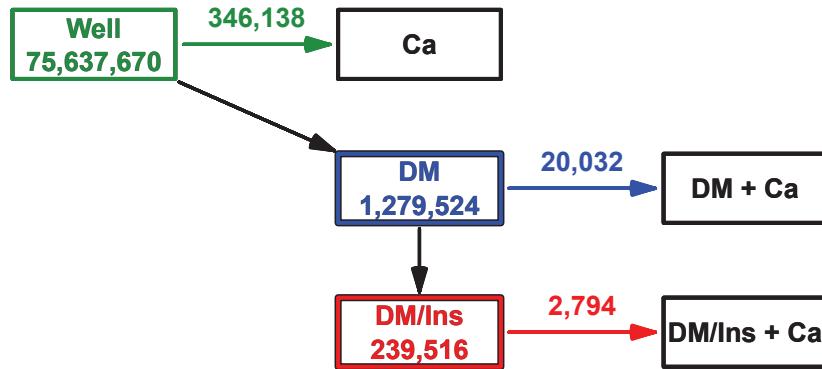
Follow-up in the population

Persons are followed 1 Jan 1995 to:

event: first primary cancer of a given type

censoring:
- diagnosis of any other primary cancer
- death
- 31 Dec 2009


DM prevalent at 1.1.1995 excluded
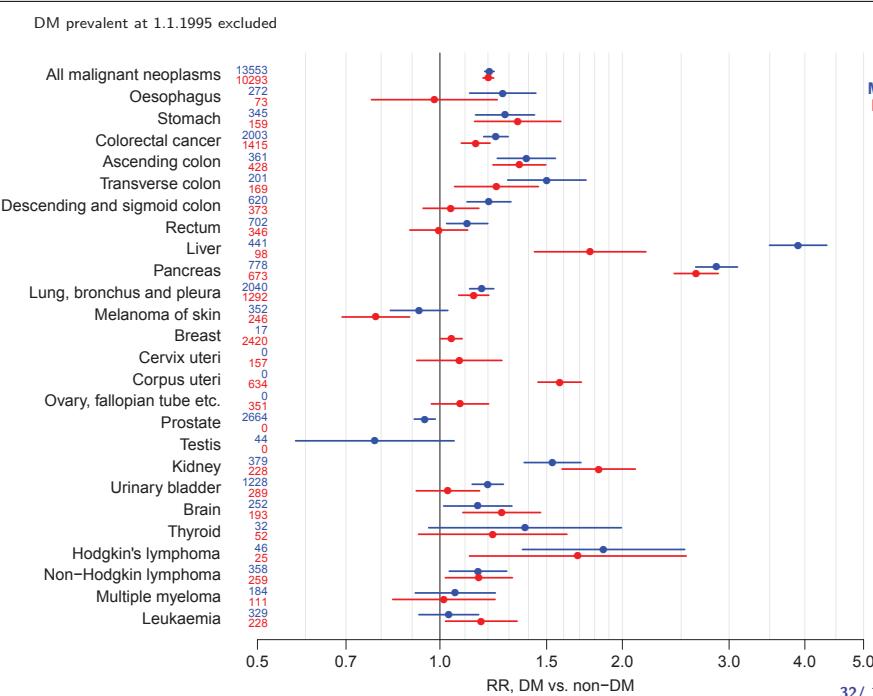
Duration model

$$\text{rate} = f(\text{age}) \times g(\text{date of FU}) \times h(\text{date of birth})$$
$$\times t(\text{DM-duration})$$
$$\times s(\text{Ins-duration})$$

All malignant neoplasms

Rate ratio DM, DM+Ins vs No DM

Diabetes duration (years)

Colorectal cancer

Rate ratio DM, DM+Ins vs No DM

Diabetes duration (years)

Lung, bronchus and pleura
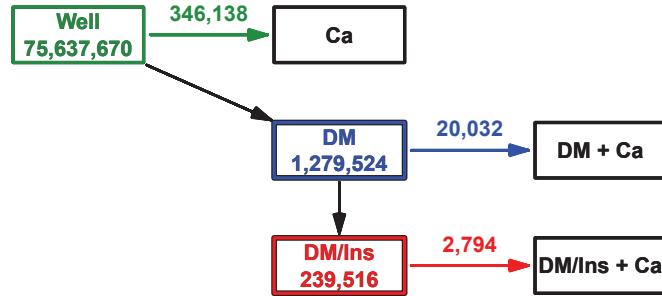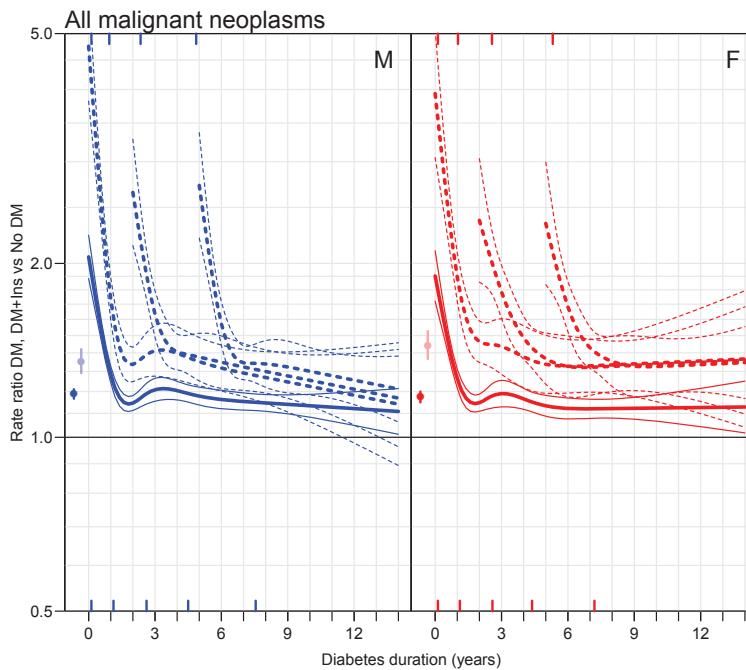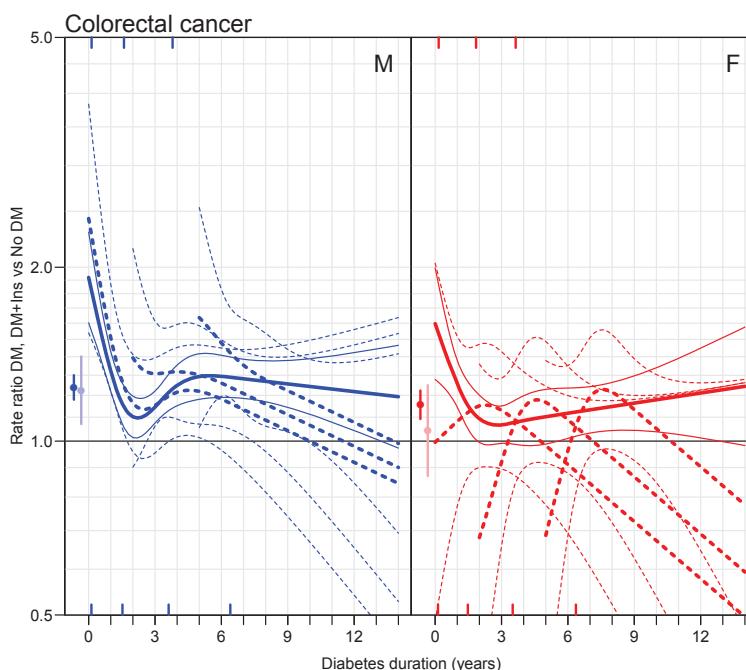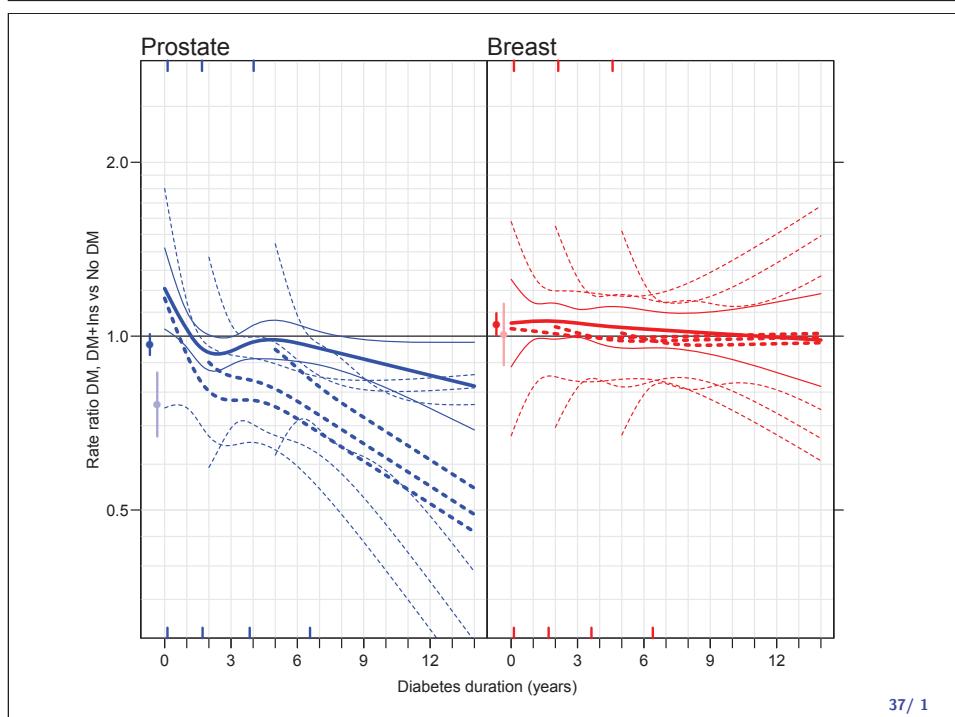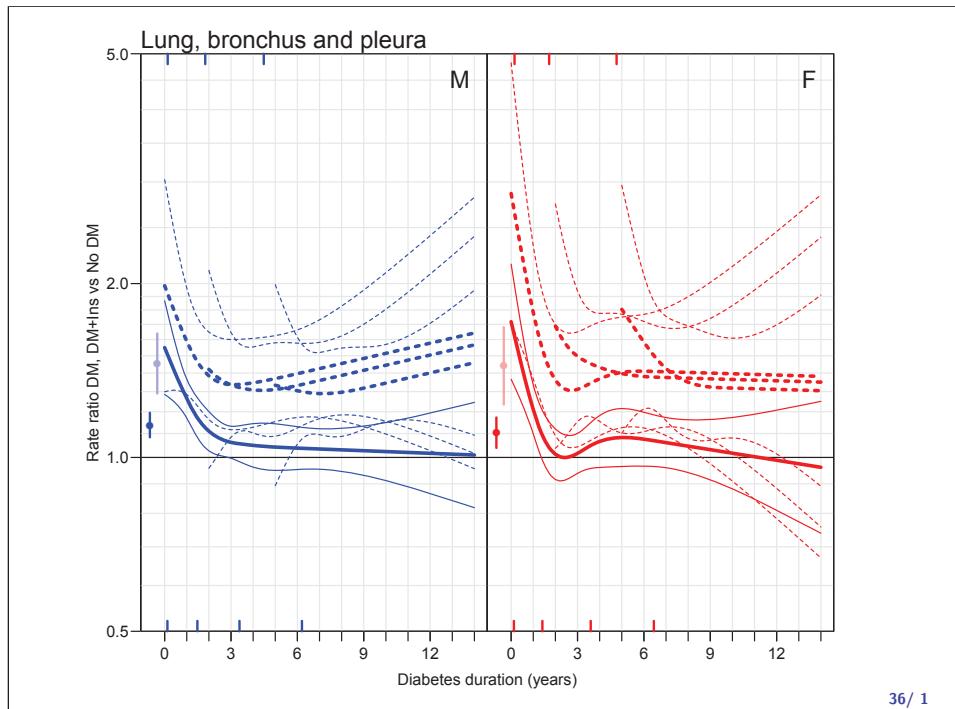
DIABETES AND OTHER DISEASES-EMERGING ASSOCIATIONS (JJ NOLAN, SECTION EDITOR)

The Epidemiology of Diabetes and Cancer

Bendix Carstensen · Marit Eika Jørgensen · Søren Friis

Abstract The literature on cancer occurrence in persons with diabetes has almost invariably been concerned with relative measures. In this paper, we briefly review this, but the aim is to quantify the absolute occurrence of diabetes and cancer in the population in order to give a fuller picture, which also includes the competing mortality risk. Overall, we find that some 35 % of the population will have a diagnosis of diabetes in their lifetime, 44 % a diagnosis of cancer, and about 15 % will have both diagnoses. The impact of differing mortality between persons with and without diabetes is illustrated by the fact that a person without diabetes at age 50 has a smaller lifetime risk of cancer than a person aged 50 with diabetes. Thus, the differences in cancer occurrence between persons

Introduction

The link between diabetes and cancer occurrence is well established, and comprehensive population-based studies have demonstrated that the association relates to both cancer incidence and mortality [1–3]. Recently, an increasing number of studies have examined cancer incidence among patients with diabetes, particularly following the report in 2009 of a potential association between the insulin analog glargine and cancer risk [4–7]. The majority of the studies have focused on comparisons of cancer incidence among diabetes patients using different antidiabetic regimes. However, these studies are prone to bias due to confounding by indication, as illus-

Demography: Life time risk
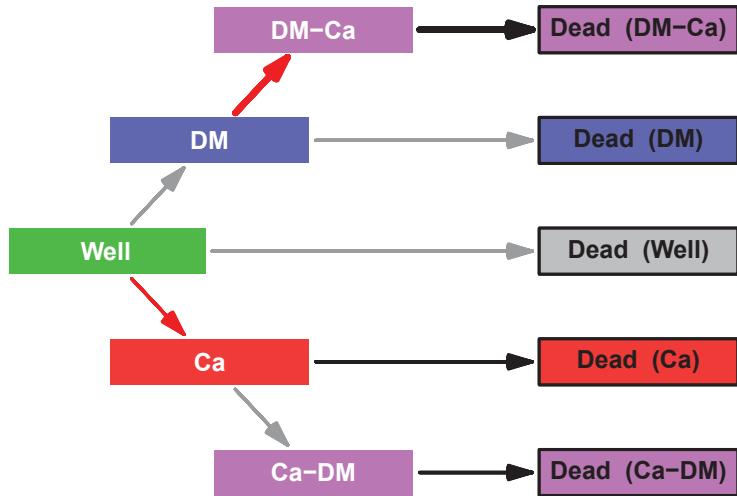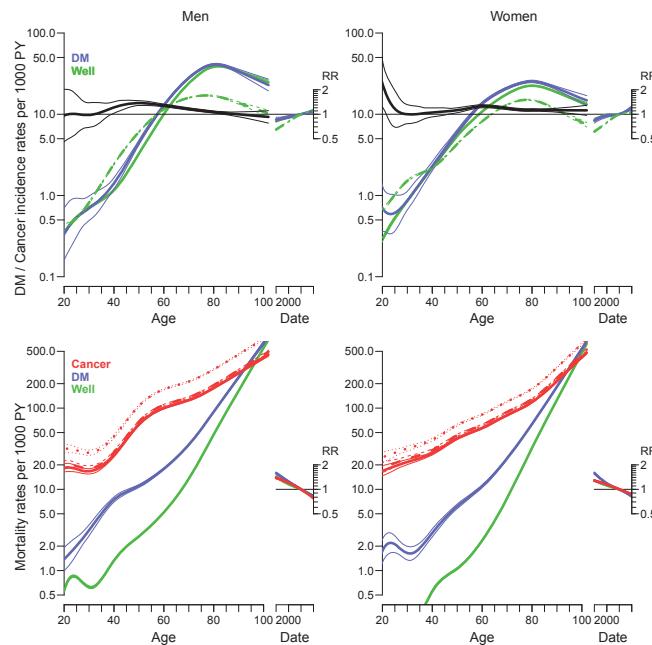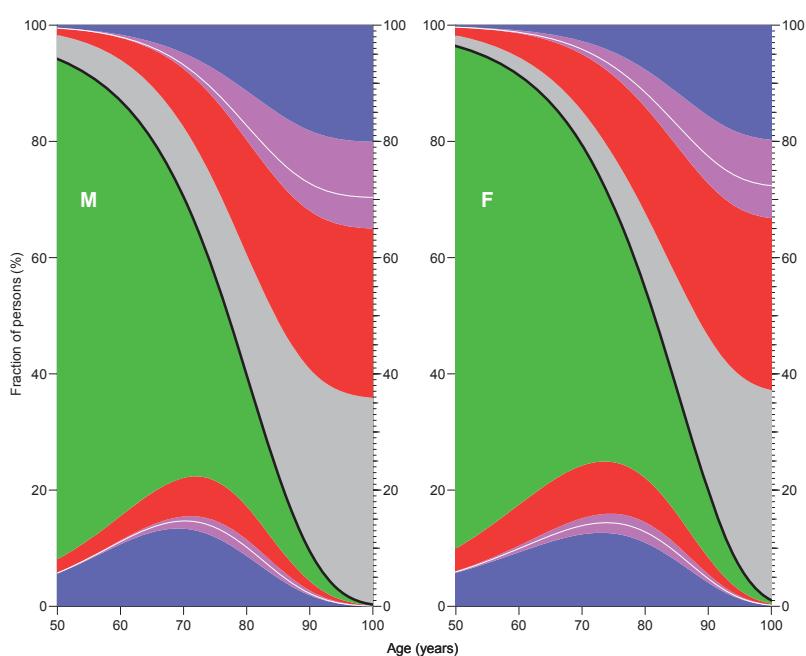
# Demography: Cumulative risk