

Cancer in patients admitted to hospital with diabetes mellitus aged 30 years and over: record linkage studies

C. J. Wotton · D. G. R. Yeates · M. J. Goldacre

Received: 28 July 2010 / Accepted: 22 October 2010 / Published online: 30 November 2010
© Springer-Verlag 2010

Abstract

Aims/hypothesis The aim of this study was to determine the risk of cancer in people admitted to hospital for diabetes mellitus when aged 30 or older.

Methods This study involved the analysis of two statistical datasets of linked hospital and mortality data, in an area in southern England, between 1963 and 1998 (the Oxford Record Linkage Study, ORLS1) and between 1999 and 2008 (ORLS2). Rates of cancer in the diabetes cohorts were compared with rates of cancer in reference cohorts and expressed as rate ratios.

Results The rate ratio for all cancer in people admitted to hospital with diabetes was 1.01 (95% CI 0.95–1.06, based on 15,898 people with diabetes) for the years 1963–1998; and 1.09 (1.00–1.19, based on 7,771 people with diabetes) in the years 1999–2008. In both datasets, there were significantly high rate ratios for cancers of the liver (ORLS1 and ORLS2, respectively, 2.0 [95% CI 1.4–2.9]; 2.5 [95% CI 1.3–4.3]), pancreas (2.2 [95% CI 1.8–2.7]; 3.5 [95% CI 2.5–4.8]) and uterus (1.5 [95% CI 1.0–2.2]; 2.6 [95% CI 1.4–4.5]). There were significantly low rate ratios for cancer of the prostate (0.6 [95% CI 0.5–0.7]; 0.7 [95% CI 0.5–0.9]) and non-melanoma skin cancer (0.6 [95% CI 0.5–0.8]; 0.8 [95% CI 0.6–0.96]).

Conclusions/interpretation Diabetes mellitus was associated with an elevated risk of some site-specific cancers and a reduction of risk of others. Considering the risk in diabetes of all cancers combined, the elevation of risk, if any, is

likely to be small and numerically less important than other known complications of diabetes.

Keywords Cancer · Diabetes mellitus · Record linkage · Type 2

Abbreviations

HES Hospital episode statistics
ICD International Classification of Disease
NHS National Health Service
OPCS Office of Population, Censuses and Surveys
ORLS Oxford Record Linkage Study

Introduction

Interest in cancer risk in people with diabetes mellitus has focused, in particular, on an increased risk of pancreatic cancer. Meta-analyses of studies of pancreatic cancer in type 2 diabetes have shown an association [1, 2], as has a recent record linkage study from Sweden [3]. It is unclear whether diabetes is a risk factor for pancreatic cancer or a consequence of it. Meta-analyses, of varying quality, have been published on diabetes and cancers of the colorectum [4], breast [5], prostate [6], endometrium [7], bladder [8] and non-Hodgkin's lymphoma [9]. Elevated risks of colorectal [4], liver [10, 11] and kidney [10] cancers have been reported, and there is evidence of an inverse association with prostate cancer [3, 6]. An important outstanding prognostic question, for people with diabetes mellitus, is whether diabetes is associated with an elevated risk of cancer, overall, as well as with an increased risk of some specific cancers and a reduced risk of others. There is also current interest in cancer risk in diabetes because

C. J. Wotton (✉) · D. G. R. Yeates · M. J. Goldacre
Unit of Health-Care Epidemiology, Department of Public Health,
University of Oxford,
Old Road Campus, Old Road,
Oxford OX3 7LF, UK
e-mail: clare.wotton@dphpc.ox.ac.uk

therapeutic agents used in its treatment, including metformin, insulin and sulfonylureas, have been implicated in altering cancer risk [12–15].

We have used the long-standing Oxford Record Linkage Study (ORLS) (1963–1998) and the Oxford subset of the more recent English national linked Hospital episode statistics (HES) (1999–2008) to investigate associations between diabetes mellitus in people aged 30 years and over for cancer overall, and for individual cancers. These datasets provide the means to study large numbers of different cancers in the same population, using the same methodology, over a long period of time.

Methods

Population and data The ORLS [16, 17] includes brief statistical abstracts of records of all hospital admissions (including day cases) in National Health Service (NHS) hospitals, and all deaths regardless of where they occurred, in defined populations within the former Oxford NHS region. The ORLS covered the years 1963 to 1998, and is referred to here as ORLS1. The hospital data were collected routinely in the NHS as the region's hospital statistics system and were similar to English national HES. The data for each individual were linked together, as they accrued, and are now anonymised and archived. A second dataset, referred to as ORLS2, has been linked and built as the Oxford regional subset of the English national HES system (version v08a), and runs from 1999 to March 2008. The data items available for linkage changed between 1998 and 1999 and the two datasets are not themselves linked together. The death data derive from death certificates. The population covered by the ORLS gradually expanded over time from an initial population of part of Oxfordshire (approximately 300,000 resident population) to all four counties of the former Oxford NHS region (resident population 2.5 million).

A cohort of people with an admission, or record of day care, for diabetes when aged 30 or older was constructed for those with diabetes mellitus recorded as the main reason for hospital care, by identifying the first admission, or episode of day case care, for the condition in an NHS hospital during 1963–1998 (ORLS1) or 1999–2008 (ORLS2). The International Classification of Disease (ICD) codes used for diabetes were 260 in ICD-7, 250 in ICD-8 and ICD-9 (www.icd9data.com/2007/Volume1/240-279/250-259/250/default.htm), and E10–E14 in ICD-10 (www.who.int/classifications/icd/en/). A comparison cohort was constructed by identifying the first admission for each individual with various medical and surgical conditions (listed in table footnotes) but excluding those with a record of diabetes. This is based on a 'reference' group of

conditions that has been used in other studies of associations between non-malignant diseases and cancer [17–22]. We followed the standard epidemiological practice, when hospital controls are used, of selecting a diverse range of conditions, rather than relying on a narrow range (in case the latter are themselves atypical in their risk of subsequent disease). People were included in the diabetes or reference cohort if they did not have an admission for cancer either before or at the same time as the admission for diabetes mellitus or the reference condition. We then searched the database for any subsequent NHS hospital care for, or death from, cancer in these cohorts. We considered that rates of cancer in the reference cohort would approximate those in the general population of the region while allowing for migration in and out of it (data on migration of individuals were not available). Type of diabetes was not reliably recorded in routine hospital statistics; we therefore used an age cut-off of '30 or older' at first recorded admission to reduce the proportion of the cohort that will have had type 1 diabetes. We chose 30 as our age cut-off as the only method available to us to minimise the inclusion of people with type 1 diabetes; it should be noted, nonetheless, that about 10% of adult diabetics have type 1 diabetes [23] and our study population is likely to be similarly composed.

Ethical approval for analysis of the record linkage study data was obtained from the Central and South Bristol Multi-Centre Research Ethics Committee (04/Q2006/176).

Statistical methods We calculated rates of each cancer based on person-years at risk. For each broad age group of people with diabetes, and for each individual cancer site, we took 'date of entry' into each cohort as the date of first admission for diabetes, or reference condition, and 'date of exit' as the date of first record of the cancer, death, or the end of the data file (31 December 1998 for ORLS1, 31 March 2008 for ORLS2), whichever was the earliest. In comparing the diabetes cohort with the reference cohort, we first calculated rates for each cancer in the diabetes and reference cohorts combined, as the standard cohort, stratified by age (in 5-year age groups), sex, calendar year of first recorded admission, and district of residence. We then applied the stratum-specific rates in the combined diabetes and reference cohorts to the number of people in each stratum in the diabetes cohort, separately, and then to those in the reference cohort, and calculated stratum-specific standardised rates in each cohort. We calculated the ratio of the standardised rate of occurrence of the cancer in the diabetes cohort relative to that in the reference cohort. The confidence interval for the rate ratio and χ^2 statistics for its significance were calculated as described elsewhere [24], as have further details of our methods [17]. Analyses were run using a

suite of programs developed in-house using SAS 9 software (SAS Institute, Cary, NC, USA).

In comparing the diabetes and reference cohorts, the precision of the rate ratio depends on the number of people with each subsequent cancer within each cohort. The size of the diabetes cohort is fixed by the number in the database with the condition. In the reference cohort, we included all the people in the database with the comparison conditions in each age group. We did this to maximise the numbers in each stratum in the reference cohort in order to maximise the precision of the rate ratios.

We studied the data with and without cases of cancer that occurred within 1 year of first recorded admission for diabetes mellitus. We did so with the aim of distinguishing longer term associations from those that might be explained by short-term ascertainment bias. In fact, we generally found no important difference between the rate ratios with or without inclusion of first year cases. Accordingly, we report the main results with all cases included.

Results

Table 1 shows the age distribution of the 15,898 people in the ORLS1, 7,771 people in the ORLS2 diabetes cohorts, and those in the respective reference cohorts, with the matching ratio of controls to cases in each age group. The matching ratio shows the number of people in the reference

cohort per person in the diabetes cohort. It shows that there were at least seven ‘matched’ cases in the reference cohort for every case of diabetes mellitus in ORLS1, at least 16 ‘matched’ cases in ORLS2, and many more in most groups. In ORLS1 (1963–1998), there were 23,779 people with cancer in the reference cohort, and an observed number of 1,355 people with cancer in the diabetes cohort compared with an expected number of 1,348. This gave a rate ratio in the diabetes cohort, compared with the reference cohort, of 1.01 (95% CI 0.95–1.06, Table 2). In ORLS2 (1999–2008), there were 9,770 people with cancer in the reference cohort and 560 people with cancer in the diabetes cohort, compared with an expected number of 516. This gave a rate ratio of 1.09 (95% CI 1.00–1.19, Table 2). These rate ratios were not significantly different from each other: the rate ratio for ORLS1 (1.01) falls within the confidence intervals of the ORLS2 result (1.00–1.19). Ten per cent of the outcome records for cancer in ORLS1 and 8% in ORLS2 were identified by linkage to death records alone; the remainder were identified by linkage to hospital records as well as to any corresponding death records.

Rate ratios and their 95% confidence intervals are shown in Table 2. There were significantly high rate ratios for cancers of the liver (2.0 in ORLS1, 2.5 in ORLS2), pancreas (respectively, 2.2, 3.5) and uterus (1.5, 2.6). There was a significantly high risk of cancer of the cervix (2.1) in ORLS1, but not in ORLS2. There was a significantly high risk of stomach cancer (2.1) in ORLS2, but not in ORLS1.

Table 1 Age distribution of study population of people admitted to hospital for diabetes mellitus

Age group (years)	ORLS1				ORLS2			
	Diabetes mellitus cohort		Reference cohort		Diabetes mellitus cohort		Reference cohort	
	Number (% men)	%	Number (% men)	Matching ratio ^a	Number (% men)	%	Number (% men)	Matching ratio ^a
30–34	719 (58)	4.5	35,656 (57)	50	460 (64)	5.9	18,218 (59)	40
35–39	641 (61)	4.0	31,320 (57)	49	499 (65)	6.4	20,043 (59)	40
40–44	732 (62)	4.6	28,895 (58)	39	564 (65)	7.3	18,270 (58)	32
45–49	844 (59)	5.3	28,033 (58)	33	525 (67)	6.8	16,515 (58)	31
50–54	1,185 (58)	7.5	27,451 (59)	23	658 (61)	8.5	17,441 (55)	27
55–59	1,436 (55)	9.0	26,037 (60)	18	726 (62)	9.3	17,963 (54)	25
60–64	1,765 (53)	11.1	24,715 (60)	14	859 (61)	11.1	16,709 (53)	19
65–69	2,113 (47)	13.3	22,093 (55)	10	879 (52)	11.3	15,856 (52)	18
70–74	2,292 (44)	14.4	19,309 (50)	8	903 (53)	11.6	14,830 (48)	16
75–79	1,937 (40)	12.2	15,370 (44)	8	758 (48)	9.7	13,061 (43)	17
≥80	2,234 (30)	14.1	16,685 (32)	7	940 (38)	12.1	16,217 (33)	17
Total	15,898 (48)	100	275,564 (55)	17	7,771 (56)	100	185,123 (53)	24

Data show the number and percentage of people in each age group at time of admission and percentage of those who were men

^a The number of people in the reference cohort per person with diabetes mellitus in each age stratum. Note that, in analysis, the diabetes and reference cohorts were stratified by age (to make them equivalent in respect of age) in calculating the age-standardised rates and rate ratios

Table 2 Occurrence of cancer in people admitted for diabetes mellitus at 30 years of age or older

Cancer (ICD code) ^a	ORLS1, ORLS2 (<i>n</i> ; <i>n</i>)	ORLS1 (1963–1998)				ORLS2 (1999–2008)			
		Observed number	Expected number	Adjusted rate ratio ^b (95% CI)	<i>p</i> value	Observed number	Expected number	Adjusted rate ratio ^b (95% CI)	<i>p</i> value
All cancers (140–208) ^c	23,779; 9,770	1,355	1,347.5	1.01 (0.95–1.06)	0.85	560	516.3	1.09 (1.00–1.19)	0.05
Oral cavity, pharynx (140–141, 143–146, 148–149)	370; 201	20	19.3	1.04 (0.63–1.63)	0.96	9	9.4	0.95 (0.43–1.84)	0.99
Larynx (161)	267; 80	9	13.3	0.67 (0.30–1.28)	0.29	3	3.7	0.80 (0.16–2.43)	0.91
Oesophagus (150)	838; 294	39	44.9	0.86 (0.61–1.19)	0.40	15	15.0	1.00 (0.55–1.68)	0.90
Stomach (151)	1,337; 230	93	84.3	1.11 (0.89–1.37)	0.36	25	12.9	2.05 (1.30–3.10)	0.001
Colon (153)	2,487; 596	152	144.4	1.06 (0.89–1.25)	0.54	49	41.6	1.19 (0.87–1.60)	0.27
Rectum (154)	1,489; 268	87	81.4	1.08 (0.85–1.34)	0.55	24	21.2	1.14 (0.72–1.74)	0.60
Liver (155)	309; 116	37	19.4	2.02 (1.40–2.85)	<0.001	14	6.1	2.46 (1.30–4.29)	0.002
Pancreas (157)	853; 267	121	58.8	2.21 (1.81–2.67)	<0.001	49	15.8	3.49 (2.52–4.75)	<0.001
Lung (162)	4,383; 1,057	244	235.8	1.04 (0.91–1.18)	0.61	63	61.2	1.03 (0.79–1.33)	0.87
Breast (women) (174)	2,412; 1,208	149	142.4	1.05 (0.88–1.24)	0.60	52	54.0	0.96 (0.71–1.27)	0.84
Cervix (180)	208; 38	27	13.9	2.07 (1.33–3.10)	<0.001	2	1.3	1.51 (0.18–5.83)	0.90
Uterus (182)	377; 118	35	23.8	1.52 (1.04–2.16)	0.02	15	6.2	2.60 (1.41–4.47)	0.001
Ovary (183.0)	476; 194	37	29.5	1.28 (0.89–1.78)	0.18	8	8.5	0.93 (0.40–1.88)	0.99
Prostate (185)	2,351; 1,271	73	121.4	0.59 (0.46–0.74)	<0.001	48	71.2	0.66 (0.49–0.88)	0.006
Kidney (189.0, 189.1)	498; 204	24	25.9	0.92 (0.59–1.39)	0.78	14	10.6	1.34 (0.72–2.30)	0.36
Bladder (188)	1,672; 619	59	86.1	0.67 (0.51–0.87)	0.003	35	33.8	1.04 (0.72–1.46)	0.90
Malignant melanoma (172)	364; 213	19	16.6	1.15 (0.68–1.82)	0.64	9	9.7	0.93 (0.42–1.79)	0.95
Other skin cancer (173)	1,833; 1,697	66	106.7	0.61 (0.47–0.77)	<0.001	65	86.1	0.75 (0.57–0.96)	0.02
Malignant brain (191)	424; 158	15	21.1	0.70 (0.39–2.98)	0.21	10	8.0	1.27 (0.60–2.39)	0.58
Bone (170)	149; 53	6	9.9	0.59 (0.21–1.32)	0.27	3	3.0	1.00 (0.20–3.07)	0.77
Non-Hodgkin's (200, 202)	794; 432	38	42.3	0.89 (0.63–1.24)	0.55	18	18.7	0.96 (0.56–1.54)	0.97
Multiple myeloma (203)	480; 170	23	28.3	0.80 (0.50–1.22)	0.36	8	8.7	0.91 (0.39–1.84)	0.94
Lymphoid leukaemia (204)	316; 172	22	20.3	1.09 (0.67–1.68)	0.78	10	8.8	1.15 (0.54–2.17)	0.80
Myeloid leukaemia (205)	359; 144	22	19.3	1.15 (0.71–1.76)	0.61	7	7.4	0.95 (0.37–2.00)	0.96
Benign brain (225)	228; 117	13	9.9	1.33 (0.70–2.32)	0.40	8	6.3	1.29 (0.54–2.63)	0.62

Data show number of people in the reference cohort with each cancer, and observed and expected number of people with cancer in the diabetes cohort, together with ratio of rate in the diabetes cohort to that in the reference cohort

Conditions used in reference cohort, with Office of Population, Censuses and Surveys (OPCS) code edition 3 for operations and ICD-9 code for diagnosis (with equivalent codes used for other coding editions): appendicectomy (OPCS3 441; excluded from analyses of colorectal cancers), tonsillectomy (230), knee arthroplasty (812), hip arthroplasty (810, 811), squint (ICD-9 378), otitis externa, otitis media (380–382), varicose veins (454), haemorrhoids (455; excluded from analyses of colorectal cancers), upper respiratory tract infections (460), deflected septum, nasal polyp (470–471), impacted tooth and other disorders of teeth (520–521), inguinal hernia (550; excluded from analyses of colorectal cancers), in-growing toenail and other diseases of nail (703), sebaceous cyst (706.2), internal derangement of knee (717), bunion (727.1), selected fractures (810–816, 823–826), dislocations, sprains and strains (830–839, 840–848), superficial injury and contusion (910–919, 920–924)

^a ICD-9 codes for each cancer (equivalent codes were used for cases coded in ICD Revisions 7, 8 and 10)

^b Adjusted for sex, age in 5-year bands, time period in single calendar years and district of residence

^c Results for all cancers were slightly different in men and women: ORLS1, males rate ratio (RR) 0.93 (0.86–1.00); females RR 1.10 (1.01–1.19); ORLS2, males RR 1.05 (0.93–1.17); females RR 1.15 (1.01–1.30)

Data not shown for cancers with non-significant results and five or fewer observed cases. The cancers that were studied in addition to those shown in the tables were those of the: salivary gland, nasopharynx, testis, nervous system, thyroid and Hodgkin's disease. Full results available on request from the authors

There were significantly low rate ratios for cancers of the bladder (0.7) in ORLS1, prostate (0.6, 0.7, respectively, in ORLS1 and ORLS2), and non-melanoma skin cancer (0.6, 0.8, respectively).

Time intervals In both ORLS1 and ORLS2 the rate ratios for cancer of the pancreas were significantly high within

5 years of diabetes admission, but not at longer time intervals (Table 3). In ORLS1, the risk of liver cancer decreased over time, with significantly high rate ratios within 5 years of diabetes admission, but not thereafter. In ORLS2, the rate ratio for liver cancer was significantly elevated at the longer time interval of 5 years or more since diabetes admission, but not within 5 years (Table 3). In

Table 3 Occurrence of selected cancers in people admitted for diabetes mellitus at 30 years of age or older

Time interval	ORLS1			ORLS2		
	Observed number	Expected number	Adjusted rate ratio (95% CI) ^a	Observed number	Expected number	Adjusted rate ratio (95% CI) ^a
Pancreas – all	121	58.8	2.21 (1.81–2.67)	52	14.7	4.00 (2.92–5.40)
<1 year	66	13.4	8.58 (6.04–12.2)	27	4.1	9.53 (5.74–15.5)
All <5 years	97	31.4	3.75 (2.95–4.73)	44	11.9	4.25 (3.00–5.91)
≥5 years	24	25.8	0.93 (0.59–1.40)	5	3.7	1.38 (0.43–3.42)
Liver – all	37	19.4	2.02 (1.40–2.85)	14	5.8	2.60 (1.38–4.54)
<1 year	13	3.3	5.65 (2.63–11.6)	4	1.0	4.36 (1.10–12.8)
All <5 years	26	9.2	3.27 (2.04–5.06)	8	4.8	1.72 (0.72–3.53)
≥5 years	11	9.6	1.16 (0.57–2.12)	6	2.3	2.94 (1.00–7.24)
Uterus – all	35	23.8	1.52 (1.04–2.16)	15	5.9	2.74 (1.49–4.71)
<1 year	6	4.0	1.58 (0.55–3.78)	6	1.7	4.02 (1.37–9.74)
All <5 years	26	12.9	2.20 (1.39–3.36)	11	4.7	2.46 (1.19–4.59)
≥5 years	9	10.8	0.83 (0.37–1.61)	4	1.1	4.74 (1.12–15.7)

Data show time between admission for diabetes and admission for cancer, observed and expected number of people with cancer in the diabetes cohort, and ratio of rate in the diabetes cohort to that in the reference cohort

^a Adjusted for sex, age in 5-year bands, time period in single calendar years and district of residence

ORLS1, the rate ratio for cancer of the uterus was significantly high within 5 years of diabetes admission, but not at longer time intervals. In ORLS2, the rate ratios for cancer of the uterus were significant both within 5 years of diabetes admission and at the longer time interval. There were no obvious temporal patterns for cancer of the cervix in ORLS1 and stomach cancer in ORLS2.

Men and women In ORLS2, most of the cases of liver cancer occurred in men. The rate ratio was significant in men (2.94 [95% CI 1.35–5.76]), but not in women (1.75 [95% CI 0.46–4.76]). These rate ratios were not significantly different from each other as one rate ratio falls within the confidence intervals of the other. There were no other notable differences between men and women in either ORLS1 or ORLS2.

Discussion

Principal findings The rate ratios for cancer overall in our study, at 1.01 and 1.09, are similar to that in a recent study in Scotland, which reported an unadjusted rate ratio of 1.01 (95% CI 0.92–1.11) and a rate ratio adjusted for deprivation of 0.99 (95% CI 0.90–1.09) [25]. They are also similar to those in record linkage studies from Denmark (rate ratio 1.1) [10] and Sweden (men 1.0, women 1.1, data on both sexes combined not given) [26]. The population of the Oxford region is comparatively healthy, as shown by health indicators such as standardised mortality ratios and life

expectancy [27], with a reasonably homogeneous population ethnically. More than 90% of the total population covered by ORLS1 was born in the UK. The population covered by ORLS2 was becoming more diverse, ethnically, but data on ethnicity or place of birth were not available in the dataset. A limitation of the datasets are that they do not include data on potential confounding variables such as social class, smoking, exercise or diet. However, the level of consistency seen across the studies noted above [10, 25, 26] makes it unlikely that our main finding, and theirs—no significant elevation of cancer risk, overall, in people aged 30 and over—is much influenced by unmeasured confounding.

The ORLS does not include data about therapy. Therapeutic agents have been implicated in altering the risk of cancer in diabetic patients, although evidence is inconsistent [12–15, 28]. Metformin [12], insulin [13–15] and sulfonylureas [15] have all been studied in an attempt to determine whether they alter the risk profile of cancer in people with type 2 diabetes. Our study is observational, as are the studies from Scotland [25], Denmark [10] and Sweden [26]. The best designs for studies intended to identify risks associated with individual treatments are randomised controlled trials. The contribution of our study is to show cancer risks in a large hospitalised population that will have been treated with the regimes that were typical in the times covered by ORLS1 and ORLS2, and to demonstrate that those risks are small.

Methodological issues: datasets, population, and multiple comparisons Strengths of the Oxford datasets include their

size and their potential for long follow-up, together with the opportunity to analyse the full range of individual cancers within a single, large population, and with the same study design. The datasets also have limitations. The cohorts are based on prevalent cases—the first recorded hospital admission or episode of day case care for each person with each condition—rather than being cohorts with follow-up from the date of first diagnosis. The datasets are limited to people who were admitted to hospital, or who received day case specialist care. This would not capture all people with diabetes, although it should identify the great majority with subsequent cancer. Data are not recorded on patients who move out of the area covered by data collection or who are treated in hospitals outside the area (mainly affecting ORLS1). Accordingly, absolute risks of cancer cannot be calculated, but relative risks can be, as shown by us in comparing the diabetes cohort with a reference cohort. In studies of cohorts of people with non-malignant disease, followed for subsequent cancer risk, and compared with the reference cohort, the reference cohort is known to give cancer rates that are highly plausible. Using the reference cohort as a comparator, we have reported on combinations of non-malignant disease and subsequent cancer that show no elevation or diminution of risks when none are expected [18–20]; high cancer rates when these are expected (e.g. high rates of colon cancer in people with inflammatory bowel disease [21], high rates of liver and pancreatic cancer in people with chronic non-malignant liver disease or pancreatitis [22]); and low rates when expected (e.g. low rates of some cancers in people with Parkinson's disease [17]).

Of people with the conditions covered by the reference cohort, 1.5% were excluded from the reference cohort because they had diabetes, i.e. there was a prevalence rate of 1.5% in those otherwise eligible to be in the reference cohort. Data from general practice show prevalence rates of diagnosed diabetes in the general population in the former Oxford NHS region. For example, the rates are 2.9% in the Oxfordshire Primary Care Trust area, 2.8% in West Berkshire and 3.0% in Milton Keynes [29]. Our 1.5% prevalence rate is an underestimate as it will not include people admitted to hospital outside of the timeframe of our study or those who have not been admitted to hospital at all with diabetes as the main reason for their admission. Our prevalence of hospitalised diabetes covers a time period from 1963 and, as diabetes prevalence is increasing year on year, there will also be some real differences in the prevalence of diabetes in our 'historical' study and in contemporary estimates. Furthermore, as we miss people with diabetes who are not hospitalised in the period covered by our study, it follows that the reference cohort is itself not a 'pure' cohort in respect of excluding people with diabetes. It includes people, unidentifiable by us, who have diabetes. Thus the comparisons shown are those between a cohort of

people, all of whom have diabetes, and a cohort of people in which maybe 2% of people have non-hospitalised and perhaps undiagnosed diabetes. The two datasets are not linked to each other because of changes in the data items available for linkage between 1998 and 1999. Consequently, it is likely that some people have been recorded as having a 'first admission' for diabetes in each of the time periods studied.

We used an age of 30 years or older as a proxy for type 2 diabetes. Although this group will mostly consist of people with type 2 diabetes, there will be some people with type 1 diabetes in the cohort as well, given that around 10% of the UK adult population has type 1 diabetes [23].

The definition of diabetes changed during the period covered by this study (1963–2008). This may have influenced the results in the two different periods studied (1963–1998 and 1999–2008). In 2000, the WHO introduced new guidelines that included the recommendation of a lower cut-off point for fasting plasma glucose for the diagnosis of diabetes. Lower admission thresholds for diabetes may have increased the number of people admitted to hospital with diabetes mellitus in the period of the ORLS2 dataset, compared with ORLS1.

We studied a large number of associations between diseases. The effect of making multiple comparisons needs to be considered. For this reason, we have given exact p values, as well as confidence intervals, so that the reader can judge the degree of significance of diabetes and each subsequent cancer. It is possible that some of the associations that are significant at a level of $p < 0.05$ or even $p < 0.01$ may result from making multiple comparisons and the play of chance. This may particularly be so where there is no prior hypothesis to support the finding. On the other hand, in a study with the number of comparisons that we have made, findings where the significance level is < 0.001 are very unlikely to be attributable to chance alone. With multiple comparisons, the play of chance can also work the other way round: rates that may truly be high can, by chance, be found to be lower than anticipated.

Individual cancers Some of the literature on the risk of specific cancers in people with diabetes mellitus is conflicting. Our findings for cancers of the pancreas, liver, uterus and prostate are similar to those reported in previous studies [1–3, 6, 7, 10, 11, 26]. The association between diabetes and elevated risks of pancreatic cancer is well established. A meta-analysis of 20 studies, performed in 1995, found a significant association between diabetes and pancreatic cancer with a pooled relative risk of 2.1 [1]. A further meta-analysis has reported on pooled risks of similar magnitude [2]. Elevated risks of pancreatic cancer may be confounded by obesity. However, two case-control studies found an elevated risk ratio for pancreatic cancer in people

with diabetes that did not materially change after adjustment for BMI [30, 31]. Although much of the risk of cancer of the pancreas in our data appears to be confined to the short-term, there is some evidence in other studies of a longer term risk of pancreatic cancer [2, 3]. In the ORLS2 dataset, there was an elevated risk of liver cancer 5 or more years after diabetes admission and a longer term risk of liver cancer has been reported by others [3]. Further studies of long term associations between type 2 diabetes and cancers of the pancreas and liver need to be done.

The association, within both datasets, with cancer of the uterus has been reported in a recent meta-analysis [7]. Any association with cancer of the uterus may be a consequence of confounding with shared risk factors, such as obesity, that may independently contribute to the development of both diabetes and endometrial cancer. For example, a recent study found that the risk of endometrial cancer in women with diabetes was increased if they were also obese or did not undertake physical exercise [32]. The associations with cancer of the cervix (ORLS1) and stomach (ORLS2) may be real (although an alternative explanation is that, though significant, they could be due to the play of chance) but they are unlikely to be a consequence of long-standing diabetes. A further possibility is that these associations, too, may be a consequence of confounding with shared risk factors. For example, chronic inflammation in response to infection has been implicated in the aetiology of both gastric (*Helicobacter pylori*) and cervical (human papilloma virus) cancers [33]; and low grade inflammation has been implicated in the risk of type 2 diabetes [34]. Rates of cervical cancer declined in the period and population covered by this study, as a result of screening and perhaps other factors. Our analysis, comparing the diabetes and reference cohorts, included adjustment by matching on year of admission in single years. Any effect of the decline in cervical cancer would have been the same in the diabetes and reference cohorts.

We found inverse associations between type 2 diabetes and both prostate cancer and non-melanoma skin cancer in both datasets. Previous studies have suggested a deficit of prostate cancer in people with diabetes [3, 6]. Although few other studies have reported on skin cancer, the Danish record linkage study found an inverse association in women (standardised incidence ratio [SIR] 0.9 [95% CI 0.8–0.9]), although not in men (1.0 [95% CI 0.9–1.1]) [10]. A study investigating non-melanoma skin cancer in type 2 diabetes concluded that the use of insulin in type 2 diabetes may provide a protective effect against this form of skin cancer [35]. It has also been suggested that insufficient levels of vitamin D may increase the risk of diabetes [36]. Because the production of vitamin D requires sunlight exposure, it is plausible that people who have relatively low exposure to sunlight may have a greater risk of diabetes as

well as a reduced risk of skin cancer. It is also possible that people with diabetes may spend less time in the sun than others; and an interpretation of the relationship between diabetes and skin cancer could only be meaningfully made with access to information about sun exposure.

Some previous meta-analyses have found elevated risks of colorectal [4], breast [5] and bladder cancers [8] and an association with non-Hodgkin's lymphoma [9]. We did not find significantly elevated risks of these cancers in the present study. Our findings for colorectal cancers, although not significant, were consistent with an elevated risk of up to 60% for colon cancer and up to 74% for rectal cancer, in ORLS2 (as judged by the upper confidence limit); these are consistent with the rate ratios found in the meta-analysis [4]. It is also possible that our rate, lower than anticipated from the literature, may be a 'chance' low (see discussion of multiple comparisons above). A lack of association between diabetes and breast cancer has been reported before [26]. It can be problematic to compare risks of different individual cancers using data from different studies with different regions. Record linkage studies seem to provide some of the most reproducible findings on individual cancer sites [3, 10, 11, 26, 32]. One of the reasons for this may be that many different cancer sites can be studied within a single population using the same methodology.

Conclusion This study adds to the evidence that type 2 diabetes is associated with elevated risks of cancers of the pancreas, liver and uterus, and to that of a possible inverse association with prostate cancer and non-melanoma skin cancer. Considering the risk in diabetes of all cancers combined, the elevation of risk, if any, is likely to be small and numerically less important than other known complications of diabetes.

Acknowledgements Over many years, the linked datafiles were built by L. Gill, G. Bettley and M. Griffith, Unit of Health-Care Epidemiology, University of Oxford. The Unit of Health-Care Epidemiology is funded by the English National Institute for Health Research to analyse the linked data.

Duality of interest The authors declare that there is no duality of interest associated with this manuscript.

References

1. Everhart J, Wright D (1995) Diabetes mellitus as a risk factor for pancreatic cancer. A meta-analysis. *JAMA* 273:1605–1609
2. Huxley R, Ansary-Moghaddam A, Berrington de González A, Barzi F, Woodward M (2005) Type-II diabetes and pancreatic cancer: a meta-analysis of 36 studies. *Br J Cancer* 92:2076–2083
3. Hemminki K, Li X, Sundquist J, Sundquist K (2010) Risk of cancer following hospitalization for type 2 diabetes. *Oncologist* 15:548–555

4. Larsson SC, Orsini N, Wolk A (2005) Diabetes mellitus and risk of colorectal cancer: a meta-analysis. *J Natl Cancer Inst* 97:1679–1687
5. Larsson SC, Mantzoros CS, Wolk A (2007) Diabetes mellitus and risk of breast cancer: a meta-analysis. *Int J Cancer* 121:856–862
6. Kaspar JS, Giovannucci E (2006) A meta-analysis of diabetes mellitus and the risk of prostate cancer. *Cancer Epidemiol Biomark Prev* 15:2056–2062
7. Friberg E, Orsini N, Mantzoros CS, Wolk A (2007) Diabetes mellitus and risk of endometrial cancer: a meta-analysis. *Diabetologia* 50:1365–1374
8. Larsson SC, Orsini N, Brisman K, Wolk A (2006) Diabetes mellitus and risk of bladder cancer: a meta-analysis. *Diabetologia* 49:2819–2823
9. Chao C, Page JH (2008) Type 2 diabetes mellitus and risk of non-Hodgkin lymphoma: a systematic review and meta-analysis. *Am J Epidemiol* 168:471–480
10. Wideroff L, Gridley G, Møller L et al (1997) Cancer incidence in a population-based cohort of patients hospitalized with diabetes mellitus in Denmark. *J Natl Cancer Inst* 89:1360–1365
11. Adami HO, Chow WH, Nyrén O et al (1996) Excess risk of primary liver cancer in patients with diabetes mellitus. *J Natl Cancer Inst* 88:1472–1477
12. Libby G, Donnelly LA, Donnan PT, Alessi DR, Morris AD, Evans JMM (2009) New users of metformin are at low risk of incident cancer. A cohort study among people with type 2 diabetes. *Diab Care* 32:1620–1625
13. Dejgaard A, Lynggaard H, Råstam J, Krosgaard Thomsen M (2009) No evidence of increased risk of malignancies in patients with diabetes treated with insulin: a meta-analysis. *Diabetologia* 52:2507–2512
14. Colhoun HM, SDRN Epidemiology Group (2009) Use of insulin glargine and cancer incidence in Scotland: a study from the Scottish Diabetes Research Network Epidemiology Group. *Diabetologia* 52:1755–1765, Erratum 52: 2469
15. Bowker SL, Majumdar SR, Veugelers P, Johnson JA (2006) Increased cancer-related mortality for patients with type 2 diabetes who use sulfonylureas or insulin. *Diab Care* 29:254–258
16. Goldacre M, Kurina L, Yeates D, Seagroatt V, Gill L (2000) Use of a large medical database to study associations between diseases. *QJM* 93:669–675
17. Fois AF, Wotton CJ, Yeates D, Turner MR, Goldacre MJ (2010) Cancer in patients with motor neuron disease, multiple sclerosis, and Parkinson's disease: record linkage studies. *J Neurol Neurosurg Psychiatry* 81:215–221
18. Goldacre MJ, Kurina LM, Wotton CJ, Yeates D, Seagroatt V (2005) Schizophrenia and cancer: an epidemiological study. *Br J Psych* 187:334–338
19. Goldacre MJ, Wotton CJ, Seagroatt V, Yeates D (2005) Cancer following hip and knee arthroplasty: record linkage study. *Br J Cancer* 92:1298–1301
20. Goldacre MJ, Wotton CJ, Seagroatt V, Yeates D (2006) Cancer and cardiovascular disease after vasectomy: an epidemiological database study. *Fertil Steril* 84:1438–1443
21. Goldacre MJ, Wotton CJ, Yeates D, Seagroatt V, Jewell D (2008) Cancer in patients with ulcerative colitis, Crohn's disease and coeliac disease: record linkage study. *Eur J Gastroenterol* 20:297–304
22. Goldacre MJ, Wotton CJ, Yeates D, Seagroatt V, Collier J (2008) Liver cirrhosis, other liver disease, pancreatitis and subsequent cancer: record linkage study. *Eur J Gastroenterol* 20:384–392
23. Department of Health (2007) About diabetes. Available from http://tna.europarchive.org/20080817162314/http://www.dh.gov.uk/en/Healthcare/NationalServiceFrameworks/Diabetes/DH_074762, accessed 15 October 2010
24. Breslow NE, Day NE (1987) Statistical methods in cancer research, Volume II. The design and analysis of cohort studies. IARC scientific publication No. 82. International Agency for Research in Cancer, Lyon, pp 103–115
25. Ogunleye AA, Ogston SA, Morris AD, Evans JMM (2009) A cohort study of the risk of cancer associated with type 2 diabetes. *Br J Cancer* 101:1199–1201
26. Adami HO, McLaughlin J, Ekblom A, Berne C, Silverman D, Hacker D et al (1991) Cancer risk in patients with diabetes mellitus. *Cancer Causes Control* 2:307–314
27. Association of Public Health Observatories—download data tables. Available from www.apho.org.uk/resource/view.aspx?RID=71961, accessed 8 November 2010
28. Hernández-Díaz S, Adami H-O (2010) Diabetes therapy and cancer risk: causal effects and other plausible explanations. *Diabetologia* 53:802–808
29. QOF Database (2007) Available from www.gpcontract.co.uk/sha.php?orgcode=Q38&year=7, accessed 15 October 2010
30. La Vecchia C, Negri E, Franceschi S, D'Avanzo B, Boyle P (1994) A case-control study of diabetes mellitus and cancer risk. *Br J Cancer* 70:950–953
31. Silverman DT, Schiffman M, Everhart J et al (1999) Diabetes mellitus, other medical conditions and familial history of cancer as risk factors for pancreatic cancer. *Br J Cancer* 80:1830–1837
32. Friberg E, Mantzoros CS, Wolk A (2007) Diabetes and risk of endometrial cancer: a population-based prospective cohort study. *Cancer Epidemiol Biomark Prev* 16:276–280
33. Moss SF, Blaser MJ (2005) Mechanisms of disease: inflammation and the origins of cancer. *Nat Clin Pract Oncol* 2:90–97
34. Crook M (2003) Type 2 diabetes mellitus: a disease of the innate immune system? An update. *Diabet Med* 21:203–207
35. Chuang T-Y, Lewis DA, Spandau DF (2005) Decreased incidence of nonmelanoma skin cancer in patients with type 2 diabetes mellitus using insulin: a pilot study. *Br J Dermatol* 153:552–557
36. Alpert PT, Shaikh U (2007) The effects of vitamin D deficiency and insufficiency on the endocrine and paracrine systems. *Biol Res Nurs* 9:117–129