Type 2 diabetes and risk of hospital admission or death for chronic liver diseases.

Sarah H Wild¹, Joanne R Morling¹, David A McAllister¹, Jan Kerssens², Colin Fischbacher², Julie Parkes³, Paul J Roderick⁴, Naveed Sattar⁵, Christopher D Byrne⁶,⁷ on behalf of the Scottish and Southampton Diabetes and Liver Disease Group and the Scottish Diabetes Research Network Epidemiology Group

¹Centre for Population Health Sciences, Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, UK

²Information Services Division, National Health Service National Services Scotland, Edinburgh, UK

³Faculty of Medicine, University of Southampton, UK

⁴Primary Care and Population Sciences, University of Southampton, UK

⁵British Heart Foundation Centre for Cardiovascular Science, University of Glasgow, UK

⁶Nutrition and Metabolism, Faculty of Medicine, University of Southampton, UK

⁷NIHR Southampton Biomedical Research Centre, University Hospital Southampton, UK

Further members of the Scottish and Southampton Diabetes and Liver Disease Group (John Dillon, Peter Hayes, David Macfarlane, John Petrie, Nick Sheron) and the Scottish Diabetes Research Network epidemiology group (Helen Colhoun, Robert Lindsay, Graham Leese, Helen Looker, John McKnight, Rory McCrimmon, Sam Philip)

Corresponding author:
Professor Sarah Wild, Centre for Population Health Sciences, Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Teviot Place, Edinburgh, EH8 9AG, UK. Tel: +44 131 651 1630 Fax: +44 131 650 6868. Email: sarah.wild@ed.ac.uk.

Word count including abstract, references, tables and figure legends: 5642

Number of figures and tables: 2 figures and 5 tables
List of abbreviations in order of appearance:

T2DM – type 2 diabetes
CLD – chronic liver disease
ALD – alcoholic liver disease
NAFLD – non-alcoholic fatty liver disease
SES – socio-economic status
RR – rate ratio
HCC – hepatocellular carcinoma
NASH - non-alcoholic steatohepatitis
SIMD –Scottish Index of Multiple Deprivation
T1DM – type 1 diabetes
BMI –body mass index
SCI-DC – Scottish Care Information – Diabetes Collaboration

Keywords: diabetes; liver disease; cohort

Conflict of interest: none

Financial support: this work was supported by funding from the Scottish Government through the Scottish Diabetes Group

Author contributions: JRM and CDB performed the literature search, SHW, JRM, DAM, JK, CF, NS and CDB designed the study and the analysis plan, data were acquired as part of routine clinical care, JK prepared the data, SHW analysed the data and prepared the figures, SHW and CDB wrote the first draft of the manuscript and all authors contributed to data interpretation, critically reviewed the manuscript for important intellectual content and approved the final version. The work was performed on behalf of the Scottish and Southampton Diabetes and Liver Disease Group and the Scottish Diabetes Research Network Epidemiology Group.
Abstract

**Background & Aims:** The impact of type 2 diabetes (T2DM) on hospital admissions and deaths due to common chronic liver diseases (CLDs) is uncertain. Our aim was to investigate associations between T2DM and CLDs in a national retrospective cohort study and to investigate the role of sex and socio-economic status (SES).

**Methods:** We used International Classification of Disease codes to identify incident alcoholic liver disease (ALD), autoimmune liver disease, haemochromatosis, hepatocellular carcinoma, non-alcoholic fatty liver disease (NAFLD) and viral liver disease from linked diabetes, hospital, cancer and death records for people of 40-89 years of age in Scotland 2004-2013. We used quasi Poisson regression to estimate rate ratios (RR).

**Results:** There were 6667 and 33624 first mentions of CLD in hospital, cancer and death records over ~1.8 and 24 million person years in people with and without T2DM, respectively. The most common liver disease was ALD among people without diabetes and was NAFLD among people with T2DM. Age adjusted RR for T2DM compared to the non-diabetic population (95 % confidence intervals) varied between 1.27 (1.04-1.55) for autoimmune liver disease and 5.36 (4.41-6.51) for NAFLD. RRs were lower for men than women and for more compared to less deprived populations for both ALD and NAFLD.

**Conclusions:** T2DM is associated with increased risk of hospital admission or death for all common CLDs and the strength of the association varies by type of CLD, sex and SES. Increasing prevalence of T2DM is likely to result in increasing burden of all CLDs.

**Keywords:** diabetes; liver disease; cohort
Introduction

Mortality from liver disease in the United Kingdom (UK) has increased by 400% since the 1970s, in contrast to reductions in mortality from cardiovascular disease and some cancers [1]. Similar increases in CLD mortality have also been observed in the United States (US)[2] although cirrhosis mortality has declined by 50-60% in France and Italy between 1980 and 2010[3]. Incidence and prevalence of common chronic liver diseases (CLDs) including alcohol-related liver disease (ALD) and obesity-related non-alcoholic fatty liver disease (NAFLD) and their complications have increased in the UK and US [4, 5]. Age-standardised incidence and mortality from hepatocellular carcinoma (HCC), a complication of many CLDs, have increased approximately two-fold in the UK in the last 40 years[4] and in other European countries (see http://globocan.iarc.fr/old/FactSheets/cancers/liver-new.asp). It is well known that type 2 diabetes (T2DM) is associated with increased risk of the whole spectrum of NAFLD and its complications including non-alcoholic steatohepatitis (NASH), cirrhosis and HCC[6-9], but the effect of T2DM on other common types of CLD is less clear.

Estimating the burden of ill health attributable to CLDs is difficult due to the lack of accurate, non-invasive diagnostic tests that can be applied within population-based studies. The available tests have been predominantly validated in high prevalence (secondary care) populations and may be less accurate in the wider population. Routine healthcare data obtained from hospital admission, cancer registration and death records provides a source of readily available data that have been used previously to assess the burden of serious gastro-intestinal disease in the US[5]. Although hospital admission and cancer registration data are likely to underestimate
prevalence of disease they provide valuable information for health care planning and an estimate of the burden of serious liver disease.

Increasing global prevalence of T2DM and the inverse influence of socio-economic status (SES) on T2DM prevalence in high income countries are well established [10, 11]. The inverse relationship between SES and ALD has also been described in several populations[12, 13] and the incidence of HCC in England between 1990 and 2009 was higher among people living in materially deprived areas than among people living in less deprived areas[14]. However, it is not known whether SES influences the association between T2DM and CLDs, as it does for T2DM and cardiovascular disease [15]. The need for further research on the contribution of SES to disparities in incidence, treatment and outcomes of liver disease has previously been identified[16, 17].

The aim of this retrospective cohort study in a national population of approximately five million people was: a) to describe numbers and rates of first record of CLD hospital admission or death for people with and without T2DM and, b) to investigate associations between T2DM and CLD hospital admissions and deaths, establishing whether associations differed by sex or SES.

**Methods**

**Datasets**

We used data from the population-based Scottish diabetes register linked to national hospital, cancer and death records in a retrospective cohort study design. The Scottish diabetes register is derived from primary and secondary care records for
people with diabetes diagnosed in normal clinical practice using blood glucose measurements and/or HbA1c with coverage of >99% since 2004[18] the start of the study period. Validation of the diabetes register for clinical practice occurs through responses to annual invitations to diabetic retinopathy screening. For example, in 2010 191,571 individuals were invited for screening of whom 2% were suspended from the screening programme due to the person not having diabetes (see http://www.sci-diabetes.scot.nhs.uk/project-overview/). In addition, validation of the research extract, performed against hospital admissions mentioning diabetes[18] or against sulphonylurea prescriptions (unpublished data) suggests that over 99% of people with diagnosed diabetes in Scotland are included in both clinical and research databases.

Definition of CLD
CLD was identified from the first mention in any position (primary/underlying or secondary diagnosis or cause of death) in national hospital in-patient, cancer and mortality records and grouped into categories using codes from the ninth (ICD-9) and tenth [ICD-10] International Classification of Diseases revisions as follows:

- **ALD:** (571.0), (571.2), (571.3), [K70];
- autoimmune hepatitis and primary biliary cirrhosis: (571.4), (571.6), [K75.4, K74.3];
- hemochromatosis: (275.0) [E83.1];
- hepatocellular carcinoma (HCC): (155.0); [C22.0];
- NAFLD: a category incorporating the following ICD-defined categories: other chronic non-alcoholic liver disease: (571.8), non-alcoholic fatty liver disease: [K76.0], non-alcoholic steatohepatitis: [K75.8], cirrhosis (571.5) [K74.6], hepatic fibrosis or sclerosis or fibrosis with sclerosis (571.9) [K74.0, K74.1,
K74.2] or portal hypertension: (572.3) [K76.6] that was not attributed to any of the other liver disease categories.

- viral hepatitis: (D070.3), (D070.5), (D070.9); [B16, B17, B18];

The study size was determined by the number of incident CLDs occurring in the national population in the study time period which ended in December 2013, based on the most recent data available to the research team.

An area-based measure of SES is assigned to residents of Scotland on the basis of where they live using the Scottish Index of Multiple Deprivation (SIMD) (for more information see [http://www.scotland.gov.uk/Topics/Statistics/SIMD](http://www.scotland.gov.uk/Topics/Statistics/SIMD)). SIMD 2009 combines 31 indicators across seven domains: income/employment, health, education, housing, geographic access, and crime. The index is generated from a weighted sum of the seven domain scores for each area defined by postcodes which contains a median of 769 people. Quintiles of the index were used to define approximately equal fifths of the distribution at a national level, and Q1 and Q5 were used to identify the most and least deprived quintiles (or fifths) respectively.

**Study design and data analysis**

We used a dynamic population-based retrospective study design using aggregated data for the total population of Scotland and individual level data on people with diabetes. Inclusion criteria for estimation of the person-years denominator were being aged 40 to 89 years in Scotland between 2004 and 2013. The only exclusion criteria were the approximately 1% of hospital admissions or death records with missing SIMD data and the events and person-years for people with type 1 diabetes.
For each CLD group, with the exception of NAFLD, the first recorded event was identified independently (so, for example, a record of hemochromatosis followed by a subsequent record of viral hepatitis contributes to event counts and person-years denominator for both disease groups). Event counts and person-years at risk for NAFLD were identified by excluding records with any mention of ALD, viral disease, hemochromatosis or auto-immune liver disease. Records that mentioned HCC and NAFLD and no other specific CLD were counted in both HCC and NAFLD groups.

Routine national healthcare data were used to identify CLD events occurring between January 1 2004 and December 31 2013 for adults aged 40-89 years. For each CLD, people with prevalent disease prior to January 2004 defined by the presence of a relevant disease code in hospital or cancer records in the previous 10 years were excluded. Follow-up time was censored for each CLD at first event or death from any cause. We included deaths from CLDs and HCC where they represented the record of CLD or HCC and there was no previous hospital or cancer record of CLD or HCC. Numbers of CLD events, person-years at risk and rates were estimated by diabetes status, age, sex, and SIMD quintile.

Mid-year population estimates for Scotland by age, sex, calendar year, and SIMD quintile for the relevant years were obtained from National Records of Scotland. Data for the non-diabetic population were identified by subtracting numbers of incident events and person-years at risk for the population of people with type 1 diabetes (T1DM) and T2DM from numbers of CLD events and person-years at risk calculated from mid-year populations for the whole of Scotland. Type of diabetes was based on clinical records and a validation algorithm that used information on age at diagnosis
of diabetes with additional validation using dates and type of diabetes treatment, dates of first HbA1c>6.5% (48mmol/mol) and invitation to diabetic retinopathy screening[19].

Absolute incidence rates were used to describe patterns of different CLDs in the study population, with direct standardisation for age, stratification by sex and, where appropriate, for SIMD quintile, using the 2007 mid-year population of Scotland as the standard. Rate ratios (RR) for people with T2DM compared to people without diabetes were estimated using quasi Poisson regression modelling to control for potential confounding by age and to allow for the over-dispersed nature of the distribution of the data. Likelihood ratio tests were used to test for interactions between T2DM and sex and T2DM and SES. Analyses were stratified by sex and also for SES where appropriate. Exploratory analyses were performed among people with type 1 diabetes.

Sensitivity analyses were performed to identify whether the associations with T2DM altered if the following alternative outcome definitions were used

1. primary diagnoses/ underlying cause of death of ALD and NAFLD (excluding mention of ALD or NAFLD in secondary positions in hospital and death records)

2. hospital admission records mentioning ALD and NAFLD in any position (excluding deaths with no prior record of hospital admission)

3. hospital and death records mentioning chronic non-alcoholic liver disease: (ICD-9 571.8), non-alcoholic fatty liver disease: [ICD-10 K76.0], and non-alcoholic steatohepatitis: [ICD-10 K75.8] in any position (excluding
Approval for the creation and analysis of the linked dataset containing no personal identifying information was obtained from the Scottish Care Information - Diabetes Collaboration (SCI-DC) steering committee, the Scottish multi-centre research ethics committee (reference number 11/AL/0225), the Privacy Advisory Committee of National Health Service National Services Scotland and Caldicott guardians. Stata version 13 (College Station, Texas) was used for the analyses.

Results
During approximately 26 million years of follow-up in total there were a total of 40291 first records of CLD events identified from hospital or death records in people aged 40 to 89 years in Scotland in the decade between 2004 and 2013. There were 6667 events in people with T2DM during 1.8 million person-years of follow-up. Approximately 61% of CLD records were identified in men regardless of diabetes status whereas men represent 54% of the population of people with T2DM in Scotland over 40 years of age and 47% of the whole population in this age group in 2008. The majority (97.4%) of people with a record of a specific incident CLD had only one of ALD, auto-immune liver disease, hemochromatosis, HCC, or viral liver disease mentioned in their records.
Table 1 shows the numbers of first records of CLDs and age-standardised rates by diabetes status, sex and type of liver disease. The majority of outcomes were identified from hospital records or cancer registrations as shown in table 2. The highest CLD rates were for ALD for people without diabetes and for NAFLD for people with T2DM. The range of rates across SIMD quintiles is given for ALD and NAFLD because of the statistically significant variation in incidence by SES (see figures 1 and 2). People with T2DM were older than those without diabetes at the time of first CLD event and a large proportion of CLD events occurred among people of working age (data not shown but available from authors).

Age-standardised rates by sex, diabetes status and SES for ALD and NAFLD are shown in figure 1 and figure 2 respectively and illustrate higher rates in more compared to less deprived populations and a more marked socio-economic gradient for ALD than NAFLD. Age-adjusted RRs for people with T2DM compared to people without diabetes for ALD and NAFLD varied with sex and SES (Table 3). We identified a statistically significant interaction between diabetes status and SES for ALD and NAFLD but there was no evidence of similar interaction for other CLDs. RRs for ALD and NAFLD were lower among more deprived than among less deprived populations. There was a statistically significant interaction between T2DM and sex for NAFLD as a consequence of markedly higher relative risk of NAFLD among women than men (p<0.0001), but no evidence for an interaction between T2DM and sex for ALD (p=0.870).

Table 4 shows adjusted RRs for each CLD for people with T2DM compared to people without diabetes. Further adjustment for SES gave similar estimates of RRs
to age-adjusted results suggesting minimal confounding by SES. T2DM was associated with statistically significantly higher risk of all types of CLD apart from viral liver disease in men. The highest RRs associated with T2DM were for HCC and NAFLD. Summary estimates of RRs are shown for ALD and NAFLD to allow comparison with RRs for other liver diseases although the interaction illustrated in the figures and reported in table 3 should be noted. The sensitivity analyses for different outcome measures for ALD and NAFLD are reported in Table 5 and show similar RRs for the associations with T2DM as for those reported for the original analyses.

Discussion

This national retrospective cohort study of a contemporary Scottish population studied over a decade between 2004 and 2013. The absolute rates and relative risks associated with T2DM varied markedly by type of CLD. The strength of the associations between T2DM and CLD varied with sex for NAFLD and with SES for both ALD and NAFLD. In keeping with previous data, T2DM was associated with particularly high relative risks of HCC[20] and NAFLD[1], but we report for the first time that there is also increased risk associated with T2DM for all other types of CLD including ALD. There was only minimal confounding of the association between T2DM and CLD by SES but there was evidence of interaction between T2DM and SES for ALD and NAFLD. The absolute risk and the relative risk of NAFLD associated with T2DM were statistically significantly higher for women than men.
Socio-economic deprivation was associated with higher incidence of ALD regardless of diabetes status.

These findings have important implications for prevention and management of CLD and for the planning and the costs of health services, given the increasing prevalence of T2DM and the considerable cost of hospital admissions. Health care costs for patients with severe complications of CLDs are particularly high as illustrated by recent data for patients with various stages of liver disease related to hepatitis C from the US[21]. Costs per-patient-per-month over a mean of 634 days of follow-up between 2002 and 2010 amounted to $1,420, $1,870 and $4,931 for people with non-cirrhotic liver disease, compensated cirrhosis and end-stage liver disease respectively ($1 = 0.93 Euros November 2015). The annual costs to the National Health Service in England in 2007 for cirrhosis and liver cancer due to alcohol consumption alone was estimated to be £566 million[22] (£1 = 1.43 Euros November 2015).

The associations between T2DM and CLDs by sex and SES are likely to reflect differences in distributions of risk factors for both CLDs and for T2DM including obesity, fat distribution and alcohol intake. For example, we have previously reported higher mean body mass index (BMI) among women than men for a given age at diagnosis of T2DM and higher prevalence of obesity (BMI\geq 30kg/m2) in the most deprived compared to the least deprived fifthths of the population around the time of diagnosis of diabetes (54% compared to 44%)[23] In the population based Scottish Health Survey 2003, obesity prevalence was 34% in the most deprived quintile and 21% in the least deprived quintile[24]. Obesity and modest alcohol intake have
previously been found to interact to further increase risk of cirrhosis in a Scottish population[25]. The higher prevalence of obesity in deprived populations is therefore likely to contribute to higher prevalence of cirrhosis regardless of diabetes status and underlying alcoholic or non-alcoholic CLD etiology.

As both CLD and T2DM may have long asymptomatic phases prior to diagnosis, it is difficult to identify which disease develops first and to identify the true incidence of either disease and whether the association with CLDs and subsequent cirrhosis is bi-directional. Detailed characterisation in cohort studies with repeated measurements is required to determine whether individual CLDs with or without cirrhosis are associated with increased risk of developing subsequent T2DM. For example, hepatic steatosis identified by ultrasound has been identified as a risk factor for subsequent incident diabetes compared to people without hepatic steatosis[6], and resolution of hepatic steatosis is associated with reduction of risk of incident diabetes compared to people with persistent hepatic steatosis[7]. Indeed, it is estimated up to 70% of patients with T2DM have hepatic steatosis and increasing evidence indicates that increases in liver fat contributes to the pathogenesis of T2DM [26].

The strengths of our study include the very large population-based nature of the electronic record of diagnosed diabetes that captures data for >99% of the population of Scotland, validation of type of diabetes, and the availability of linkage to quality-assured hospital admission, cancer registration and mortality data for the whole population (for more information, see http://www.isdscotland.org/Products-and-Services/Data-Quality/Assessments/).
We identified first hospital admission and deaths in people in Scotland mentioning CLD whether or not this was recorded as the primary diagnosis in hospital or underlying cause of death in death records for our primary outcome measures. We believe that this is the most appropriate approach for identifying CLD from routine health care data at present because co-morbidities such as cardiovascular disease [27, 28] may be coded as the primary diagnosis in hospital records or underlying cause of death in people with CLDs. Accuracy and completeness of recording of secondary conditions on health records may differ by diabetes status which could introduce bias in either direction. People with diabetes may be more likely to have liver disease diagnosed as a consequence of larger numbers of health service contacts and are also be more likely to be admitted to hospital than people without diabetes but other co-morbidities may be recorded in preference to CLD among people with diabetes. Our sensitivity analyses identified that the associations with T2DM were similar when the outcomes were limited to primary diagnosis in hospital records or underlying cause of death suggesting that bias in recording did not influence our original findings. We are not able to describe the population-based incidence of each CLD in people with and without diabetes, because our measure of data capture only identifies non-fatal diagnoses of CLD recorded during a hospital admission. Consequently our data capture omits diagnoses established solely in primary or ambulatory care that are likely to represent less severe forms of each liver disease. Our identification of the first CLD event and the interpretations of the associations with diabetes are based on the assumption that most of the previous events can be identified in Scottish records and that proportions of events occurring outside Scotland in people with and without diabetes are similar.
Validation of coding of CLD and its complications is clearly required to facilitate population-based research. For example, the use of the narrow National Center for Health Statistics definition of CLD which excludes viral hepatitis and hepatobiliary cancer has been reported to systematically under-estimate liver-related mortality in the United States[30]. We found similar rate ratios associated with T2DM whether or not we included cirrhosis, hepatic fibrosis/sclerosis and portal hypertension in the NAFLD outcome but it is not clear without validation whether it would be appropriate to include these outcomes in estimates of NAFLD resource use. Validated coding of CLD would also improve future research including cohort studies of the association between different CLDs and HCC, cirrhosis, transplantation and mortality in people with and without diabetes and the investigation of effects of different treatments for diabetes on CLD outcomes.

There is potential for mis-classification of ALD and NAFLD which will co-exist in a large proportion of people in populations where alcohol consumption and obesity are common. The approach we have taken means that records that include both diagnoses are identified as ALD, with consequent under-counting of records of NAFLD. It is likely therefore that we have under-estimated the strength of the association between T2DM and hospital admissions and deaths for people with a record of NAFLD. There is an argument for combining CLD into a single category but we elected not to do this because of the different patterns of association between T2DM and different CLDs.

The use of an area-based measure of deprivation, rather than an individual-based measure of SES is a potential limitation of the study. However, the small average
population size of each area used to create the SIMD (median <800 people) means that in general it performs well as a measure of material deprivation. The inclusion of the health domain in SIMD does not seem to have a major effect on the pattern of health inequalities by SIMD[29] The predominantly white ethnic distribution of the Scottish population means that it is not feasible to investigate ethnic differences in CLD incidence. We are not able to describe the population-based incidence of each CLD in people with and without diabetes because we were not able to identify diagnoses of CLD established in primary or ambulatory care.

Key limitations of this analysis are that records of lifetime alcohol intake and measures of diet and physical activity are not available in this database and BMI is only available for people with diabetes. We were therefore not able to estimate what proportion of the apparent effect of T2DM could be explained by the association with obesity. The study had limited power to detect statistically significant interactions for less common liver diseases than ALD and NAFLD and to investigate the association between T1DM and CLD. However, based on 115 ALD events among people with T1DM in our population, we found age, sex and SES adjusted RRs (95% confidence intervals) associated with type 1 diabetes of 1.09 (0.69-1.72). There was a statistically significant interaction between type 1 diabetes and sex for the association with NAFLD, with age and SES adjusted RR of 3.03 (2.68-3.43) for men and 5.36 (4.41-6.51) for women based on 75 events in men and 78 in women.

Conclusions
These findings show that people with T2DM in Scotland are at higher risk of first hospital admission, hepato-cellular cancer registration or death from all types of common chronic liver diseases, compared to people without diabetes. Adjustment for an area-based measure of SES only modestly attenuated the contribution of T2DM to hospital admission or death from CLD. SES modified the associations between T2DM and ALD and NAFLD and sex modified the association between T2DM and NAFLD. We suggest that there may be a role for targeted case finding of CLD and appropriate intervention in high risk populations including people with T2DM. The increasing global prevalence of T2DM can be expected to result in an increasing burden of all CLDs.

Acknowledgements

This work was supported by funding from the Scottish Government through the Scottish Diabetes Group. Diabetes data for Scotland are available for analysis thanks to numerous healthcare staff who enter the data and people and organisations (people with diabetes, the Scottish Care Information – Diabetes Collaboration [SCI-DC] Steering Group, the Scottish Diabetes Group, the Scottish Diabetes Survey Group, the managed clinical network managers and staff in each Health Board) involved in providing data, setting up, maintaining and overseeing SCI-DC. CDB is supported in part by the Southampton National Institute for Health Research Biomedical Research Centre.
Figure Legends

Figure 1. Age-standardised incidence of hospital admissions and deaths for alcoholic liver disease per 100000 person years by sex, diabetes status (circles represent data for people without diabetes, squares for people with T2DM) and quintile of socio-economic status (SES, where 1 indicates the most deprived quintile)

Figure 2. Age-standardised incidence of hospital admissions and deaths for non-alcoholic fatty liver disease per 100000 person years by sex, diabetes status (circles represent data for people without diabetes, squares for people with T2DM) and quintile of socio-economic status (SES, where 1 indicates the most deprived quintile)
Reference List


