

# Matched population-based study examining the risk of type 2 diabetes in people with and without diagnosed hepatitis C virus infection

Schnier, Christian; Wild, S.; Kurdi, Z.; Povey, C.; Goldberg, David J.; Hutchinson, Sharon J.

*Published in:*
Journal of Viral Hepatitis

*DOI:*
10.1111/jvh.12520

*Publication date:*
2016

*Document Version*
Peer reviewed version

Title:
Matched population-based study examining the risk of type 2 diabetes in people with and without diagnosed hepatitis C infection

Running title:
Type 2 diabetes and HCV

Authors:
C. Schnier, School of Health and Life Sciences, Glasgow Caledonian University, Glasgow G4 0BA & Health Protection Scotland, Meridian Court, 5 Cadogan Street, Glasgow G2 6QE, UK
S. Wild, Centre for Population Health Sciences, University of Edinburgh, Edinburgh EH8 9AG
Z. Kurdi, Centre for Population Health Sciences, University of Edinburgh, Edinburgh EH8 9AG
C. Povey, Information Service Division, Scotland, 1 South Gyle Crescent, Edinburgh EH12 9EB, UK
D. J. Goldberg, Health Protection Scotland, Meridian Court, 5 Cadogan Street, Glasgow G2 6QE, UK
S. J. Hutchinson, School of Health and Life Sciences, Glasgow Caledonian University, Glasgow G4 0BA & Health Protection Scotland, Meridian Court, 5 Cadogan Street, Glasgow G2 6QE, UK

Corresponding Author:
Sharon Hutchinson
School of Health and Life Sciences
Glasgow Caledonian University
Cowcaddens Road
Glasgow
G4 0BA
sharon.hutchinson@gcu.ac.uk
Tel: +44(0)141 273 1949

# Matched population-based study examining the risk oftype 2 diabetes in people with and without diagnosed hepatitis C infection


## Abstract

Meta-analyses have found hepatitis C virus (HCV) infection to be associated with an increased risk of type 2 diabetes mellitus (T2DM). Here, we examine this association within a large population-based study, according to RNA status.

A data-linkage approach was used to examine the excess risk of diagnosed T2DM in people diagnosed with HCV-antibodies in Scotland (21,929 $Ab^{+ves}$; involving 15,827 $RNA^{+ves}$, 3927 $RNA^{-ves}$ and 2175 with unknown RNA-status) compared to that of a three-fold larger general population sample matched for sex, age and postcode (65,074 $Ab^{-ves}$). To investigate effects of ascertainment bias the following periods were studied: up to one year before (*pre*-HCV)/within one year of (*peri*-HCV)/more than one year post (*post*-HCV) the date of HCV-diagnosis.

T2DM had been diagnosed in 2.9% of $Ab^{+ves}$ (including 3.2% of $RNA^{+ves}$ and 2.3% of $RNA^{-ves}$) and 2.7% of $Ab^{-ves}$. A higher proportion of T2DM was diagnosed in the *peri*-HCV period (i.e. around the time of HCV-diagnosis) for the $Ab^{+ves}$ (22%) compared to $Ab^{-ves}$ (10%). In both the *pre*-HCV and *post*-HCV periods, only those $Ab^{+ves}$ living in less deprived areas (13% of the cohort) were found to have a significant excess risk of T2DM compared to $Ab^{-ves}$ (adjusted odds ratio in the *pre*-HCV period: 4.0 for females and 2.3 for males; adjusted hazard ratio in the *post*-HCV period: 1.5). These findings were similarly observed for both $RNA^{+ves}$ (chronic) and $RNA^{-ves}$ (resolved).

In the largest study of T2DM among chronic HCV-infected individuals to date, there was no evidence to indicate that infection conveyed an appreciable excess risk of T2DM at the population level.

**Keywords:**

Hepatitis C, Type 2 Diabetes, Matched cohort study, Data linkage

# 1 Introduction

A large consistent body of evidence from several observational studies suggests that Hepatitis C virus (HCV) infection is associated with

insulin resistance (IR) and Type 2 diabetes mellitus (T2DM (1-4). In addition, several plausible pathways have been suggested to explain how HCV influences IR and T2DM) (5-7). However, estimates of the size of the effect of HCV on T2DM risk vary between different studies. Two different meta-analyses of a total of 47 different studies both showed approximately 70% increased odds/hazards of having diabetes for individuals with HCV infection compared to individuals without HCV infection (adjusted Odds Ratio (OR), 1.7; 95% Confidence Interval (CI), 1.2-2.5 (3) ; and 1.7; 95% CI, 1.2-2.2 (8)). A recent population based cross sectional study from the US (9), however, found little evidence of increased risk to test diabetes positive in people with, compared to without, either current HCV-infection (OR, 1.1; 95% CI, 0.6-1.9) or with current or past HCV-infection (OR, 1.0; 95% CI, 0.6-1.7). In addition, a large population based cohort study from Southern Italy showed that compared to $HCV^{-ve}$ controls only people with HCV and elevated alanine aminotransferase (ALT) levels were at higher odds of developing T2DM (OR,1.5; 95% CI, 1.0-2.2), while those with HCV and normal baseline ALT levels were at lower odds (OR, 0.6; 95% CI, 0.3-1.1) (10). Another study of people enrolled in a community-based cohort in the US showed that HCV infection increased the risk of developing diabetes (adjusted hazard ratio (HR), 11.6; 95%CI 1.4-96.6),

but only among those at high risk of diabetes (based on body mass index and age) (11). Finally, a recent meta analysis suggested, on the basis of limited evidence, that having diabetes can also be a risk factor for contracting HCV (12).

The heterogeneity of findings from the different studies indicates that, at a population level, the effect of HCV on T2DM risk is comparably low and varies between different strata of the population. Therefore, studies to estimate the size of the effect of HCV on T2DM in the general population need to be sufficiently large to allow examination of different strata of the population and need careful control of confounding. Factors that increase the risk for diabetes and that might differ between those with HCV and those without HCV include low socioeconomic status (13,14), a history of heroin dependence (15) and methadone treatment (16), high alcohol consumption (17), smoking (18), increasing age (19), male sex (19), non white ethnicity (20) and higher body mass index (14).

To study the relationship between HCV infection and T2DM at a population level, we compared the risk of T2DM diagnosis in all people who have been diagnosed HCV antibody$^{+ve}$ with the risk of T2DM diagnosis in a three-times larger cohort of controls matched for the major confounding factors of sex, neighbourhood and age. To ascertain

wether any difference in the risk of T2DM was related to the virus itself or to factors associated with HCV-infection, we compared the relationship between HCV infection and T2DM in all people who had tested (i) HCV antibody$^{+ve}$, (ii) HCV antibody$^{+ve}$ and RNA$^{+ve}$ and (iii) HCV antibody$^{+ve}$ and RNA$^{-ve}$. To reduce the potential effect of ascertainment bias associated with being diagnosed for HCV infection, we studied three different periods of T2DM diagnosis: (i) a diabetes diagnosis at least 1 year *prior* to HCV diagnosis; (ii) within ±1 year of HCV diagnosis and; (iii) later than one year *post* HCV diagnosis.

# Patients and Methods

### Data sources for diagnosis of HCV and T2DM

Scotland has comprehensive national disease databases of people diagnosed with HCV-antibodies and of people diagnosed with diabetes. The database of people diagnosed with HCV-antibodies held at Health Protection Scotland holds information on more than 30,000 people from all over Scotland who have tested HCV antibody$^{+ve}$ between 1985 and 2011 (see (21) for a description of the database). The Scottish Care Information – Diabetes Collaboration (SCI-DC) manages a national register that holds information on individuals with diagnosed diabetes

(over 300,000) in Scotland and is estimated to have included over 99% of people with diagnosed diabetes since 2004. Individuals are included on SCI-DC if they have a Read code[1] for diabetes assigned in primary care or if they are seen in a hospital diabetes clinic (for a description of the database, see (22)).

## HCV antibody$^{+ve}$ cohort

For the period up to the end of 2011, 31,468 records of HCV antibody$^{+ve}$ people from all over Scotland were held in the HCV diagnoses database. From the database, information was extracted on forename initial, a soundex encrypted version of the surname (soundex is a phonetic algorithm for indexing names by sound, as pronounced in English), date of birth, sex, RNA test results at first diagnosis (positive, negative, unknown) and date of first HCV$^{+ve}$ antibody test (hereafter referred to as date of HCV diagnosis).

To enable linkage of the partially anonymised data in the HCV database to other databases, 24,975 (79%) records from the HCV database were probabilistically linked to the database of the community health index (CHI), a unique identifier used in medical records (23).

[1] Read codes are the standard clinical terminology system used in General Practice in the United Kingdom.

After linkage, information from CHI was added to the HCV antibody$^{+ve}$ cohort including full personal identifiers, postcode sector of residence at the time of HCV diagnosis, an indicator for social deprivation of the area of residence (Scottish Index of Multiple Deprivation, SIMD) (24) and an indicator and date for migration from Scotland. We then excluded 107 people younger than 16 and a total of 588 individuals with missing or unclear information on SIMD, sex and diagnosis date. After these exclusions, 24,280 individuals remained in the study population (see Figure 1 in the Appendix).

**HCV antibody$^{-ve}$ cohort**

For every person in the HCV antibody$^{+ve}$ cohort, up to three people were randomly sampled without replacement from the CHI database who were (i) born within one calendar year; (ii) of the same sex; (iii) alive at the time of diagnosis of the matched person on the HCV database; (iv) lived in the same postcode sector (but not in the same postcode) at the time of HCV diagnosis; and (v) were not included in the HCV antibody$^{+ve}$ cohort. Given the low prevalence of HCV in the Scottish population (25), less than 2% of the HCV antibody$^{-ve}$ cohort will likely have undiagnosed HCV-infection; thus, this misclassification will have negligible influence on the results. For 2118 people in the HCV

antibody$^{+ve}$ cohort, no matching individual could be identified in the CHI database; these people were excluded from the HCV antibody$^{+ve}$ cohort. As a result, 22,162 matched groups were available for analysis. People in the HCV antibody$^{-ve}$ cohort were assigned an index date which corresponded to the diagnosis date of their matched cohort member.

**Diabetes**

To identify diagnosed diabetes status in both cohorts (HCV antibody$^{+ve}$ and HCV antibody$^{-ve}$), data were deterministically linked to the SCI-DC database based on CHI number. After linkage, information from SCI-DC was added to the data including type of diabetes (T1DM, T2DM and other/unknown) and date of diabetes diagnosis. For 11 HCV antibody$^{+ve}$ people, diabetes was diagnosed but date of diabetes diagnosis was not available; these individuals, together with their 31 matched individuals from the HCV antibody$^{-ve}$ cohort, were removed from analysis. An additional three individuals from the HCV antibody$^{-ve}$ cohort with a diabetes diagnosis were removed as they had no date for their diagnosis. A further 219 HCV antibody$^{+ve}$ individuals were removed together with 652 matched individuals from the HCV

antibody$^{-ve}$ cohort because they had been diagnosed with a type of diabetes other than T2DM. Additionally, 451 people from the HCV antibody$^{-ve}$ cohort were excluded because they had been diagnosed with a type of diabetes other than T2DM.

**Morbidity and mortality**

To identify further censoring dates in both cohorts, data were then linked deterministically to mortality data from the General Registrars Office of Scotland (GRO, see (26) for a description of the database) and the date of death was added to the cohort data. Cohort members were additionally linked deterministically to hospital databases, to ascertain whether, prior to the HCV-diagnosis date, they had been in hospital for an alcohol-related admission (ICD9: 571.[0-3], 291.[0-9], 535.3, 425.5, 357.5, 305.0, 303.9; ICD10: E24.4, E51.2, F10.[0-9], G31.2, G62.1, G72.1, I42.6, K29.2, K70.[0-9], K86.0, O35.4, P04.3, Q86.0, R78.0, T51.[0,1,9], X[4,6]5, Y15, Y57.3, Y90.[3-8], Y91, Z50.2, Z71.4, Z72.1) or for an obesity-related admission (ICD9: 278.[0-9]; ICD10: E66). Three members of the HCV antibody$^{+ve}$ cohort matched to two different death records and were subsequently removed from analysis, leaving 21,929 for analysis (Fig. 1).

## Information Governance

Data linkages were approved by the NHS National Services Scotland Privacy Advisory Committee and use of the CHI database was approved by the CHI Advisory Group. All linkages were undertaken at Information Service Division, Scotland and all personal identifiable information removed from the outputs *prior* to release of data to the research team for analysis.

## Statistical analysis

The probability of T2DM diagnosis for those in the HCV antibody$^{+ve}$ compared to the HCV antibody$^{-ve}$ cohort was determined for the following three time periods: i) up-to one year before HCV diagnosis (*pre-HCV*); ii) from one year before HCV diagnosis to one year after HCV diagnosis (*peri-HCV*); and iii) from one year after HCV diagnosis to the earlier of either the end of follow-up (November 1st, 2011), death, migration out of Scotland or diagnosis of T2DM (*post-HCV*).

Generalized linear mixed models (R, package lme4) were used for the analysis of the odds of T2DM diagnosis *pre-HCV* and *peri-HCV*. Mixed effects Cox models (R, package coxme) were used for the analysis of the hazard of T2DM diagnosis *post-HCV*. In all three regression models, the year of HCV diagnosis (grouped into prior to

2000 and later than 1999), sex, social deprivation (grouped into three groups using the original quintiles: 1-2=high, 3=medium and 4-5=low deprivation) and age at HCV diagnosis were included as explanatory variables. Fractional polynomials were used to model age at HCV diagnosis (R, package mfp). To adjust for correlation within matched groups, a random group effect was added to all three models.

To study if the estimated effect of HCV-infection on the probability of T2DM diagnosis was modified by period of HCV diagnosis, sex, social deprivation or age at HCV diagnosis, interaction-terms between these variables and HCV were added to the full model. Likelihood ratio tests were used for testing the statistical significance of interaction terms and those interaction terms that were not statistically significant ($P>0.05$) were removed. For statistically significant interaction terms, a synergy index ($S$) was calculated to demonstrate the excess risk from exposure (to both exposures) when there is interaction relative to the risk from exposure (to both exposures) without interaction. Influential values, outliers and model fit were ascertained in the final models excluding random group effects (R, package boot). The assumption of proportionality of hazards in the survival analysis was tested using Schoenfeld residuals (R, package survival).

To study the effect of chronic and resolved HCV infection, all final models were re-run separately for those in the HCV antibody$^{+ve}$ cohort who were initially tested (i) RNA-positive (indicative of chronic HCV) and (ii) RNA-negative (indicative of resolved HCV). Here, the HCV antibody$^{-ve}$ cohorts were composed only of people who were matched to RNA-positive (for ( i)) and RNA-negative (for (ii)) individuals.

# Results

### Characteristics of the study population

Table 1 shows the composition of the study population comprising 21,929 people in the HCV antibody$^{+ve}$ cohort and 65,074 people in the matched HCV antibody$^{-ve}$ cohort. Reflecting the composition of the HCV antibody $^{+ve}$ population in Scotland, people in the HCV antibody$^{+ve}$ cohort were predominantly male (68%), born between 1960 and 1980 (68%), were diagnosed with HCV after the year 2000 (70%) and were living at the time of HCV diagnosis in areas of highest deprivation (75%). 72% of the people in the HCV antibody$^{+ve}$ cohort were HCV-RNA$^{+ve}$, 18% were HCV-RNA$^{-ve}$ and in 10% the RNA status was unknown. More than 97% of people in the HCV antibody$^{+ve}$

cohort could be matched to three HCV antibody$^{-ve}$ people from the CHI database, while for people born before 1950 fewer matches were identified.

Median follow-up time from HCV-diagnosis to censoring or end of follow-up was 6.4 years in the HCV antibody$^{+ve}$ cohort and 6.6 years in the HCV antibody$^{-ve}$ cohort; median age at HCV diagnosis was 33 years. During a total follow-up period of 151,020 person-years from HCV-diagnosis to censoring in the HCV antibody$^{+ve}$ cohort, 4016 people died (2.66 per 100 person-years). In the HCV antibody$^{-ve}$ cohort, the total follow-up period was 463,977 person-years with 2633 deaths recorded (0.57 per 100 person-years). The proportion of people who have had an alcohol-related hospitalization prior to HCV-diagnosis was considerably higher in the HCV antibody$^{+ve}$ cohort (22%) than in the HCV antibody$^{-ve}$ cohort (4.5%), while there was not much difference in the proportion of people who have had an obesity-related hospitalization (both 0.3%) prior to HCV-diagnosis.

**Diagnosis of T2DM in the HCV antibody$^{+ve}$ cohort compared to the HCV antibody$^{-ve}$ cohort**

Of 21,929 people in the HCV antibody$^{+ve}$ cohort, 628 (2.86%) had been diagnosed with T2DM, of whom 187 (30%) had been diagnosed with T2DM more than a year before they had been diagnosed HCV-positive and 141 (22%) had been diagnosed with T2DM within one calender year of their HCV diagnosis (Table 2). This compares to 1772 out of 65,074 (2.72%) in the HCV antibody$^{-ve}$ cohort who have been diagnosed with T2DM, of whom 524 (30%) had been diagnosed with T2DM more than a year before the matched person in the HCV antibody$^{+ve}$ cohort had been diagnosed HCV-positive and 184 (10%) had been diagnosed with T2DM within one calender year of their HCV diagnosis (Table 2). The difference between both cohorts in the proportion of people who were diagnosed with T2DM (0.14%) indicates an excess of 32 cases in HCV antibody$^{+ve}$ study population or 14 per 10,000 HCV-infected people, while for those who tested RNA$^{+ve}$ and RNA$^{-ve}$, excess risks of 34 and 20 per 10,000, respectively, were found. In both HCV antibody$^{+ve}$ and HCV antibody$^{-ve}$ cohorts the median age at diagnosis with T2DM was 45 years.

**Odds of T2DM diagnosis up to one year *prior* to HCV diagnosis**

In the HCV antibody$^{-ve}$ cohort, male sex and high social deprivation were associated with increased risks of having a diagnosis of T2DM in the period up to one year *prior* to HCV diagnosis. However, in the HCV antibody$^{+ve}$ cohort, the same variables were associated with decreased risk (Table 3). The 4345 women in the HCV antibody$^{-ve}$ cohort who resided in areas of lowest deprivation had the lowest risk of having a diagnosis of T2DM (0.4%), while the 941 women in the HCV antibody$^{+ve}$ cohort who resided in areas of lowest deprivation had the highest risk (2.4%; OR, 4.02; 95% CI, 2.29-7.04 $P$<0.01). The 28,267 men in the HCV antibody$^{-ve}$ cohort who resided in areas of highest deprivation had a higher risk of having a diagnosis of T2DM (0.9%) than the 11,131 men in the HCV antibody$^{+ve}$ cohort who resided in areas with the same high deprivation (0.5%; OR, 0.61; 95% CI, 0.43-0.87 $P$<0.01). The synergy indices show negative interaction on an additive scale, indicating that the combined effects of male sex and HCV-infection and deprivation and HCV-infection were less than the sum of the effects of male sex and HCV-infection and deprivation and HCV-infection.

Similar ORs were estimated when restricting the HCV-positive cohort to either only people who have tested RNA$^{+ve}$ (indicative of

chronic infection) or those who have tested RNA-negative (indicative of past infection; Table 3).

**Odds of T2DM diagnosis within ±one year of HCV diagnosis**

In the HCV antibody$^{-ve}$ cohort, male sex was associated with increased risks of having a diagnosis of T2DM in the period within one year of HCV diagnosis. However, in the HCV antibody$^{+ve}$ cohort, there was little difference between men and women (Table 4). The lowest risk of having a diagnosis of T2DM was observed for the 20,626 women in the HCV antibody$^{-ve}$ cohort (0.2%) while the highest risk was observed for the 6996 women in the HCV antibody$^{+ve}$ cohort (0.7%; OR, 3.78; 95% CI, 2.29-6.24 $P$<0.01). Increased risks of having a diagnosis of T2DM were also observed in the 14,746 men in the HCV antibody$^{+ve}$ cohort (0.6%) compared to men in the HCV antibody$^{-ve}$ cohort (0.3%), but because of the increased risk in males in the HCV antibody$^{-ve}$ cohort, the estimated adjusted OR was lower than in women (OR, 1.97; 95% CI, 1.46-2.65; $P$<0.01). Again, the synergy index indicates negative interaction on an additive scale between the effect of male sex and HCV-infection ($S$=0.71).

The estimated increased odds for women in the HCV antibody$^{+ve}$ cohort compared to those in the HCV antibody$^{-ve}$ cohort further increased when only women were included in the data set who had tested RNA-positive (OR, 4.57). Increased odds were also calculated for those women who tested RNA-negative (OR, 2.89). For men, estimates for the effect of HCV-infection on the odds of having a diagnosis of T2DM were similar in the full data set (OR, 1.97), the RNA-positives (OR, 2.07) or RNA-negatives (OR, 2.02). However, restricting the cohort to RNA-negatives, the variance for estimates increased and some of the differences in the odds between people in the HCV-positive cohort and the HCV antibody$^{-ve}$ cohort were not statistically significant (Table 4).

**Hazard of T2DM diagnosis later than one year after HCV diagnosis**

In the HCV antibody$^{-ve}$ cohort, increasing social deprivation was associated with an increased hazard of having a diagnosis of T2DM in the period later than one year after HCV diagnosis. However, in the HCV antibody$^{+ve}$ cohort, increasing social deprivation was associated with a decreased hazard of having a diagnosis of T2DM (Table 5). The lowest hazard of having a diagnosis of T2DM was observed for the 14,298 people in the HCV antibody$^{+ve}$ cohort who lived in areas of

highest deprivation (1.4%) which was (non-significantly) lower than the hazard for the 34,470 members of the HCV antibody$^{-ve}$ cohort living in the same areas of high deprivation (1.9%; HR, 0.88; 95% CI, 0.75-1.03 $P$=0.11). The highest hazard was observed for the 2401 people in the HCV antibody$^{+ve}$ cohort who lived in areas of lowest deprivation (2.5%) which was (significantly) higher than the hazard for the 10,957 members of the HCV antibody$^{-ve}$ cohort living in the same areas of low deprivation (1.6%; HR, 1.53; 95% CI, 1.14-2.04 $P$<0.01). The synergy indices indicate negative interaction on an additive scale between the effect of deprivation and HCV-infection.

Slightly higher effects of HCV-infection on the hazard of being diagnosed with T2DM more than one year after HCV diagnosis were estimated when restricting the HCV-positive cohort to those who have tested RNA$^{+ve}$ (indicative of chronic infection). Increased hazards were also estimated for those HCV antibody$^{+ve}$ who tested RNA-negative and who lived in areas with high or low deprivation; however, due to the small sample size, those differences were not statistically significant (Table 5).

# Discussion

This study compares the risk of receiving a diagnosis of T2DM in a cohort of all people who have been diagnosed HCV antibody$^{+ve}$ in Scotland (the vast majority of whom will have acquired infection through injecting drug use) with that of a three times larger HCV antibody$^{-ve}$ cohort matched on year of birth, sex and neighbourhood. The HCV antibody$^{+ve}$ cohort was further stratified by RNA-status to check whether any additional risk attributed to HCV infection was related to the virus infection itself or to other factors related to the infection. Further the effect of HCV infection in three time periods - *pre*-HCV, *peri*-HCV and *post*-HCV diagnosis was studied to investigate any bias due to increased testing for T2DM at the time of HCV diagnosis.

This study shows that nationwide over a time-period of approximately 12 years there were approximately 14 additional cases of T2DM for every 10,000 HCV-infected people compared to what would have been observed in a HCV antibody$^{-ve}$ cohort of identical size and characteristic. The excess risk was similarly low among RNA$^{+ve}$ when taking into account the excess risk among RNA$^{-ve}$. Including those with HCV who are undiagnosed (nationwide approximately 50%, (25) ), we

would expect that the total excess number of people with HCV-antibody infection who have developed HCV-related T2DM up to this point in time is less than 100.

While this is the first study to estimate the extra number of HCV-related T2DM cases for a whole nation, increases in risk of those with HCV have been reported elsewhere (1-4). For the national health system of Scotland, compared to total number of people reported to have been diagnosed with T2DM (265,000 between 2000 and 2012), the increase of less than 100 cases in a 12-year period is relatively small. Similarly, for the HCV-infected individual, compared to lifestyle choices related to an increase in T2DM risk, the increase in risk related to HCV-infection from 2.7% to 2.9% seems comparably low. The relatively small difference in risk observed in our study indicates the necessity to study the association between HCV-infection and T2DM in large, well-defined study populations. Ruhl et al. (9) found no association between HCV and either diabetes or insulin resistance (IR) in their US population based study, involving 277 HCV antibody$^{+ve}$ individuals (compared to the 21,929 studied here); a relationship between HCV and diabetes could only be found among those with elevated enzyme activity. Ruhl et al. thus suggest that the previously reported findings of a strong relationship with diabetes may have

417 resulted from the increased liver enzyme activity in the HCV
418 populations studied (9). Further, a recent meta-analysis has found an
419 association between the presence of IR and advanced fibrosis in those
420 with HCV genotype 1 (the most common genotype in the US), but not
421 for genotype 3 (27). We lacked data on liver enzyme activity, IR and
422 HCV genotype in this database linkage study to be able to investigate
423 this further in a larger cohort.

424 Matching allowed us to control for the effects of age, sex and
425 neighbourhood; the latter being a proxy for social deprivation and
426 regional differences in testing and recording for both conditions.
427 However, estimates of the number of additional cases of T2DM in those
428 with HCV-infection could have been biased from other risk factors for
429 T2DM for which information was not available. Ethnicity is known to be
430 related to T2DM, with people of South Asian background living in the
431 UK having 3-4 times higher risk of developing T2D during their life
432 compared to the majority white population (20). Moreover, people of
433 South Asian ethnicity are known to have a higher prevalence of HCV
434 (28), so a higher proportion of people with South Asian ethnicity would
435 be expected in the HCV-positive cohort. However, the South Asian
436 population in Scotland is very small (≈1% in the 2001 census), so that
437 confounding from a varying ethnic composition of the HCV-positive

cohort and the HCV-negative cohort can be expected to be small. Body-mass is a further known risk factor for T2DM, and it is possible that differences in BMI may confound the association between diagnoses of HCV and T2DM. However, since social deprivation and obesity are closely correlated in Scotland (14), matching by neighbourhood should have increased comparability of both cohorts, as indicated by similar proportions of people with a record of an obesity related hospitalization in the HCV antibody$^{+ve}$ and the HCV antibody$^{-ve}$ cohort. Similarly, alcohol consumption is a known risk factor for T2DM (29) and because alcohol consumption is positively related to HCV-status it could be expected that the proportion of people with high alcohol consumption was higher in the HCV-positive cohort compared to the HCV-negative cohort. Indeed, compared to people in the HCV-negative cohort, people in the HCV-positive cohort had a 4.6-times higher risk of having an alcohol-related hospitalization. This bias from other risk factors related to T2DM might explain the observation in our study that compared to people in the HCV antibody $^{-ve}$ cohort, people with resolved HCV-infection (RNA-negative) were still at higher risk of having a diagnosis of T2DM.

The study also shows that the effect of diagnosed HCV-infection on the relative proportions of people with a diagnosis of T2DM was time

dependent. Partitioning of the risk period clearly showed that the increased risk is mainly due to increased T2DM diagnosis around the time of HCV diagnosis, while the 10% increased risk more than one year *prior* to HCV diagnosis and one year *post* HCV diagnosis were considerably lower than the estimate from the meta-analyses. Interestingly, the estimate of a 10% increased relative risk is very similar to that from the largest cohort study that had been included in the meta-analyses (30) although the estimate of absolute T2DM prevalence in the HCV antibody$^{-ve}$ cohort in our study (3.2%) was much lower than that in the US study (13%) or indeed any other cohort study but one included in the meta-analyses. Increased T2DM within ±1 year is likely related to ascertainment bias. However, neither guidelines by the Scottish Intercollegiate Guideline Network (SIGN guidelines 116) nor by the National Institute of Clinical Excellence recommend testing for HCV infection in people diagnosed with T2DM and guidelines by the European Association for the Study of the Liver only recommend testing for T2DM *prior* to treatment for HCV infection, since 'poorly controlled diabetes' is a contra-indication for treatment with interferon containing regimens. Therefore, the most likely reason for the increased T2DM diagnosis *peri*-HCV diagnosis is related to people showing clinical symptoms indicative of liver disease. It seems likely that for people with

480 signs of liver disease, a blood sample for glucose testing is collected at
481 the same time as samples for HCV tests and liver function
482 measurements. We do not have access to laboratory test databases in
483 order to investigate the potential for ascertainment bias further. While
484 there was a highly significant correlation between increasing age and the
485 risk of T2DM diagnosis, there was no significant increase with age in the
486 effect of HCV infection on the risk of T2DM ($P$=0.34 for inclusion of an
487 HCV*age interaction term). This result indicates that the observed effect
488 of HCV infection on the risk of T2DM is more likely caused by other
489 factors related to HCV infection than by the (slowly progressing) action
490 of the virus. However, our HCV infected cohort is still relatively young
491 (median age at HCV diagnosis was 33 years) and has been followed up
492 for a relatively short time (median of 6.4 years), thus the excess risk of
493 T2DM may still change as our cohort advances in age and duration of
494 infection.

495 Male sex and living in areas of highest deprivation decreased effects
496 of HCV infection on the risk of T2DM diagnosis. This effect
497 modification was not related to follow-up time, age at HCV-infection or
498 RNA-status since those did not differ within sex and social deprivation.
499 Since male sex and high deprivation are positively related to T2DM risk,
500 our observation does not confirm the suggestion from (11) that relative

effects of HCV on T2DM risk are higher in people at increased risk of T2DM. However, the effect modification could be explained by different uptake of health care (and thereby testing for diabetes) in men living in areas of high deprivation. The effect modification could explain some of the heterogeneity that both meta analyses found, since few of the reviewed studies stratified by sex and none by social deprivation. However, widely accepted biological models of the effects of HCV infection on T2DM risk (5-7) do not explain the observed effect modification. Moreover, while sex, social deprivation and year of birth were included in our matched analysis to increase efficiency of the study (31), the analysis of effect modification by sex, social deprivation, year and age was purely exploratory.

Ideally, every person in the HCV-positive cohort should have been followed-up from the date of HCV-infection to development of T2DM or censoring. However, because date of HCV-infection was unknown, the follow-up period and thereby the risk of T2DM diagnosis *pre*-HCV diagnosis was heterogeneous. Additionally, the T2DM database is only approximately complete from 2004 onwards, with regional differences in the date from which diagnoses of T2DM were reported to the database. By matching people in the HCV antibody$^{-ve}$ cohort to those in the HCV antibody$^{+ve}$ cohort by year of birth and place of residence and

by adequately controlling for the effect of matching in the analysis we managed to reduce the potential bias for the odds ratio from heterogeneous follow-up times. However, the estimated odds of T2DM diagnosis *pre*-HCV diagnosis in both cohorts are difficult to interpret. In addition, since date of HCV-infection and date of onset of T2DM both were unknown to us, the temporal relationship of onset of HCV infection and T2DM is not known. Indeed, T2DM has been described as a risk factor for contracting HCV (12). However, an estimated 86% of HCV-infection in Scotland is related to injecting drug use (32) and a large fraction of those diagnosed HCV-positive will have been infected in their early drug using career. Given that the risk of developing T2DM increases with age, it is unlikely that the increased risk for HCV in those with T2DM was responsible for the results of our study.

Our study has demonstrated that on the population level the size of the effect of HCV antibody status on T2DM is smaller than effects of many life style choices (e.g., obesity, smoking and alcohol consumption) and therefore not as significant a public health concern as previously suggested from predominantly clinic based studies. Findings were similarly observed for both $RNA^{+ves}$ (chronic) and $RNA^{-ves}$ (resolved) which further indicates that the observed differences in risk of T2DM diagnosis were not related to the virus itself but to factors related to the

infection (e.g., factors related to drug abuse). However, given the increased risk for HCV-related disease progression in those affected by both conditions (33), further research is required to identify whether screening and earlier treatment for T2DM improves outcomes among people with a diagnosis of chronic HCV. Socio-economic status, sex and a history of alcohol use and injecting drug use modify the effect of HCV on T2DM which could explain some of the discrepancies between different studies given the different patterns of these factors in different populations.

## Acknowledgement

We are grateful to the following virologists for their support with the HCV diagnosis database: Dr Kate Templeton (East of Scotland Specialist Virology Centre, Royal Infirmary of Edinburgh, Edinburgh), Dr Celia Aitkin (West of Scotland Specialist Virology Centre, Gartnavel General Hospital, Glasgow), Dr Paul McIntyre (Department of Medical Microbiology, Ninewells Hospital and Medical School, Dundee), and Dr Pamela Molyneaux (Department of Medical Microbiology, University Medical School, Foresterhill, Aberdeen).

## Literature

(1) Arao M, Murase K, Kusakabe A, Yoshioka K, Fukuzawa Y, Ishikawa T, et al. Prevalence of diabetes mellitus in Japanese patients infected chronically with hepatitis C virus. J Gastroenterol 2003;38(4):355-360.

(2) Serfaty L, Capeau J. Hepatitis C, insulin resistance and diabetes: clinical and pathogenic data. Liver Int 2009 Mar;29 Suppl 2:13-25.

(3) Naing C, Mak JW, Ahmed SI, Maung M. Relationship between hepatitis C virus infection and type 2 diabetes mellitus: meta-analysis. World J Gastroenterol 2012 Apr 14;18(14):1642-1651.

(4) Dai CY, Yeh ML, Huang CF, Hou CH, Hsieh MY, Huang JF, et al. Chronic hepatitis C infection is associated with insulin resistance and lipid profiles. J Gastroenterol Hepatol 2013 Jun 28. Epub ahead of print.

(5) Bose SK, Ray R. Hepatitis C virus infection and insulin resistance. World J Diabetes 2014 Feb 15;5(1):52-58.

(6) Alexander GJ. An association between hepatitis C virus infection and type 2 diabetes mellitus: what is the connection? Ann Intern Med 2000 Oct 17;133(8):650-652.

(7) Naing C, Mak JW, Wai N, Maung M. Diabetes and infections-hepatitis C: is there type 2 diabetes excess in hepatitis C infection? Curr Diab Rep 2013 Jun;13(3):428-434.

(8) White DL, Ratziu V, El-Serag HB. Hepatitis C infection and risk of diabetes: a systematic review and meta-analysis. J Hepatol 2008 Nov;49(5):831-844.

(9) Ruhl CE, Menke A, Cowie CC, Everhart JE. Relationship of Hepatitis C Virus Infection with Diabetes in the U.S. Population. Hepatology 2014 Oct; 60 (4):1139-1149.

(10) Montenegro L, De Michina A, Misciagna G, Guerra V, Di Leo A. Virus C hepatitis and type 2 diabetes: a cohort study in southern Italy. Am J Gastroenterol 2013 Jul;108(7):1108-1111.

(11) Mehta SH, Brancati FL, Strathdee SA, Pankow JS, Netski D, Coresh J, et al. Hepatitis C virus infection and incident type 2 diabetes. Hepatology 2003 Jul;38(1):50-56.

(12) Guo X, Jin M, Yang M, Li J. Type 2 diabetes mellitus and the risk of hepatitis c virus infection: a systematic review. Scientific Reports 2013;3(2981):1-8.

(13) Espelt A, Borrell C, Roskam AJ, Rodriguez-Sanz M, Stirbu I, Dalmau-Bueno A, et al. Socioeconomic inequalities in diabetes mellitus

across Europe at the beginning of the 21st century. Diabetologia 2008 Nov;51(11):1971-1979.

(14) Evans JM, Newton RW, Ruta DA, MacDonald TM, Morris AD. Socio-economic status, obesity and prevalence of Type 1 and Type 2 diabetes mellitus. Diabet Med 2000 Jun;17(6):478-480.

(15) Pereska Z, Bozinovska C, Dimitrovski C, Cakalarovski K, Chibishev A, Zdravkovska M, et al. Heroin dependence duration influences the metabolic parameters: mechanisms and consequences of impaired insulin sensitivity in hepatitis C virus seronegative heroin dependents. J Addict Med 2012 Dec;6(4):304-310.

(16) Fareed A. Predictors of diabetes mellitus and abnormal blood glucose in patients receiving opioid maintenance treatment. American journal on addictions 2013;22(4):411.

(17) Beulens JW. Estimating the mediating effect of different biomarkers on the relation of alcohol consumption with the risk of type 2 diabetes. Ann Epidemiol 2013;23(4):193-197.

(18) Kim SJ, Jee SH, Nam JM, Cho WH, Kim JH, Park EC. Do early onset and pack-years of smoking increase risk of type II diabetes? BMC Public Health 2014 Feb 19;14(1):178.

(19) Wild S, Roglic G, Green A, Sicree R, King H. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. Diabetes Care 2004 May;27(5):1047-1053.

(20) Fischbacher CM, Bhopal R, Steiner M, Morris AD, Chalmers J. Is there equity of service delivery and intermediate outcomes in South Asians with type 2 diabetes? Analysis of DARTS database and summary of UK publications. J Public Health (Oxf) 2009 Jun;31(2):239-249.

(21) Shaw L, Taylor A, Roy K, Cameron S, Burns S, Molyneaux P, et al. Establishment of a database of diagnosed HCV-infected persons in Scotland. Communicable Disease and Public Health / PHLS 2003;6(4):305-310.

(22) Scottish Diabetes Group. Diabetes in Scotland. Available at: http://www.diabetesinscotland.org.uk, 2014.

(23) NHS National Services Scotland. What is the Community Health Index (CHI)? Available at: http://www.shsc.scot.nhs.uk/upload/file/national_committee_services/chiag/2010_10_19_what_is_the_community_health_index.pdf, 2014.

(24) Donnan PT, Leese GP, Morris AD, Diabetes Audit and Research in Tayside, Scotland/Medicine Monitoring Unit Collaboration. Hospitalizations for people with type 1 and type 2 diabetes compared

639 with the nondiabetic population of Tayside, Scotland: a retrospective
640 cohort study of resource use. Diabetes Care 2000
641 Dec;23(12):1774-1779.
642 (25) Harris H, editors. Hepatitis C in the UK: Annual Report 2014.
643 Available at:
644 https://www.gov.uk/government/uploads/system/uploads/attachment_d
645 ata/file/337115/HCV_in_the_UK_2014_24_July.pdf; 2014.
646 (26) General Register Office for Scotland. Vital events: Technical
647 Report. Available at:
648 http://www.gro-scotland.gov.uk//statistics//theme//vital-events//deaths//
649 index.html. (Accessed: April 2014).
650 (27) Patel S, Jinjuvadia R, Patel R, Liangpunsakul S. Insulin resistance
651 is associated with significant liver fibrosis in chronic hepatitis C
652 patients: a systemic review and meta-analysis. J Clin Gastroenterol.
653 2016 Jan; 50(1): 80-4.
654 (28) Uddin G, Shoeb D, Solaiman S, Marley R, Gore C, Ramsay M, et
655 al. Prevalence of chronic viral hepatitis in people of south Asian
656 ethnicity living in England: the prevalence cannot necessarily be
657 predicted from the prevalence in the country of origin. J Viral Hepat
658 2010;17(5):327-35.

(29) Wannamethee SG, Shaper AG, Perry IJ, Alberti KG. Alcohol consumption and the incidence of type II diabetes. J Epidemiol Community Health 2002 Jul;56(7):542-548.

(30) Butt AA, Evans R, Skanderson M, Shakil AO. Comorbid medical and psychiatric conditions and substance abuse in HCV infected persons on dialysis. J Hepatol 2006 May;44(5):864-868.

(31) Greenland S, Morgenstern H. Matching and efficiency in cohort studies. Am J Epidemiol 1990 Jan;131(1):151-159.

(32) McDonald SA, Hutchinson SJ, Schnier C, McLeod A, Goldberg DJ. Estimating the number of injecting drug users in Scotland's HCV-diagnosed population using capture-recapture methods. Epidemiol Infect 2014 Jan;142(1):200-207.

(33) Arase Y, Kobayashi M, Suzuki F, Suzuki Y, Kawamura Y, Akuta N, et al. Effect of type 2 diabetes on risk for malignancies includes hepatocellular carcinoma in chronic hepatitis C. Hepatology 2013 Mar;57(3):964-973.

676

677 Table 1: Characteristics of the study population

678

| Variable | Level | HCV Ab$^{+ve}$ cohort (N) | HCV Ab$^{-ve}$ cohort (N) | % | % Complete Matches[1] |
|---|---|---|---|---|---|
| Sex | Women | 7067 | 20,956 | 32 | 97 |
| | Men | 14,862 | 44,118 | 68 | 97 |
| Year of birth | <1950 | 1335 | 3859 | 6 | 90 |
| | 1950-1959 | 2876 | 8521 | 13 | 96 |
| | 1960-1969 | 7246 | 21,545 | 33 | 97 |
| | 1970-1979 | 7616 | 22,656 | 35 | 98 |
| | ≥1980 | 2856 | 8493 | 13 | 97 |
| Year of diagnosis | <2000 | 6592 | 19526 | 30 | 96 |
| | ≥2000 | 15,337 | 45,548 | 70 | 97 |
| Deprivation | Low | 2824 | 13,604 | 13/21[3] | 96 |
| | Medium | 2678 | 9628 | 12/15[3] | 96 |
| | High | 16,427 | 41,842 | 75/64[3] | 97 |
| Alcohol-related hospitalization[2] | Yes | 4812 | 2942 | 22/4.5[3] | |
| Obesity-related hospitalization[2] | Yes | 60 | 209 | 0.3/0.3[3] | |
| *Total* | | 21,929 | 65,074 | | 97 |

[1] A complete match is 1 person in the HCV antibody$^{+ve}$ cohort and 3 people in the HCV antibody$^{-ve}$ cohort matched on year of birth, sex and postcode sector of residence.

[2] Alcohol and obesity related hospitalization prior to HCV diagnosis; ICD9 codes and ICD10 codes as listed in patients and methods.

[3] HCV antibody$^{+ve}$ and HCV antibody$^{-ve}$, respectively.

679

680

Table 2: Number (and proportion) of people with T2DM in the HCV antibody$^{+ve}$ cohort (including for those PCR$^{+ve}$ and PCR$^{-ve}$) and in the HCV antibody$^{-ve}$ cohort according to time since HCV diagnosis.

| Period since HCV diagnosis[1] | HCV Ab$^{-ve}$ (N=65,074) | | HCV Ab$^{+ve}$ (N=21,929) | | HCV Ab$^{+ve}$ & PCR$^{+ve}$ (N=15,827) | | HCV Ab$^{+ve}$ & PCR$^{-ve}$ (N=3,927) | |
|---|---|---|---|---|---|---|---|---|
| | Diabetes$^{+ve}$ | % | Diabetes$^{+ve}$ | % | Diabetes$^{+ve}$ | % | Diabetes$^{+ve}$ | % |
| >1 year pre | 524 | 0.81 | 187 | 0.85 | 157 | 0.99 | 23 | 0.59 |
| ± 1 year | 184 | 0.28 | 141 | 0.64 | 115 | 0.73 | 18 | 0.46 |
| >1 year post | 1064 | 1.64 | 300 | 1.37 | 234 | 1.48 | 49 | 1.25 |
| *Total* | 1772 | 2.72 | 628 | 2.86 | 506 | 3.20 | 90 | 2.29 |

[1] For those in the HCV antibody$^{-ve}$ cohort, HCV diagnosis data was taken to be the same as their respective HCV antibody$^{+ve}$ cohort members, for the purpose of analysis.

690

691 Table 3: Odds of having been diagnosed with T2DM in the period up to 1 year before
692 HCV diagnosis in the HCV antibody$^{+ve}$ cohort (total and broken down by PCR status)
693 compared to the HCV antibody$^{-ve}$ cohort[1,2]

694 .

| Sex | Deprivation | Diabetes$^{+ve}$ /HCV Ab$^{-ve}$ | Diabetes$^{+ve}$ /HCV Ab$^{+ve}$ | aOR[3] (95% CI; *P*) | S[4] |
|---|---|---|---|---|---|
| Antibody$^{+ve}$ | | | | | |
| F | Low | 17/4345 (0.4%) | 23/941 (2.4%) | 4.02 (2.32-6.96); *P*<0.01 | |
| F | Medium | 16/3036 (0.5%) | 10/830 (1.2%) | 1.92 (0.95-3.86); *P*=0.08 | 0.42 |
| F | High | 101/13,575 (0.7%) | 38/5296 (0.7%) | 1.05 (0.66-1.69); *P*=1.00 | 0.32 |
| M | Low | 77/9259 (0.8%) | 40/1883 (2.1%) | 2.33 (1.42-3.83); *P*<0.01 | 0.62 |
| M | Medium | 57/6592 (0.9%) | 19/1848 (1.0%) | 1.11 (0.58-2.11); *P*=0.99 | 0.28 |
| M | High | 256/28,267 (0.9%) | 57/11,131 (0.5%) | 0.61 (0.43-0.87); *P*<0.01 | 0.15 |
| Antibody$^{+ve}$ and PCR$^{+ve}$ | | | | | |
| F | Low | 12/3067 (0.4%) | 18/661 (2.7%) | 4.35 (2.33-8.13); *P*<0.01 | |
| F | Medium | 10/2117 (0.5%) | 7/575 (1.2%) | 2.05 (0.93-4.50); *P*=0.09 | 0.42 |
| F | High | 80/9098 (0.9%) | 33/3576 (0.9%) | 1.14 (0.67-1.93); *P*=0.96 | 0.35 |
| M | Low | 59/6886 (0.9%) | 34/1375 (2.5%) | 2.61 (1.50-4.55); *P*<0.01 | 0.63 |
| M | Medium | 44/4877 (0.9%) | 16/1360 (1.2%) | 1.23 (0.60-2.54); *P*=0.93 | 0.30 |
| M | High | 202/20,936 (1.0%) | 49/8280 (0.6%) | 0.68 (0.46-1.01); *P*=0.06 | 0.19 |
| Antibody$^{+ve}$ and PCR$^{-ve}$ | | | | | |
| F | Low | 0/841 (0.0%) | 4/169 (2.4%) | 6.14 (1.38-27.21); *P*<0.01 | |
| F | Medium | 4/669 (0.6%) | 3/175 (1.7%) | 2.69 (0.55-13.23); *P*=0.42 | 0.63 |
| F | High | 18/3343 (0.5%) | 4/1294 (0.3%) | 0.74 (0.21-2.61); *P*=0.96 | 0.09 |
| M | Low | 11/1339 (0.8%) | 5/267 (1.9%) | 2.45 (0.63-9.55); *P*=0.36 | 0.54 |
| M | Medium | 9/1010 (0.9%) | 3/283 (1.1%) | 1.07 (0.25-4.66); *P*=1.00 | 0.32 |
| M | High | 33/4450 (0.7%) | 4/1739 (0.2%) | 0.29 (0.09-0.98); *P*=0.04 | 0.01 |

695 [1]For those in the HCV antibody$^{-ve}$cohort, HCV diagnosis date was taken to be the
696 same as their respective HCV antibody$^{+ve}$cohort members, for the purpose of
697 analysis.
698 [2]Based on the likelihood-ratio test comparing the antibody$^{+ve}$cohort to the
699 antibody$^{-ve}$cohort, interaction-terms other than sex × HCV and deprivation ×
700 HCV were deemed not statistically significant and therefore excluded from the
701 final model.
702 [3]Adjusted OR and *P* for exposure to HCV-infection within strata of sex and social
703 deprivation. Odds ratios adjusted for age at HCV diagnosis, year of HCV
704 diagnosis and the extra correlation due to the matching.
705 [4]Synergy Index.

706

Table 4: Odds of having a diagnosis of T2DM in the period within ±1 year of the time of HCV diagnosis in the HCV antibody$^{+ve}$ cohort (total and broken down by PCR status) compared to the HCV antibody$^{-ve}$ cohort[1,2]

| Sex | Diabetes$^{+ve}$ /HCV Ab$^{-ve}$ | Diabetes$^{+ve}$ /HCV Ab$^{+ve}$ | aOR[3] (95% CI; *P*) | S[4] |
|---|---|---|---|---|
| Antibody$^{+ve}$ | | | | |
| F | 36/20,626 (0.2%) | 46/6996 (0.7%) | 3.78 (2.29-6.25); *P*<0.01 | |
| M | 142/43,406 (0.3%) | 95/14,746 (0.6%) | 1.97 (1.46-2.65); *P*<0.01 | 0.71 |
| Antibody$^{+ve}$ and PCR$^{+ve}$ | | | | |
| F | 25/13,230 (0.2%) | 38/4486 (0.8%) | 4.57 (2.56-8.18); *P*<0.01 | |
| M | 111/30,223 (0.4%) | 77/10,273 (0.7%) | 2.07 (1.48-2.90); *P*<0.01 | 0.66 |
| Antibody$^{+ve}$ and PCR$^{-ve}$ | | | | |
| F | 6/4591 (0.1%) | 6/1555 (0.4%) | 2.89 (0.52-16.01); *P*=0.31 | |
| M | 18/6283 (0.3%) | 12/2131 (0.6%) | 2.02 (0.67-6.10); *P*=0.29 | 1.01 |

[1]For those in the HCV antibody$^{-ve}$cohort, HCV diagnosis date was taken to be the same as their respective HCV antibody$^{+ve}$cohort members, for the purpose of analysis.

[2]Based on the likelihood-ratio test comparing the antibody$^{+ve}$cohort to the antibody$^{-ve}$cohort, interaction-terms other than sex × HCV were deemed not statistically significant and therefore excluded from the final model.

[3]Adjusted OR and *P* for exposure to HCV-infection within strata of sex. Odds ratios adjusted for age at HCV diagnosis, year of HCV diagnosis, social deprivation and the extra correlation due to the matching.

[4]Synergy Index.

713

714 Table 5: Hazard of being diagnosed with T2DM in the period >1 year after the time of
715 HCV diagnosis in the HCV antibody$^{+ve}$ cohort (total and broken down by PCR status)
716 compared to the HCV antibody$^{-ve}$ cohort[1,2]

717

| Deprivation | Diabetes$^{+ve}$ /HCV Ab$^{-ve}$ | Diabetes$^{+ve}$ /HCV Ab$^{+ve}$ | aHR[3] (95% CI; *P*) | S[4] |
|---|---|---|---|---|
| Antibody$^{+ve}$ | | | | |
| Low | 175/10,957 (1.6%) | 61/2401 (2.5%) | 1.53 (1.14-2.04); *P*<0.01 | |
| Medium | 137/7740 (1.8%) | 43/2308 (1.9%) | 1.14 (0.81-1.60); *P*=0.47 | 0.74 |
| High | 646/34,470 (1.9%) | 196/14,298 (1.4%) | 0.88 (0.75-1.03); *P*=0.11 | 0.36 |
| Antibody$^{+ve}$ and PCR$^{+ve}$ | | | | |
| Low | 118/8158 (1.4%) | 47/1750 (2.7%) | 1.71 (1.21-2.40); *P*<0.01 | |
| Medium | 100/5659 (1.8%) | 35/1677 (2.1%) | 1.26 (0.86-1.86); *P*=0.24 | 0.70 |
| High | 470/25,027 (1.9%) | 152/10,448 (1.5%) | 0.89 (0.74-1.07); *P*=0.22 | 0.39 |
| Antibody$^{+ve}$ and PCR$^{-ve}$ | | | | |
| Low | 25/1459 (1.7%) | 9/376 (2.4%) | 1.46 (0.68-3.13); *P*=0.33 | |
| Medium | 19/1395 (1.4%) | 4/401 (1.0%) | 0.70 (0.24-2.05); *P*=0.51 | (-)[5] |
| High | 91/6489 (1.4%) | 36/2657 (1.4%) | 1.10 (0.75-1.62); *P*=0.62 | 0.53 |

[1]For those in the HCV antibody$^{-ve}$cohort, HCV diagnosis date was taken to be the same as their respective HCV antibody$^{+ve}$cohort members, for the purpose of analysis.

[2]Based on the likelihood-ratio test comparing the antibody$^{+ve}$cohort to the antibody$^{-ve}$cohort, interaction-terms other than deprivation × HCV were deemed not statistically significant and therefore excluded from the final model.

[3]Adjusted HR and *P* for exposure to HCV-infection within strata of social deprivation. Odds ratios adjusted for age at HCV diagnosis, sex, year of HCV diagnosis and the extra correlation due to the matching.

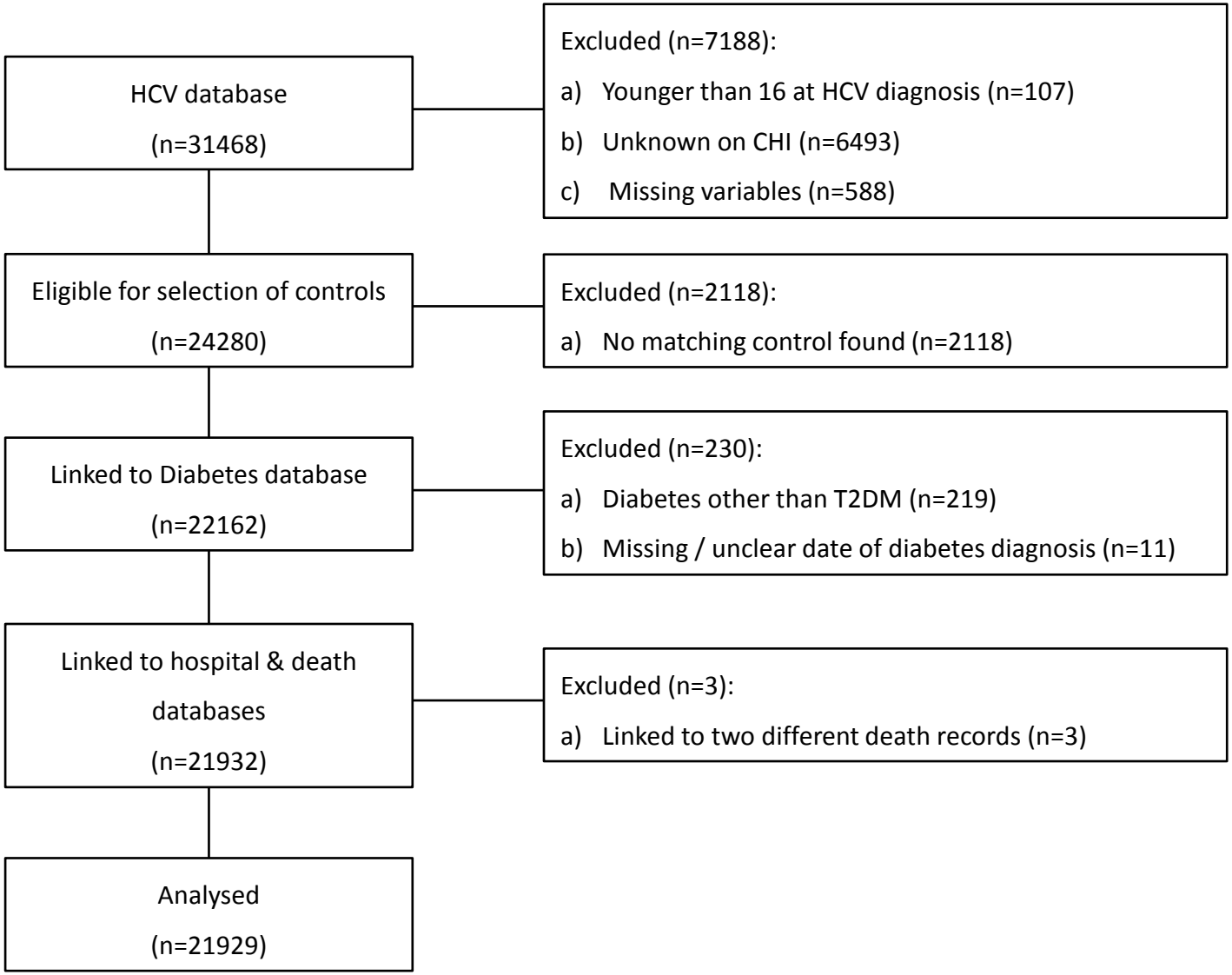
[4]Synergy Index.

[5]To ease comparison between different models, the reference category (antibody$^{-ve}$ and low deprivation) was fixed between models. This caused a negative (invalid) synergy index.

718

719

720 Figure 1: Flowchart describing inclusion (boxes in the left column) and
721 exclusion criteria (boxes in the right column) for the HCV+ve cohort


HCV database
(n=31468)
Excluded (n=7188):
a) Younger than 16 at HCV diagnosis (n=107)
b) Unknown on CHI (n=6493)
c) Missing variables (n=588)
Eligible for selection of controls
(n=24280)
Excluded (n=2118):
a) No matching control found (n=2118)
Linked to Diabetes database
(n=22162)
Excluded (n=230):
a) Diabetes other than T2DM (n=219)
b) Missing / unclear date of diabetes diagnosis (n=11)
Linked to hospital & death databases
(n=21932)
Excluded (n=3):
a) Linked to two different death records (n=3)
Analysed
(n=21929)