Risk of Hepatocellular Carcinoma Among Individuals with Different Aetiologies of Cirrhosis: a Population-Based Cohort Study

Professor Joe West[1*], Dr Timothy R Card[1,2], Professor Guruprasad P Aithal[2], Dr Kate M Fleming[13]

Professor Joe West was supported by a University of Nottingham/Nottingham University Hospitals NHS Trust Senior Clinical Research Fellowship that funded this work

(1) Division of Epidemiology and Public Health, Clinical Sciences Building 2, City Hospital Campus, The University of Nottingham, Nottingham, NG5 1PB, United Kingdom.

(2) NIHR Nottingham Digestive Diseases Biomedical Research Unit at Nottingham University Hospitals NHS Trust and University of Nottingham, Nottingham, UK

(3) Public Health Institute, Liverpool John Moores University, Liverpool L3 2ET, UK

*Corresponding author

Joe West

Professor of Epidemiology

Honorary Consultant Gastroenterologist

Room B113 Clinical Sciences Building 2

University of Nottingham

City Hospital Campus

Hucknall Road

Nottingham NG5 1PB

+44(0)115 8231345 joe.west@nottingham.ac.uk

Abbreviations: GPRD – General Practice Research Database; CumI Cumulative Incidence; HR - Hazard Ratio; 95% CI - 95% Confidence Interval, HCC – Hepatocellular carcinoma

Keywords: epidemiology; cancer; cirrhosis; hepatocellular carcinoma; incidence; alcohol; cryptogenic; chronic viral hepatitis

Author contribution:

Author contribution area	JW	TC	GA	KF
study concept and design	√	√	√	√
acquisition of data	√			
analysis and interpretation of data	√	√	√	√
drafting of the manuscript	√	√	√	√
critical revision of the manuscript for important intellectual content	√	√	√	√
statistical analysis	√			√
obtained funding	√			
technical, or material support	√			
study supervision	√			

No conflict of interest

I, the designated Corresponding Author, on behalf of myself and my co-authors declare that no relevant conflicts of interest exist.

Financial support: JW was supported by a University of Nottingham/Nottingham University Hospital's NHS Trust Senior Clinical Research Fellowship

Word count 3221

Keywords: epidemiology; liver cancer; alcohol; chronic viral hepatitis; risk factor; cryptogenic

Number of tables: 3

Number of figures: 1

Abstract.

Background: Among patients with cirrhosis, only those determined to be at risk for hepatocellular carcinoma (HCC) should undergo surveillance. However, little is known about how different aetiologies of cirrhosis affect risk for HCC.

Aim: To quantify the cumulative incidence of HCC among a representative population of people with cirrhosis of the liver of varying aetiology.

Methods: We identified subjects with hepatic cirrhosis from the UK's General Practice Research Database (1987–2006). Diagnoses of HCC were obtained from linked national cancer registries (1971–2006). Cox proportional hazards regression was used to estimate hazard ratios. The predicted 10-year cumulative incidence of HCC for each aetiology of cirrhosis was estimated while accounting for competing risks of death from any cause and liver transplant.

Results: Among 3107 people with cirrhosis the adjusted relative risk of HCC was increased 2- to 3-fold among people with viral and autoimmune/metabolic aetiologies, compared to those with alcohol-associated cirrhosis. The 10-year predicted cumulative incidence estimates of HCC for each aetiology were: alcohol, 1.2%; chronic viral hepatitis 4.0%; autoimmune or metabolic disease 3.2%; and cryptogenic 1.1%.

Conclusions: In a population-based study in the UK, people with cirrhosis have an estimated cumulative 10-year incidence of HCC of 4% or lower. Cumulative incidence varies with aetiology such that individuals with alcohol or cryptogenic cirrhosis have the lowest risk for
Risk of HCC in cirrhosis

HCC. These findings provide important information for cost-effectiveness analyses of HCC surveillance.

Introduction

Surveillance for hepatocellular carcinoma (HCC) has been suggested by some as an explicit indicator of quality of care in patients with cirrhosis[1]. It remains however a highly controversial topic and a key aspect of such surveillance activities is whether or not they are cost-effective[2, 3]. It is self-evident that the incidence of HCC critically impacts on whether surveillance is cost-effective, and guidance from the American Association for the Study of Liver Diseases (AASLD) based on studies evaluating cost-effectiveness[4-6] recommends that surveillance should only be undertaken in those whose risk of HCC is 1.5% per year or greater (or in hepatitis B 0.2% or greater)[7]. While cirrhosis is the most common underlying condition associated with HCC, the incidence of HCC in cirrhosis due to different aetiologies is not fully known[7]. Whilst the most recent AASLD guidance suggests that the thresholds for HCC incidence to be cost-effective are exceeded in cirrhosis due to hepatitis B or C, primary biliary cirrhosis – now known as Primary Biliary Cholangitis (PBC), genetic haemochromatosis and alpha-1 antitrypsin deficiency, it is explicitly recognized in this guidance that the risk of HCC is not accurately known in many relevant groups.

There is limited evidence to support the reported incidence of HCC; which may explain some of the documented lack of uptake of these guidelines[8]. The available evidence is based principally on studies conducted in tertiary care centers on a small scale [9-13]. These studies are prone to significant biases both in case selection, favouring the inclusion of those with more severe cirrhosis, and with respect to HCC ascertainment, employing active case finding. Recently, Danish evidence derived from a large population based cohort reports 5-year cumulative incidence of only 1% in patients with cirrhosis of an alcoholic aetiology, with HCC barely

Risk of HCC in cirrhosis

contributing to the high mortality seen in these patients[14]. Many studies suggest that other aetiologies of cirrhosis, particularly viral hepatitis, carry a greater risk of HCC. However, there is no study to date that has been able to accurately estimate the rate of HCC in patients with cirrhosis of varying aetiologies drawn from the same underlying population.

We therefore carried out a comprehensive population based study of the risk of HCC in cirrhosis of all aetiologies with a view to improving the evidence-base through which recommendations to current HCC surveillance guidelines can be made to improve their cost effectiveness.

Methods

We conducted a cohort study using linked data from three sources. The General Practice Research Database (GPRD; now the Clinical Practice Research Datalink - CPRD) is a prospectively gathered, anonymised primary care database using data from more than 600 GP practices in the UK, between 1987 to the present[15]. In brief, it provides all recorded primary care data on patients including clinical diagnoses, treatments, and outcomes. Its validity has been tested in numerous studies; for example a systematic review of 357 validation studies showed that overall, a high proportion of cases were confirmed for all diseases with a median of 89%, i.e. 89 of 100 cases with a computerized diagnosis were confirmed based on additional internal or external information[16]. Cancer diagnoses specifically have been validated directly against cancer registration information giving positive predictive values of a GPRD cancer diagnosis of 96% for lung cancer, 92% for urinary tract cancer, 97% for gastro-oesophageal cancer and 98% for colorectal cancer[17]. Hospital Episodes Statistics (HES) is a secondary care database containing data for all hospitalizations in England, including diagnoses and procedures. 51% of English GPRD practices are linked to HES, from April 1997 onwards. Cancer registry data are provided by the National Cancer Intelligence Network and consist of two databases; the Merged Cancer Registry data (1990 to 2006, from English registries only) and the Office for National Statistics (ONS) minimum cancer dataset (1971 to 2006).

We identified people with cirrhosis of the liver from subjects in the whole GPRD who had their first incident recording of cirrhosis, oesophageal varices or portal hypertension within their up to research standard GPRD data between 1987 and 2006 as we have previously described[18]. In this previous study we carried out a validation of the diagnosis in which, in order to assess the

accuracy of the recording of the diagnosis of cirrhosis, paper records from the GPs were requested from a stratified random sample of patients with a diagnostic or therapeutic code for cirrhosis. The patients' paper records (that includes letters from Consultant Hepatologists, liver biopsy results etc) were examined by a consultant hepatologist (GPA). Information was gathered on whether there was any record of cirrhosis, whether this had been confirmed by biopsy and whether there was any record of presumed aetiology of the cirrhosis. Three-quarters of these patients had definite evidence of cirrhosis in the available paper records. Of the 25% of cases where cirrhosis could not be confirmed, all bar one had evidence of chronic liver disease; they were cases of PBC, alcoholic liver disease, Budd-Chari syndrome and autoimmune hepatitis. In subsequent work we have demonstrated that approximately three quarters of those people with a diagnosis of cirrhosis in their primary care record have an inpatient hospitalisation related to cirrhosis[19]. Given that there is a reasonably high proportion of cases identified at a compensated stage of their disease and not all patients will require inpatient hospitalization this provides further evidence of the robustness our definition.

We then restricted our population to only those who were registered in practices with linked cancer registry data. Presumed aetiology of cirrhosis of either alcohol-related, viral hepatitis (B and C), autoimmune or metabolic liver disease (i.e. PBC, haemochromatosis, alpha-1 anti-trypsin deficiency) or other unspecified causes of cirrhosis was defined using appropriate Read codes for these aetiologies. We also used information in the available laboratory results (for example hepatitis B and C positive results, anti-mitochondrial antibody) and linked Hospital Episodes Statistics (using International Classification of Diseases (ICD) 10 codes)[20]. We defined excess alcohol use if there was evidence in the primary or secondary care records of evidence of for

example alcohol abuse, addiction or dependence, 'problem drinking' or referral to alcohol cessation services. Similarly if the weekly alcohol consumption in their primary care records exceeded the Chief Medical Officer's recommended amount (14 units for women, 21 units for men) these patients were ascribed as having alcohol-related cirrhosis. Aetiologies were assigned in a hierarchical fashion ordered chronic viral hepatitis, autoimmune disease, metabolic disease and alcohol excess. Those without any of these aetiologies were grouped together as cryptogenic cirrhosis. Once categorized, these groups were considered mutually exclusive for analysis purposes.

We identified people with HCC using the linked cancer registry data (data available from 1971 – 2006) using ICD 10 and ICD10-O-3 oncology codes[20]. Where necessary ICD9[21] codes were mapped to ICD10. We defined incident HCC as the first occurrence of a record in cancer registry data of a diagnosis coded with a 4 character ICD10 code of C22.0 (malignant neoplasm, liver cell carcinoma) coupled with a histological classification of either 81703 (hepatocellular carcinoma NOS) or 80003 (neoplasm, malignant) in ICD-O-3.

Statistical analysis

Person-time at risk commenced at the first record of cirrhosis in the people with cirrhosis and ended when patients left a participating GP practice or died or the end of cancer registration follow up (31st December 2006) or when liver transplant occurred, whichever came first. We assessed several baseline characteristics including whether the person with cirrhosis had evidence of decompensation (prior to and up to 30 days after entry) or diabetes mellitus.

Incidence rates of HCC were calculated by dividing the number of cases of HCC by total person

Risk of HCC in cirrhosis

years of follow-up and are presented per 1000 person years with 95% confidence intervals. Hazard ratios for HCC were estimated comparing incidence rates by presumed aetiology using Cox proportional hazard's regression adjusted for sex and age at the start of follow up, smoking status, body mass index (BMI) and presence of diabetes mellitus, extracted prior to start of follow up in the study. Model assumptions were checked by plotting proportional hazard and log minus log plots. We fitted a semiparametric proportional hazards model (Fine-Gray method[22, 23]) to estimate the predicted cumulative incidence function for occurrence of HCC accounting for the competing risks of death from any cause and liver transplant. These estimates were calculated at the mean value of all covariates in the model (age, sex, BMI, smoking status and diabetes mellitus) except the primary exposure i.e. aetiology of cirrhosis. All data management and statistical analysis were performed using Stata 14 MP2 (Statacorp, 4905 Lakeway Drive, College Station, Texas 77845 USA).

Results

We identified 3,107 people with cirrhosis from practices with linked cancer registry data available. These subjects contributed 12977 person years respectively to the analyses. Of the people with cirrhosis, 56% were classified as having a presumed aetiology of alcohol, approximately 12% chronic viral hepatitis, 11% autoimmune or metabolic disease and the rest (21%) were classified as cryptogenic. Baseline characteristics such as age and sex varied statistically depending on which aetiology category people were in (table 1). This was also true of all the other factors we measured. As expected, the aetiology with the greatest proportion of those with decompensation was alcohol, and in those with diabetes cryptogenic. More transplants occurred during follow up in those with chronic viral hepatitis than any other group whereas more deaths occurred in the alcohol and cryptogenic groups compared to the others.

Absolute rate of HCC and variation with aetiology

There were 51 incident cases of HCC in the whole population. Overall the incidence rate among people with cirrhosis of all causes was 3.9 per 1000 person years or on average 0.4% per annum. Absolute rates of HCC varied by age, sex and aetiology of disease and are displayed in table 2. As expected they were higher in men compared to women, at older ages and among those with a chronic viral aetiology. When mutually adjusted for age, sex, smoking status, BMI, diabetes mellitus and aetiology using a Cox proportional hazards model people with a chronic viral aetiology were 3 times more likely (HR 3.22 95% CI 1.56-6.65) to develop HCC than those with alcohol related cirrhosis. Those with metabolic or autoimmune diseases were also at increased risk compared to the alcohol group whereas those with the assignation of cryptogenic cirrhosis had a similar incidence of HCC to the alcohol group.

Risk of HCC in cirrhosis

Estimated predicted cumulative incidence of HCC by aetiology after accounting for competing risks

The estimated predicted cumulative incidence of HCC at 1, 5 and 10-years by aetiology among the people with cirrhosis is shown in table 3. For alcohol and cryptogenic aetiology the 10-year risk was less than 2%. The cumulative incidence functions for each aetiology are shown in figure 1.

Discussion

In this study we have quantified the 10-year cumulative incidence of HCC among people with cirrhosis of the liver resulting from alcohol excess; chronic viral hepatitis; autoimmune or metabolic diseases; or of unknown cause using a large, representative, population based cohort study. Overall the incidence of HCC in all these groups was low regardless of aetiology. We found the highest 10-year cumulative incidence of HCC among those with cirrhosis due to chronic viral hepatitis; people with either chronic viral hepatitis or autoimmune/metabolic diseases underlying their cirrhosis had a 2-3 fold increased risk of HCC compared to those with alcoholic cirrhosis. However, in those people we identified as having alcohol as the presumed cause of their cirrhosis or no specific cause (i.e. cryptogenic cirrhosis) the 10-year cumulative incidence rates were less than 2% indicating that surveillance for HCC in these particular groups is unlikely to be cost-effective regardless of other parameters that could influence its cost or outcome.

Strengths and limitations

If there is imprecision in our definition of cirrhosis, the presumed aetiology we have ascribed or the ascertainment of incident HCC our results may be incorrect to some extent. If we have either included people without cirrhosis in our disease cohort or missed people with the disease we may have respectively overestimated or underestimated the incidence of HCC. For example, if we have included patients with alcoholic hepatitis or non-alcoholic steatohepatitis incorrectly as having cirrhosis when they don't, we will have underestimated the incidence of HCC in the alcohol and cryptogenic group respectively. For the definition of cirrhosis we have relied upon the accuracy of recording made by primary care physicians in the electronic health records of

their patients following communication from hepatologists in secondary care about the diagnosis of cirrhosis the latter have made. We have previously validated this approach[18] and shown that it is reliable. In this the recording of cirrhosis mirrors that of a number of other chronic diseases for which validation studies have been conducted[16]. In addition to this, our cohort is of roughly the same age and sex distribution as those reported previously from similar population based or hospital registries from England, Denmark and Sweden[14, 24-26]. For these reasons we think it unlikely we have included many subjects without cirrhosis in our cirrhosis cohort. It is possible however that those people diagnosed with decompensated alcoholic cirrhosis via an emergency admission to hospital who then died rapidly while an inpatient may not have had their diagnosis transmitted to primary care for retrospective addition to their records. By this mechanism we might fail to include some cases of cirrhosis. In the context of our study, i.e. determining the risk of HCC for the purposes of deciding whether or not to carry out surveillance among people with alcoholic cirrhosis, the impact of having potentially excluded these individuals is minimal as they would contribute very little person time at risk and few events during their subsequent follow-up time under surveillance. For the presumed aetiology of disease we have comprehensively searched the primary and secondary care electronic records of the people with cirrhosis which include not only diagnostic and procedure records but also, where available, laboratory and test results. However, we must acknowledge that small variations in the number of cancers diagnosed among each of the aetiologies of liver disease due to misclassification of the aetiology could have led to some differences in our findings. However, with respect to the classification of aetiology, our approach is similar if not more comprehensive than previous work. For example, our ascertainment of excess alcohol use is likely to have been more comprehensive than studies reliant solely on secondary care data. Despite the challenges of

assigning aetiology our distribution of the aetiology of cirrhosis is very similar to that reported from northern European countries that have assembled similar cohorts. By assuming that where a specific aetiology is recorded, for example autoimmune liver disease, that it is solely the cause of cirrhosis in a hierarchical manner we will have inevitably introduced some misclassification. We have chosen to do this purposefully as despite the large size of our cohort it is not large enough to permit us to determine precise rates of HCC among those with multiple aetiologies (for example those with a recording of both alcohol excess and an autoimmune liver disease). The effect of our mutually exclusive categorization is that the rates we have provided may be overestimates of the risk in those with a single aetiology further up our hierarchy if, as has been suggested, those with more than one aetiology have an increased risk[25, 27]. For the ascertainment of incident HCC we have used the linked national cancer registry data which is a method analogous to that carried out in previous reports from Sweden and Denmark[14, 25]. We have used a specific ICD 10 code for HCC coupled with an oncology classification of histology in our definition to avoid, as far as possible, misclassification of, for example, metastatic liver cancer or cholangiocarcinoma which can otherwise occur[28, 29]

We were able to adjust for some important confounders (smoking status, BMI and diabetes mellitus[30, 31]) in our multivariate Cox regression model but we did not have good data available on other potential confounding factors such as ethnicity which may have led to some residual confounding being present by this covariate. In addition, due to the small numbers of events within each mutually exclusive aetiological category, we were unable to present meaningful stratified cumulative incidence rates by any of these covariates to assess for evidence of interactions. We have however taken account of the potential competing risks of death from any

cause and liver transplant on the incidence of HCC via the predicted cumulative incidence function estimated in our analysis.

Other literature

Few studies have been able to study the risk of HCC for these aetiologies among one cohort identified from the same population based source in the manner that we have. The best data for comparison we believe are those derived from the Swedish and Danish registry studies. In 1998 Sørensen et al reported HCC risks among people with cirrhosis diagnosed in Denmark between 1977 and 1989 of alcoholic, chronic hepatitis, PBC and cryptogenic aetiologies[24]. Their approximate crude rates for both alcohol and cryptogenic cirrhosis appear fairly similar to ours (3.4 and 2.5 per 1000 person years respectively). In addition, a more recent analysis of the same data but limited to patients with alcoholic cirrhosis diagnosed between 1993 and 2005 by Jepsen et al., reported annual and cumulative 5 year incidence rates of 0.4% (95% CI 0.34%-0.47%) and 1% (95% CI 0.8%-1.8%) respectively having excluded the first year of follow up[14]. Kuper et al., carried out a similar study using Swedish data and reported cumulative 15 year risks of HCC of 6.2% (95% CI 1%-12.5%) for those with chronic viral hepatitis and 1.1% (95% CI 0.8%-1.5%) for those with alcoholic cirrhosis[25]. Studies from elsewhere in Europe, Japan and the United States of America have all reported higher rates of HCC for the same aetiologies we have examined[9-13]. This is probably partly due to differences in the selection of their cohorts (all being clinic based and therefore likely to have selected more severe cases of cirrhosis), and/ or a differing distribution of aetiology of cirrhosis in those countries, favoring populations with HBV and HCV-related cirrhosis. On this latter point our findings may not be so generalizable to some geographical areas due to their different case-mix of cirrhosis in terms of severity of disease at

Risk of HCC in cirrhosis

diagnosis, quantity of alcohol consumption and underlying prevalence of diseases such as non-alcoholic fatty liver disease.

Clinical implications

Our study contributes important information to the ongoing debate about the utility and implementation of surveillance for HCC among people with cirrhosis[2, 3, 8]. In the AASLD guidelines on this subject[7] it is stated that "for patients with cirrhosis of varying aetiologies, surveillance should be offered when the risk of HCC is 1.5% per year or greater" based on cost-effectiveness modelling[4-6]. In the United Kingdom a Health Technology Assessment economic model[32] found that annual surveillance with a willingness to pay threshold of £30,000 per Quality Adjusted Life Year was only just cost-effective for alcoholic liver disease. Given that our study has found far lower risks of HCC than were used in these economic models it seems highly likely that if they were repeated they would find that surveillance was not cost-effective. Though there may be particular patients with combinations of risk factors where surveillance is warranted our results imply that universal surveillance should not be undertaken on the basis of alcoholic aetiology or in cryptogenic cirrhosis and is likely to be of debatable value in autoimmune and metabolic causes of cirrhosis.

Figure Legends. Figure 1. Estimated predicted cumulative incidence (cumulative proportion with HCC during follow up) for HCC in the cirrhosis cohort by aetiology

Tables.

Table 1. Baseline characterstics, follow up and events among the cirrhosis cohort, presented by aetiology group (n=3107)

	Viral Hepatitis	Autoimmune/metabolic	Alcohol	Cryptogenic	Chi Squared
Total number	374	343	1,743	647	
% Aetiology	12.0	11.0	56.1	20.8	
Median Follow up, years	2.6	3.1	2.6	3.0	
Follow up IQR, years	5.0	5.1	4.9	5.6	
Male	61.5	32.1	65.5	47.8	<0.001
Age					
18-	27.3	10.2	18.7	10.4	
45-	34.5	19.2	28.9	12.4	
55-	20.6	26.2	30.6	21.0	
65-	17.7	44.3	21.9	56.3	<0.001
BMI categories					
<25	30.0	28.6	25.7	20.4	
>=25 to 30	23.3	26.5	20.4	22.3	
>=30	12.8	12.0	12.3	15.9	
Missing	34.0	32.9	41.6	41.4	<0.001
Smoking status					
Current	35.0	16.0	39.8	13.8	
Ex	11.2	18.1	11.2	15.6	
No	28.9	42.3	19.3	36.6	
Missing	24.9	23.6	29.8	34.0	<0.001
Diabetes mellitus	13.6	11.1	13.0	20.1	<0.001
Decompensated at start of follow up	30.8	23.0	35.6	18.7	<0.001
Events					
None	65.8	61.5	57.3	55.5	
Hepatocellular carcinoma	3.2	2.3	1.3	1.4	
Death	27.3	33.5	40.9	42.7	
Liver transplant	3.7	2.6	0.6	0.5	<0.001

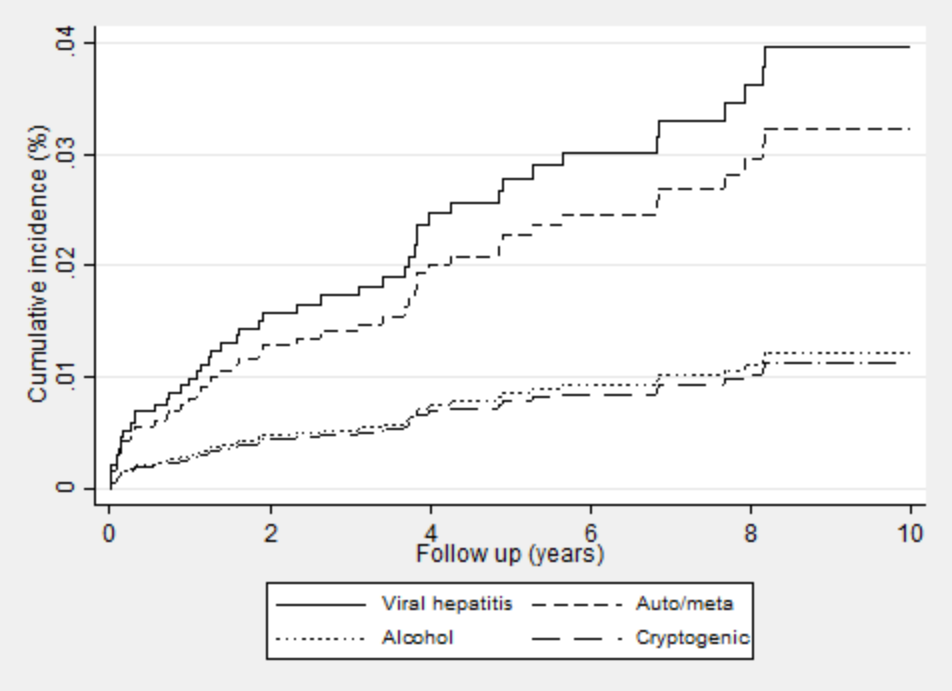
Table 2. Absolute incidence rates of HCC for all follow up time and Hazard Ratios (for HCC incidence) and their 95% confidence intervals for the cirrhosis cohort by age, sex and aetiology

	HCCs during follow up	Person years	Incidence rate per 1000 person years (95% CI)	Hazard Ratio (95% CI)	
Sex					
Male	42	7,146	5.9 (4.3-8.0)	1	
Female	9	5,831	1.5 (0.8-3.0)	0.2	(0.10-0.44)
Age groups					
18-44 years	6	2,390	2.5 (1.1-5.6)	1	
45-54 years	9	3,292	2.7 (1.4-5.2)	0.85	(0.30-2.41)
55-64 years	13	3,674	3.5 (2.1-6.1)	1.27	(0.47-3.42)
65+ years	23	3,621	6.4 (4.2-9.6)	2.73	(1.05-7.10)
Aetiology					
Alcohol	22	6,977	3.2 (2.1-4.8)	1	
Chronic viral hepatitis	12	1,572	7.6 (4.3-13.4)	3.22	(1.56-6.65)
Autoimmune and metabolic diseases	8	1,520	5.3 (2.6-10.5)	2.7	(1.15-6.30)
Cryptogenic	9	2,908	3.1 (1.6-5.9)	0.92	(0.42-2.05)

* adjusted for sex, age groups, smoking status, BMI, diabetes mellitus and aetiology

Table 3. Estimated cumulative incidence (%) of HCC accounting for competing risks of death and liver transplant by aetiology at 1, 5 and 10-years of follow up

Follow time (years)	Viral Hepatitis	Autoimmune/metabolic	Alcohol	Cryptogenic
1	1.0	0.8	0.3	0.3
5	2.8	2.3	0.9	0.8
10	4.0	3.2	1.2	1.1



Footnote: Viral hepatitis = hepatitis B or C; Auto/Meta = Autoimmune or metabolic liver disease; Alcohol = alcoholic; Cryptogenic = no other distinct aetiology identified. Values on the y axis represent proportions i.e. the risk of HCC at 10 years of follow up among those people with cirrhosis with chronic viral hepatitis (B or C) is 4%

Risk of HCC in cirrhosis

References

1. Kanwal F, Kramer J, Asch SM, et al. An explicit quality indicator set for measurement of quality of care in patients with cirrhosis. *Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association* 2010;**8**(8):709-17.

2. Lederle FA, Pocha C. Screening for liver cancer: the rush to judgment. *Ann Intern Med* 2012;**156**(5):387-9.

3. Sangiovanni A, Colombo M. Surveillance for hepatocellular carcinoma: a standard of care, not a clinical option. *Hepatology* 2011;**54**(6):1898-900.

4. Arguedas MR, Chen VK, Eloubeidi MA, Fallon MB. Screening for hepatocellular carcinoma in patients with hepatitis C cirrhosis: a cost-utility analysis. *The American journal of gastroenterology* 2003;**98**(3):679-90.

5. Lin OS, Keeffe EB, Sanders GD, Owens DK. Cost-effectiveness of screening for hepatocellular carcinoma in patients with cirrhosis due to chronic hepatitis C. *Alimentary pharmacology & therapeutics* 2004;**19**(11):1159-72.

6. Sarasin FP, Giostra E, Hadengue A. Cost-effectiveness of screening for detection of small hepatocellular carcinoma in western patients with Child-Pugh class A cirrhosis. *Am J Med* 1996;**101**(4):422-34.

7. Bruix J, Sherman M, American Association for the Study of Liver D. Management of hepatocellular carcinoma: an update. *Hepatology* 2011;**53**(3):1020-2.

8. Davila JA, Morgan RO, Richardson PA, Du XL, McGlynn KA, El-Serag HB. Use of surveillance for hepatocellular carcinoma among patients with cirrhosis in the United States. *Hepatology* 2010;**52**(1):132-41.

9. Fattovich G, Pantalena M, Zagni I, Realdi G, Schalm SW, Christensen E. Effect of hepatitis B and C virus infections on the natural history of compensated cirrhosis: a cohort study of 297 patients. *The American journal of gastroenterology* 2002;**97**(11):2886-95.

10. Velazquez RF, Rodriguez M, Navascues CA, et al. Prospective analysis of risk factors for hepatocellular carcinoma in patients with liver cirrhosis. *Hepatology* 2003;**37**(3):520-7.

11. Mancebo A, Gonzalez-Dieguz ML, Cadahia V, et al. Annual incidence of hepatocellular carcinoma among patients with alcoholic cirrhosis and identification of risk groups. *Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association* 2012;**11**(1):7.

12. Ikeda K, Saitoh S, Koida I, et al. A multivariate analysis of risk factors for hepatocellular carcinogenesis: a prospective observation of 795 patients with viral and alcoholic cirrhosis. *Hepatology* 1993;**18**(1):47-53.

13. Mair RD, Valenzuela A, Ha NB, et al. Incidence of Hepatocellular Carcinoma Among US Patients With Cirrhosis of Viral or Nonviral Etiologies. *Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association* 2012;**10**(12):1412-7.

14. Jepsen P, Ott P, Andersen PK, Sorensen HT, Vilstrup H. Risk for hepatocellular carcinoma in patients with alcoholic cirrhosis: a Danish nationwide cohort study. *Ann Intern Med* 2012;**156**(12):841-7, W295.

15. Herrett E, Gallagher AM, Bhaskaran K, et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). *International journal of epidemiology* 2015.

16. Herrett E, Thomas SL, Schoonen WM, Smeeth L, Hall AJ. Validation and validity of diagnoses in the General Practice Research Database: a systematic review. *British journal of clinical pharmacology* 2010;**69**(1):4-14.

17. Dregan A, Moller H, Murray-Thomas T, Gulliford MC. Validity of cancer diagnosis in a primary care database compared with linked cancer registrations in England. Population-based cohort study. *Cancer Epidemiol* 2012.

18. Fleming KM, Aithal GP, Solaymani-Dodaran M, Card TR, West J. Incidence and prevalence of cirrhosis in the United Kingdom, 1992-2001: A general population-based study. *Journal of hepatology* 2008;**49**(5):732-738.

19. Ratib S, West J, Crooks CJ, Fleming KM. Diagnosis of liver cirrhosis in England, a cohort study, 1998-2009: A comparison with cancer. *American Journal of Gastroenterology* 2014;**109**(2):190-198.

20. International Statistical Classification of Diseases and Related Health Problems. Tenth Revision ed. Geneva: World Health Organisation; 1992.

21. International Classification of Diseases: manual of the international statistical classification of diseases, injuries and causes of death. 9th revision ed. Geneva: WHO; 1975.

22. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. *Journal of the American Statistical Association* 1999;**94**(446):496-509.

23. Austin PC, Lee DS, Fine JP. Introduction to the Analysis of Survival Data in the Presence of Competing Risks. *Circulation* 2016;**133**(6):601-9.

24. Sorensen HT, Friis S, Olsen JH, et al. Risk of liver and other types of cancer in patients with cirrhosis: a nationwide cohort study in Denmark. *Hepatology* 1998;**28**(4):921-5.

25. Kuper H, Ye W, Broome U, et al. The risk of liver and bile duct cancer in patients with chronic viral hepatitis, alcoholism, or cirrhosis. *Hepatology* 2001;**34**(4 Pt 1):714-8.

26. Roberts SE, Goldacre MJ, Yeates D. Trends in mortality after hospital admission for liver cirrhosis in an English population from 1968 to 1999. *Gut* 2005;**54**(11):1615-21.

27. Donato F, Tagger A, Gelatti U, et al. Alcohol and hepatocellular carcinoma: the effect of lifetime intake and hepatitis virus infections in men and women. *American journal of epidemiology* 2002;**155**(4):323-31.

28. West J, Wood H, Logan RF, Quinn M, Aithal GP. Trends in the incidence of primary liver and biliary tract cancers in England and Wales 1971-2001. *Br J Cancer* 2006;**94**(11):1751-8.

29. Khan SA, Emadossadaty S, Ladep NG, et al. Rising trends in cholangiocarcinoma: is the ICD classification system misleading us? *Journal of hepatology* 2012;**56**(4):848-54.

30. Trichopoulos D, Bamia C, Lagiou P, et al. Hepatocellular carcinoma risk factors and disease burden in a European cohort: a nested case-control study. *Journal of the National Cancer Institute* 2011;**103**(22):1686-95.

31. Davila JA, Morgan RO, Shaib Y, McGlynn KA, El-Serag HB. Diabetes increases the risk of hepatocellular carcinoma in the United States: a population based case control study. *Gut* 2005;**54**(4):533-9.

32. Thompson Coon J, Rogers G, Hewson P, et al. Surveillance of cirrhosis for hepatocellular carcinoma: systematic review and economic analysis. *Health Technol Assess* 2007;**11**(34):1-206.

1 STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation	Done
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	Yes
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	Yes
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	Yes
Objectives	3	State specific objectives, including any prespecified hypotheses	Yes
Methods			
Study design	4	Present key elements of study design early in the paper	Yes
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	Yes
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	Yes
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	Yes
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	Yes
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	Yes
Bias	9	Describe any efforts to address potential sources of bias	Yes
Study size	10	Explain how the study size was arrived at	Yes

Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	Yes
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	Yes
		(*b*) Describe any methods used to examine subgroups and interactions	Yes
		(*c*) Explain how missing data were addressed	Yes
		(*d*) If applicable, explain how loss to follow-up was addressed	Yes
		(*e*) Describe any sensitivity analyses	Yes
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	Yes
		(b) Give reasons for non-participation at each stage	Yes
		(c) Consider use of a flow diagram	
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	Yes
		(b) Indicate number of participants with missing data for each variable of interest	Yes
		(c) Summarise follow-up time (eg, average and total amount)	Yes
Outcome data	15*	Report numbers of outcome events or summary measures over time	Yes
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	Yes
		(*b*) Report category boundaries when continuous variables were categorized	Yes
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	Yes

Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	Yes
Discussion			
Key results	18	Summarise key results with reference to study objectives	Yes
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	Yes
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	Yes
Generalisability	21	Discuss the generalisability (external validity) of the study results	Yes
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	Yes

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.