# UNIVERSITY OF BIRMINGHAM

## University of Birmingham Research at Birmingham

# Toronto HCC Risk Index:

Sharma, Suraj A.; Kowgier, Matthew; Hansen, Bettina E.; Pieter Brouwer, Willem; Maan, Raoel; Wong, David; Shah, Hemant; Khalili, Korosh; Yim, Colina; Heathcote, E. Jenny; Janssen, Harry L. A.; Sherman, Morris; Hirschfield, Gideon; Feld, Jordan J.

Link to publication on Research at Birmingham portal

# Accepted Manuscript

Toronto HCC Risk Index: A validated scoring system to predict 10-year risk of HCC in patients with cirrhosis

Suraj A Sharma, Matthew Kowgier, Bettina E Hansen, Willem Pieter Brouwer, Raoel Maan, David Wong, Hemant Shah, Korosh Khalili, Colina Yim, E. Jenny Heathcote, Harry L.A. Janssen, Morris Sherman, Gideon M. Hirschfield, Jordan J Feld

Please cite this article as: Sharma, S.A., Kowgier, M., Hansen, B.E., Pieter Brouwer, W., Maan, R., Wong, D., Shah, H., Khalili, K., Yim, C., Jenny Heathcote, E., Janssen, H.L.A., Sherman, M., Hirschfield, G.M., Feld, J.J., Toronto HCC Risk Index: A validated scoring system to predict 10-year risk of HCC in patients with cirrhosis, *Journal of Hepatology* (2017), doi: http://dx.doi.org/10.1016/j.jhep.2017.07.033

**Toronto HCC Risk Index: A validated scoring system to predict 10-year risk of HCC in patients with cirrhosis**

Suraj A Sharma[1], Matthew Kowgier[1,5], Bettina E Hansen[2], Willem Pieter Brouwer[2] , Raoel Maan[2] David Wong[1], Hemant Shah[1], Korosh Khalili[3], Colina Yim[1], E. Jenny Heathcote[1], Harry L.A. Janssen[1,2], Morris Sherman[1], Gideon M Hirschfield[4]* and Jordan J Feld[1]*

*These authors contributed equally to the supervision of this study

1. Toronto Centre for Liver Disease, University Health Network, University of Toronto

2. Department of Gastroenterology and Hepatology, Erasmus MC University Medical Centre, Rotterdam, The Netherlands

3. Department of Medical Imaging, University Health Network, University of Toronto

4. Centre for Liver Research and NIHR Biomedical Research Unit, University of Birmingham, Birmingham, UK

5. Dalla Lana School of Public Health, University of Toronto

**Address for correspondence:**

Attn: Dr Jordan J. Feld,
Toronto Center for Liver Disease
Sandra Rotman Centre for Global Health
University Health Network, University of Toronto
6B-Fell Pavilion Rm 158
399 Bathurst Street, Toronto, Ontario
M5T 2S8, Canada

Jordan.feld@uhn.ca

**Manuscript characteristics:** Abstract=364; Text= 3637 words; Tables= 4; Figures=4, Supplemental tables=2, Supplemental Figures=2

**Abbreviations:** HCC – hepatocellular carcinoma, HBV – Hepatitis B virus, HCV – Hepatitis C virus, AST – aspartate aminotransferase, ALT – alanine aminotransferase, APRI – AST to platelet ratio index, FIB-4 – Fibrosis-4 score, AIH – Autoimmune hepatitis, PBC – Primary biliary cirrhosis, PSC – Primary sclerosing cholangitis, NAFLD – Non-alcoholic fatty liver disease, ALD –

1

alcoholic liver disease, THRI – Toronto HCC Risk Index, HR – Hazard Ratio, CI – confidence interval

**Abstract**

2

**Background:** Current guidelines recommend biannual surveillance for hepatocellular carcinoma (HCC) in all patients with cirrhosis, regardless of etiology. However, HCC incidence is not well established for many causes of cirrhosis.

**Aim:** To assess the disease-specific incidence of HCC in a large cohort of patients with cirrhosis and to develop a scoring system to predict HCC risk.

**Methods:** A derivation cohort of patients with cirrhosis diagnosed by biopsy or non-invasive measures was identified through retrospective chart review. The disease-specific incidence of HCC was calculated according to etiology of cirrhosis. Factors associated with HCC were identified through multivariable Cox regression and used to develop a scoring system to predict HCC risk. The scoring system evaluated in an external cohort for validation.

**Results:** Of 2,079 patients with cirrhosis and ≥6 months follow-up, 226 (10.8%) developed HCC. The 10-year cumulative incidence of HCC varied by etiologic category from 22% in patients with viral hepatitis, to 16% in those with steatohepatitis and 5% in those with autoimmune liver disease (p<0.001). By multivariable Cox regression, age, sex, etiology and platelets were associated with HCC. Points were assigned in proportion to each hazard ratio to create the Toronto HCC Risk Index (THRI). The 10-year cumulative HCC incidence was 3%, 10% and 32% in the low (<120 points) medium (120-240) and high (>240) risk groups respectively, values that remained consistent after internal validation. External validation was performed on a cohort of patients with PBC, HBV and

3

HCV cirrhosis (n= 1,144) with similar predictive ability (Harrell's *c-statistic* 0.77) in the validation and derivation cohorts.

**Conclusion:** HCC incidence varies markedly by etiology of cirrhosis. The THRI, using readily available clinical and laboratory parameters, has good predictive ability for HCC in patients with cirrhosis, and has been validated in an external cohort. This risk score may help to guide recommendations regarding HCC surveillance among patients with cirrhosis.

**Keywords:** Cirrhosis, hepatocellular carcinoma, HCC, Toronto hepatoma risk index (THRI), cumulative incidence

4

**Introduction**

Ambulatory management of cirrhosis is of increasing importance with rising rates of chronic liver disease and the associated complications of end-stage cirrhosis. One such complication, hepatocellular carcinoma (HCC), is the fifth most common malignancy among men globally, and seventh most common malignancy among women, leading to more than 700,000 deaths annually[1-3]. In the US, the age-adjusted incidence of HCC rose from 1.6 to 4.5 per 100,000 people between 1975 and 2005[4]. Currently, one-year survival for HCC is still less than 50%, despite improvements in early recognition using surveillance and the availability of better therapies[4-6].

Current guidelines recommend twice-yearly ultrasounds for HCC surveillance in all patients with cirrhosis [6,7]. These recommendations[8] arise from studies evaluating HCC doubling time[9,10], cost-effectiveness[11], as well as one randomized trial of HCC surveillance in patients with chronic hepatitis B (CHB) infection from China[12]. AASLD guidelines state that HCC surveillance is cost-effective at an annualized incidence of 1.5% or above in patients with cirrhosis[6].

However, the risk of HCC in cirrhotic patients is known to vary with etiology, age, gender and other factors[13-16]. Scoring systems have been developed for HCC-risk prediction in patients with specific causes of liver disease such as CHB, however etiology-independent risk stratification is currently not possible. As a result, HCC surveillance is recommended for all patients with cirrhosis, regardless of the etiology of liver disease or the presence of other risk factors.

5

Using sequential data from a large cohort of well-characterized patients with cirrhosis, we aimed to assess the disease-specific incidence of HCC. Combining etiology of liver disease with other risk factors for HCC, we developed and both internally and externally validated a scoring system to accurately predict the 5- and 10-year risk of HCC among patients with cirrhosis.

**Methods**

**Patients**

Complete records were obtained for all patients seen at the Toronto Western Hospital Liver Centre between January 1, 2000 and December 31, 2009. The study was approved by the Research Ethics Board at the University Health Network.

**Diagnosis of cirrhosis and HCC**

To identify a cohort of patients with cirrhosis, patients were evaluated in a step-wise manner. The AST to platelet ratio index (APRI) [(AST/upper limit of normal/platelet count)× 100] and FIB-4 [age $\times$ AST/platelet count $\times \sqrt{}$ALT] were used to identify patients with probable cirrhosis. To maximize specificity for the diagnosis of cirrhosis, only those with an APRI value ≥ 1.0 (Specificity 75%, sensitivity 89%) [17] *and* a FIB-4 value ≥ 3.25 (Specificity 98%, Sensitivity 37%) [18] at either the first or last clinic visit were included in the analysis. To confirm the diagnosis of cirrhosis, the charts of all patients identified using non-invasive markers, were reviewed to document at least one of: clinical (varices, variceal hemorrhage, ascites), pathological (F3 or F4 on biopsy) or radiological

6

(coarse/nodular/lobar redistribution on ultrasound) evidence of cirrhosis. Patients with no confirmatory features were excluded from further analysis. The primary etiology of cirrhosis was assigned according to the 'diagnosis field' in the electronic patient record, as determined by the consultant hepatologist. Patients with steatohepatitis were categorized as having alcoholic cirrhosis if alcohol was identified as the cause of liver disease by the treating physician in the clinical record. Patients with both, chronic viral hepatitis, and a history of alcohol intake or non-alcoholic steatohepatitis were categorized as having viral hepatitis for the purpose of this analysis. Treatment status for HCV and HBV patients was also recorded. The category of 'other' included patients with a diagnosis of Wilson disease, hereditary hemochromatosis, alpha-1 antitrypsin deficiency and cryptogenic cirrhosis.

The diagnosis of HCC was made in accordance with AASLD guidelines[6] and required either a positive ultrasound with confirmation on a second dynamic imaging modality (CT or MR) or a diagnostic biopsy. Demographic, clinical and laboratory parameters were recorded for all visits including: age, sex, etiology of cirrhosis, body mass index (BMI), liver enzyme (AST, ALT, ALP), liver function (bilirubin, albumin, INR) and platelet levels. Missing data were imputed using last observation carried forward within a window of 3 months. Follow-up time was censored at the last clinic visit or diagnosis of HCC. HCCs diagnosed within 6 months of referral were excluded from further analysis. Patients with primary sclerosing cholangitis (PSC) were censored at the time of diagnosis of cholangiocarcinoma and these were not included for HCC incidence.

7

**Statistical Analysis**

The incidence of HCC was determined for each etiology of cirrhosis using the Kaplan-Meir (K-M) method, with curves compared using the log-rank test. Continuous variables were compared using Student's t-test or Wilcoxon Rank Sum test. Categorical variables were compared using Fisher's exact test.

Cox proportional hazards regression was used to identify factors at the time of the diagnosis of cirrhosis that were significantly associated with development of HCC. For continuous covariates, threshold values were determined to improve clinical utility. Covariates associated by univariate analysis (p<0.15) were combined in a backward stepwise multivariable model. Factors not significant but of potential clinical importance (MELD, BMI) and interaction terms were also assessed to exclude important confounding. The final model was selected by minimizing the Akaike information criterion (AIC). Schoenfeld's residuals were evaluated to ensure validity of the assumptions used in Cox methodology. Statistical analysis was carried out using Stata 12 (Texas, USA), and R version 3.1.2 (R core development team, 2010).

**Internal validation**

To assess the degree of optimism, or over-fitting, in the performance of the final predictive model, we estimated the optimism using Harrell's method[19] and then computed an optimism-corrected C-index. To quantify the degree of over-fitting of the final model we also estimated the shrinkage factor using the bootstrap[19].

8

The bootstrap analysis was done using R V3.2.2 (RMS package) with 1000 iterations performed.

**Derivation of Toronto HCC Risk Index**

To improve clinical utility of the model, a simple scoring system was derived by assigning points for each covariate in proportion to the hazard ratios in the final multivariable model, as previously described[20]. The THRI score was divided into three equally spaced categories to obtain low (<120), medium (120-240) and high (>240) risk thresholds. Using the 5- and 10-year cumulative HCC incidence at each point score, a nomogram was developed to predict HCC risk by total point score. The performance of the original multivariable model and of the THRI strata were then evaluated using an optimism-corrected Harrell's *c*-statistic[21,22]. A c-statistic of 0.7-0.8 is considered good, while >0.8 is considered excellent [18].

**External Validation**

The THRI score was validated in an external cohort of cirrhotic patients from Rotterdam, Netherlands. Validation was performed in the external cohort with the variables in categorical format with the pre-specified cut-offs determined in the discovery cohort[23]. Data were obtained from four published cohorts[24-27]. Cirrhosis was defined by the presence of clinical or radiologic features, or liver biopsy. The THRI and the predicted HCC incidence (calculated using the formula $S_1 = 1 - S_0(t)^{\exp(\beta * THRI)}$) were calculated for each patient. $S_0(t)$ is the baseline HCC free-survival in the derivation cohort and $\beta$ is the estimated effect of THRI in the

9

derivation cohort. Harrell's *c*-statistic was calculated to evaluate the performance

of the model. Discrimination of the THRI risk score were assessed by plotting the

observed HCC incidences of the risk groups using the Kaplan-Meier method,

together with the predicted HCC incidence from the formula above (figure 4).

## Results

### Identification of the study cohort

A total of 12,199 patients (Supplementary Figure 1) were seen at the Liver

Centre between January 1, 2000 and December 31, 2009. Of these, 3,064 (25%)

patients had an APRI≥1.0 *and* FIB-4≥3.25, suggesting probable cirrhosis. The

charts of all 3,064 patients were reviewed. Of these, 2,416 patients (79%) had

evidence of varices/ascites (n=1,567), F3/F4 on biopsy (n=1,178) and/or a

coarse/nodular liver or lobar redistribution on ultrasound (n=2,121). Of those with

confirmed cirrhosis (n=2,416), 2,079 (98%) patients were followed for ≥ 6 months,

and formed the study cohort. Using AASLD criteria, 226 patients (10.8%) were

diagnosed with HCC during follow-up.

### Baseline characteristics

The baseline characteristics of the study population are shown in Table 1. The

most common causes of cirrhosis were chronic hepatitis C (CHC, n=883),

chronic hepatitis B (CHB, n=396), alcoholic (ALD, n=228) and non-alcoholic fatty

liver disease (NAFLD, n=111). A large number of patients with autoimmune liver

10

diseases (n=260) were also included: autoimmune hepatitis (AIH) 112, primary biliary cirrhosis (PBC) 108 and PSC 40.

Patients who developed HCC were more likely to be male (HCC: 81% vs. no HCC: 58%, (p<0.001) and older at baseline (HCC: 57±11 vs. non-HCC: 54±13, p<0.001) than patients who did not develop HCC during follow-up. The mean duration of follow-up was significantly shorter in the HCC group (HCC: 3.9 ± 3.6 years vs. No HCC: 6.1 ± 2.6, years; p<0.001), presumably because follow-up was censored at HCC diagnosis. Among patients who developed HCC during follow-up, a higher proportion had chronic viral hepatitis (79.6%) and fewer had an autoimmune liver disease (AIH/PBC/PSC; 3.1%) or 'other' etiology (4.9%) as the cause of cirrhosis.  As expected, there were differences in demographic (sex, age) and clinical features (BMI) between patients with different etiologies of liver disease (Table 4).

**HCC incidence by etiology**

The cumulative incidence of HCC was calculated for each etiology of cirrhosis. Figure 1 shows the 5- and 10-year cumulative incidence for HCC according to etiology. HCC occurred most frequently in patients with chronic viral hepatitis with a 10-year cumulative incidence of 23.2% among patients with CHB and 21.1% among those with CHC (p=0.21). Viral clearance is associated with a significant reduction in the incidence of HCC in patients with CHC-related cirrhosis [27].   Patients with CHC who achieved SVR had a 10-year incidence of

11

HCC of 7%, significantly lower than that seen in patients who remained viremic (24.6%). This translated to an incidence rate of 26.2 (CHB), 27.0 (CHC without SVR) and 7.0 (CHC with SVR) per 1,000 patient-years respectively.

The HCC incidence was in the intermediate range in the steatohepatitis cohort (ALD and NAFLD). Patients with ALD had 10-year cumulative incidence of HCC of 17.7% compared to 12.8% among those with NAFLD (p=0.59). The corresponding incidence rates were 18.4 (ALD, 95% CI 11.9-28.5) and 14.4 (NAFLD, 95% CI 7.2-28.8) per 1,000 patient-years, respectively.

Patients with 'other' liver diseases had a cumulative HCC incidence of 4.5% at 5 years and 8.0% at 10 years, corresponding to an incidence rate of 10.1 (95% CI 5.6-18.3) per 1,000 patient-years.

By contrast, HCC incidence was lower in those with autoimmune liver disease. In patients with AIH, no cases of HCC were diagnosed at 5 years of follow-up, and the cumulative incidence was 1.7% at 10 years. For patients with PBC the cumulative incidence of HCC was 6.1% at 10 years, while in patients with PSC, it was 6.9%. The corresponding incidence rates were 1.3 for AIH (95% CI 0.2-9.4), 5.5 for PBC (95% CI 2.0-14.6) and 7.2 for PSC (95% CI 1.8-28.9) per 1,000 patient-years of follow-up (log-rank test p=0.09). One patient was diagnosed with mixed HCC-Cholangiocarcinoma, but was excluded because the tumor was diagnosed within 6 months of the start of follow-up.

The cumulative incidence was also compared between four broad etiologic categories of cirrhosis: viral hepatitis (CHB and CHC), steatohepatitis (ALD and

12

NAFLD), autoimmune (AIH, PBC and PSC) and other liver diseases. Patients with viral hepatitis had the highest 10-year cumulative incidence of HCC at 21.7% compared to 16.3% in those with steatohepatitis, 4.6% in those with autoimmune liver disease and 8.0% in those with other liver disease. By the log-rank test, the cumulative incidence of HCC was significantly higher in patients with viral hepatitis or steatohepatitis compared to those with autoimmune liver disease (p<0.001). The difference in incidence between viral etiologies and steatohepatitis did not reach statistical significance (p=0.09) (Figure 1).

Cox regression was used to compare the incidence of HCC by etiology of liver disease. Using autoimmune liver disease as a reference, the hazard ratios (HR) and 95% confidence intervals (CI) for HCC were 5.8 (95% CI 2.8-12.4) for viral hepatitis and 4.1 (95% CI 1.8-9.5) for steatohepatitis. After adjusting for age and sex, the HR for HCC was 4.0 (95% CI 1.9-8.6) for all viral hepatitis, 2.7 for steatohepatitis (95% CI 1.2-6.3) and 2.5 for other etiologies (95% CI 0.95-6.31) compared to autoimmune liver disease (HR=1).

**Identification of risk factors for HCC and derivation of the Toronto HCC Risk Index (THRI)**

By univariate Cox regression, age, sex, etiology, APRI, FIB-4, MELD and platelet count were significantly associated with the development of HCC (Table 2). Schoenfeld's residuals confirmed the validity of the assumptions used in the Cox methodology.

13

Variables were combined in a backward stepwise multivariable Cox model avoiding collinear variables like FIB-4 and APRI. The final multivariable model with the lowest AIC included age, male sex, etiology of cirrhosis (HBV, HCV, HCV with SVR, Steatohepatitis, Autoimmune and Other), and platelet count (Table 2). No significant interactions were detected between variables. Treatment for HBV was not an independent predictor of HCC in this cohort, since the majority (76%) of patients were on treatment.

The Toronto HCC risk index (THRI) was then derived using HRs of the variables in the multivariable model (Table 3). Point scores were assigned to each covariate in proportion to the HR for that variable, using a common denominator to compare the HR of significant covariates and then rounding scores to the nearest integer, following the method of Yang and colleagues [28].

Patients were then stratified into three equally-spaced groups based on the THRI score (<120, 120-240, and >240 points). These cutoffs were chosen in order to create maximum separation between the risk groups. Patients with a low THRI score (<120) had a 5-year cumulative HCC incidence of 1.2% (95% CI 0.13-2.2%) and a 10-year cumulative incidence of 2.7% (95% CI 0.3-5.1%). Patients in the intermediate-risk group (120-240 points) had a cumulative incidence of 4.4% (95% CI 3.1-5.6%) at 5 years and 9.8% (95% CI 7.3-12.3%) at 10 years. The high-risk group (THRI >240) had a higher 5- and 10-year cumulative incidence of HCC of 15.4% (95% CI 12.7-17.9%) and 32.1% (95% CI 27.4-36.4%), respectively. The cumulative HCC incidence by THRI score is shown in figure 2. The modified Harrell's c-statistic for the multivariable model was 0.76

14

(95% CI: 0.72-0.79) with platelets as a continuous variable, and 0.75 with platelets as a categorical variable. This analysis was repeated after excluding patients with F3 (advanced fibrosis) and no other clinical, radiologic or laboratory manifestations of cirrhosis. The c-statistic for this subset was 0.74.

The 5- and 10-year HCC incidence was obtained for each THRI point score. These data were used to develop a nomogram for HCC incidence by point score to allow for more individualized risk assessment (Figure 3).

**External Validation**

External validation of the model was performed in a cohort of 1,144 cirrhotic patients from Rotterdam, Netherlands (PBC n=408, HBV n=253, HCV n=301, HCV-SVR n=182). A total of 107 patients (9.4%) among the validation cohort developed HCC during follow-up. Cumulative HCC incidence in the low, moderate and high risk groups at 5 years was 1.1% (95%CI 0.0- 2.3%), 4.9% (95%CI 2.4- 7.5%), 13.1% (95%CI 9.6- 16.6%) and at 10 years 3.6% (95%CI 1.1-6.2%), 8.9% (95%CI 4.6- 13.2%), 24.3% (95%CI 18.4- 30.1%) respectively (Figure 4). The predicted mean HCC incidence coincides with the observed Kaplan-Meier HCC-incidence of the 3 risk groups (Figure 4) reflecting a good calibration (the overall calibration slope is 0.95 (SE 0.05), not significantly different from 1) and with similar predictive ability (Harrell's *c-statistic* 0.77) in the validation cohort as seen in the derivation cohort. The calibration slope of the prognostic risk score in the validation dataset was 0.93 (SE 0.10), which is not significantly different from 1 (p=0.48), and thus the calibration seems to be preserved.

15

**Discussion**

HCC is one of the most common causes of morbidity and mortality for patients with chronic liver disease. HCC screening protocols consume significant time and resources and currently do not stratify patients based on their individualized risk of HCC.  Our large clinic-based cohort of cirrhotic patients (n=2,079), with long-term follow-up (median 71 months), encompassing a comprehensive range of liver disease, provides important data on HCC incidence with development of a simple scoring system to stratify patients based on their risk of HCC over time.

Current guidelines for HCC surveillance advise biannual surveillance for all patients with cirrhosis using ultrasound[6].  This recommendation applies to all patients regardless of the etiology of cirrhosis and other risk factors.  The 6-month interval was selected based on tumor doubling time and in part from cost-effectiveness analyses showing that an annualized HCC incidence of 1.5% or more would make surveillance cost-effective[11]. However, the disease-specific risk for HCC has not been clearly established for all etiologies of cirrhosis. Our data clearly show that the etiology of cirrhosis is a critical determinant of HCC risk that can be summarized with etiologic categories.

Patients with chronic viral hepatitis had the highest cumulative risk of HCC, whereas those with autoimmune liver diseases had the lowest risk. For many etiologies of cirrhosis, the annualized incidence of HCC was lower than the AASLD recommended threshold for biannual screening to be cost-effective (1.5% per year). In our cohort, these included AIH (0.17% per year), PBC (0.6% per year), PSC (0.7% per year) and possibly NAFLD (1.3% per year). Notably

16

patients with CHC who were successfully treated and achieved a sustained virological response (SVR) had an HCC incidence similar to that in patients with autoimmune liver disease.

There is some variability in the published estimates of HCC incidence by etiology of cirrhosis, partially because many studies report incidence over the study period rather than annualized risk. The observed annualized incidence of HCC in this study was similar to estimates reported previously for each specific etiology of cirrhosis[14,29-31].

To aid in risk stratification for individual patients, an easy-to-use scoring system, using a combination of routine clinical and laboratory parameters (age, sex, etiology and platelets), was developed to predict the risk of HCC in patients with cirrhosis. The THRI was able to stratify patients into low, intermediate and high risk of HCC. Patients in the low-risk group had an annual HCC incidence of approximately 0.3% - well below the threshold for cost-effective surveillance – whereas the incidence in the high-risk group was 3.2% per year, suggesting that these patients require close follow-up.

To assess whether patients with cryptogenic cirrhosis should have been classified with the 'Other' group vs the 'Fatty liver' group, we evaluated the risk of HCC in the 'Other' group versus the 'Cryptogenic' group using the log-rank test and found that there was no significant difference. Of the 35 patients with cryptogenic cirrhosis in this cohort, only 2 developed HCC, which is lower than would be expected if they all had NAFLD. Therefore, although it may be that some of these patients had NAFLD, none had associated clinical conditions and

17

thus it seems more appropriate to classify those with 'Cryptogenic' cirrhosis in the 'Other' category

The strengths of this study include the large patient cohort, the long-term follow-up and the large number of patients with autoimmune liver disease included. Importantly external validation confirmed the good predictive ability of the scoring system in an independent cohort of cirrhotic patients of varying etiology. While previous studies have developed etiology-specific HCC-risk scores[20,28], the THRI allows for HCC risk stratification for all cirrhotic patients, regardless of etiology. Importantly, as patients age, and disease progresses, the THRI score will also increase. Alternatively, if a patient with CHC is treated and achieves SVR, the THRI and corresponding risk of HCC will go down.

This study has limitations, largely relating to its retrospective design.  The scoring system was developed in a single large tertiary care liver clinic, however it was validated in an independent population with similar predictive ability. Given the retrospective nature of the THRI score development further prospective studies including data on other risk factors and causes of morbidity and mortality are required to confirm the predictive ability of the THRI score.  Notably, 119 patients were followed for between 6-12 months meaning that they only underwent one surveillance ultrasound during follow-up. These patients were included in the final analysis to maximize the power of the scoring system, noting that their short follow-up time is accounted for using survival analysis. The diagnosis of NAFLD and alcoholic liver disease was recorded based on the clinical opinion of the hepatologist that assessed the patient rather than by using standardized alcohol

18

assessment tools, which may have led to some degree of misclassification. However, in the final risk score, both NAFLD and alcoholic liver disease are categorized in one subgroup (Steatohepatitis), since the HCC risk among these groups is similar. Data on the race/ethnic background of patients was not available in the databases used nor was data on competing risks for mortality and as such the risk of HCC was not adjusted for competing risks, which may be relevant given that all patients had cirrhosis.

Data on adherence to 6-monthly ultrasound were not available for all patients. Among patients surveyed at UHN, the adherence to screening guidelines was poor (35%), but in line with prior estimates[32]. However, given the rapidly progressive nature of untreated HCC, it is very unlikely that this affected the incidence estimates significantly, and likely did not affect the comparison between etiologic categories. Lastly, data on competing risks of mortality from other causes was not available to us. This may potentially impact the estimates of HCC incidence obtained in this study. However, given that patients were cirrhotic at entry, we do not expect this to have a marked differential impact on estimates between etiologic categories.

In conclusion, in a large cohort of cirrhotic patients with a representative spectrum of underlying liver disease, HCC risk varied widely according to the etiology of cirrhosis. Routine clinical and laboratory data were used to develop an easy-to-use scoring system to predict HCC risk over time, which was validated in an independent cohort of patients with cirrhosis. This tool may be used to identify patients at very low risk of HCC, who may not require surveillance.

19

**Table 1: Baseline characteristics of patients with cirrhosis (follow-up >6months) and patients with HCC**

| | Derivation cohort | | Validation cohort | |
|---|---|---|---|---|
| | **All patients (n=2,079)** | **HCC (n=226)** | **All patients (n=1,144)** | **HCC (n=107)** |
| **Mean Age (yrs)** ±SD | 53.9±12.4 | 57.3±11.0 | 51.2±11.6 | 54.2±10.0 |
| (Range) | (15-93) | (20-83) | (14-81) | (26-76) |
| **Sex** | | | | |
| **Female n (%)** | 828 (39.9) | 44 (19.5) | 569 (49.7) | 30 (28.0) |
| **Male n (%)** | 1251 (60.1) | 182 (80.5) | 575 (50.3) | 77 (72.0) |
| **Mean follow-up (yrs)**±SD | 5.9±3.5 | 3.9±2.6 | 7.3±5.1 | 5.4±4.4 |
| Median | 5.4 | 3.4 | 6.2 | 4.2 |
| (Range) | (0.5-18.6) | (0.5-15.3) | (0.5-26.7) | (0.5-22.0) |
| **Etiology, n (%)** | | | | |
| **HBV** | 396 (19.0) | 69 (30.5) | 253 (22.1) | 23 (21.5) |
| **HCV-No SVR** | 692 (33.3) | 102 (45.1) | 301 (26.3) | 60 (56.1) |
| **HCV-SVR** | 191 (9.2) | 9 (3.9) | 182 (15.9) | 7 (6.5) |
| **ALD** | 228 (11.0) | 20 (8.8) | | |
| **NAFLD** | 111 (5.3) | 8 (3.5) | | |
| **AIH** | 112 (5.4) | 1 (0.4) | | |
| **PBC** | 108 (5.2) | 4 (1.8) | 408 (35.7) | 17 (15.9) |
| **PSC** | 40 (1.9) | 2 (0.9) | | |
| **Other** | 201 (9.7) | 11 (4.9) | | |
| **Mean BMI (kg/m²)**±SD | 26.4±5.2 | 26.1±4.6 | Not available | Not available |
| (Range) | (14.9-52.5) | (18.3-40.7) | | |

21

**Table 2: Factors associated with HCC by Cox proportional hazards regression**

| Covariate | Univariable HR (95% CI) | p-value | Multivariable HR (95% CI) | p-value |
|---|---|---|---|---|
| **Etiology** | | | | |
| Autoimmune | 1 | | 1 | |
| HCV-SVR | 1.74 (0.65-4.67) | 0.27 | 1.04 (0.38-2.83) | 0.94 |
| Other | 2.46 (0.95-6.34) | 0.06 | 1.65 (0.63-4.27) | 0.31 |
| Steatohepatitis | 4.14 (1.80-9.51) | 0.001 | 2.12 (0.91-4.92) | 0.08 |
| HCV | 6.74 (3.13-14.50) | <0.001 | 3.83 (1.76-8.33) | 0.001 |
| HBV | 6.50 (2.99-14.16) | <0001 | 3.66 (1.66-8.07) | 0.001 |
| | | | | |
| **Age** | 1.03 (1.02-1.05) | <0.001 | 1.04 (1.03-106) | <0.001 |
| | | | | |
| **Age category** | | | | |
| <45 | 1 | | 1 | |
| 45-60 | 1.91 (1.27-2.87) | 0.002 | 1.94 (1.28-2.93) | 0.002 |
| >60 | 2.94 (1.94-4.46) | <0.001 | 3.77 (2.45-5.80) | <0.001 |
| | | | | |
| **Male Sex** | 2.84 (2.04-3.94) | <0.001 | 2.92 (2.07-4.12) | <0.001 |
| | | | | |
| **Platelets (x10^9/L)** | 0.99 (0.99-0.99) | <0.001 | 0.99 (0.99-1.00) | <0.001 |
| | | | | |
| **Platelet Category (x10^9/L)** | | | | |
| <80 | 1 | | 1 | |
| 80-139 | 0.78 (0.57-1.07) | 0.12 | 0.78 (0.57-1.07) | 0.13 |
| 140-200 | 0.40 (0.27-0.59) | <0.001 | 0.40 (0.27-0.60) | <0.001 |
| >200 | 0.25 (0.16-0.41) | <0.001 | 0.31 (0.19-0.50) | <0.001 |
| | | | | |
| **APRI** | | | | NS |
| <1.5 | 1 | | | |
| 1.5-4 | 1.69 (1.26-2.28) | <0.001 | | |
| >4 | 1.57 (1.09-2.26) | 0.01 | | |
| | | | | |
| **FIB-4** | | | | NS |
| 0-1.5 | 1 | | | |
| 1.5-5.5 | 2.55 (1.44-4.53) | 0.001 | | |
| >5.5 | 4.10 (2.29-7.34) | <0.001 | | |
| | | | | |
| **MELD>8** | 1.61 (1.23-2.11) | 0.001 | | NS |

22

**Table 3: Components of the Toronto HCC Risk Index**

| Risk Factor | Score |
| --- | --- |
| **Age** | |
| <45 | 0 |
| 45-60 | 50 |
| >60 | 100 |
| **Etiology** | |
| Autoimmune | 0 |
| HCV SVR | 0 |
| Other | 36 |
| Steatohepatitis | 54 |
| HCV | 97 |
| HBV | 97 |
| **Gender** | |
| Female | 0 |
| Male | 80 |
| **Platelets** | |
| >200 | 0 |
| 140-200 | 20 |
| 80-139 | 70 |
| <80 | 89 |
| Total | 0-366 |

23

**Table 4: Crude HCC incidence according to etiology**

| | N at baseline | 5-year cumulative incidence (95% CI) | 10-year cumulative incidence (95% CI) | Incidence per 1000 person-years (95% CI) |
|---|---|---|---|---|
| **Etiology** | | | | |
| HBV | 396 | 13.0% (9.9-16.9) | 23.2 % (18.5-28.8) | 26.2 (20.7-33.1) |
| HCV | 883 | 9.0% (7.1-11.4) | 21.1% (17.5-25.3) | 21.9 (18.2-26.4) |
| ALD | 228 | 9.9% (5.9-16.1) | 17.7% (10.6-28.9) | 18.4 (11.9-28.5) |
| NAFLD | 111 | 7.7% (3.5-16.6) | 12.8% (6.3-25.1) | 14.4 (7.2-28.8) |
| AIH | 112 | 0% | 1.7% (0.2-11.4) | 1.3 (0.2-9.4) |
| PBC | 108 | 3.4% (1.1-10.3) | 6.1% (2.1-16.9) | 5.5 (2.0-14.6) |
| PSC | 40 | 2.5% (0.4-16.6) | 6.9% (1.7-25.9) | 7.2 (1.8-28.9) |
| Other | 201 | 4.5% (2.1-9.2) | 8.0% (4.2-14.9) | 10.1 (5.6-18.3) |
| **Etiologic category** | | | | |
| Viral | 1279 | 10.3% (8.6-12.3) | 21.7% (18.8-25.0) | 23.3 (20.2-27.0) |
| Steatohepatitis | 339 | 9.1% (5.9-13.8) | 16.3% (10.6-24.5) | 17.0 (11.8-24.7) |
| Autoimmune | 260 | 1.8% (0.7-4.7) | 4.6% (2.0-10.1) | 4.0 (1.9-8.3) |
| Other | 201 | 4.5% (2.1-9.2) | 8.0% (4.214.9) | 10.1 (5.6-18.3) |

24

**Supplementary Table 1: Baseline characteristics of patients with cirrhosis and follow-up >6 months**

| Etiology of cirrhosis | N (total 2079) | Mean Age ±SD [years] | Sex (% male) | Mean follow-up, ±SD [years] | Mean BMI ±SD [Kg/m$^2$] |
|---|---|---|---|---|---|
| HBV | 396 | 52.4±12.6 | 74% | 6.7±3.6 | 24.9±4.5 |
| HCV | 883 | 53.9±10.8 | 63% | 5.7±3.4 | 26.9±5.1 |
| ALD | 228 | 55.3±11.2 | 77% | 4.7±3.5 | 27.7±5.8 |
| NAFLD | 111 | 60.5±10.6 | 51% | 5.0±3.2 | 30.0±6.0 |
| AIH | 112 | 50.7±17.1 | 27% | 6.7±3.3 | 26.9±5.1 |
| PBC | 108 | 57.3±12.4 | 12% | 6.8±3.8 | 24.6±4.2 |
| PSC | 40 | 46.4±17.2 | 62% | 6.9±3.5 | 25.2±4.4 |
| Other | 201 | 53.5±14.6 | 57% | 5.6±3.6 | 25.3±5.5 |
| Viral | 1279 | 53.4±11.4 | 66% | 6.0±3.5 | 26.2±5.0 |
| Steatohepatitis | 339 | 57.0±11.2 | 65% | 4.8±3.4 | 28.6±5.9 |
| Autoimmune | 260 | 52.8±15.8 | 26% | 6.8±3.5 | 25.8±4.8 |

**Supplementary Table 2: C-statistic and 95% CI by subgroup in development and validation cohorts**

| Etiology | C stat | Lower limit | Upper limit |
|---|---|---|---|
| **Derivation Cohort** | | | |
| Autoimmune | 0.73 | 0.71 | 0.78 |
| Steatohepatitis | 0.79 | 0.76 | 0.80 |
| HBV | 0.72 | 0.69 | 0.78 |
| HCV | 0.71 | 0.68 | 0.79 |
| **Validation Cohort** | | | |
| PBC | 0.68 | 0.56 | 0.80 |
| HBV | 0.77 | 0.67 | 0.87 |
| HCV | 0.71 | 0.64 | 0.77 |
| HCV-SVR | 0.73 | 0.63 | 0.83 |

25

**Figure Legends**

**Figure 1: Cumulative incidence of HCC by etiology and disease category**:
The cumulative incidence of HCC is shown for (a) patients with viral hepatitis
(chronic HBV and chronic HCV), (b) patients with steatohepatitis (alcohol or
NASH) and (c) patients with autoimmune liver disease (AIH, PBC and PSC).
The cumulative incidence of HCC is also compared across etiologic categories
(d). Cumulative incidence is compared between groups using the log-rank test.

**Figure 2: Cumulative HCC incidence by Toronto HCC Risk index (THRI)**
The THRI stratified patients into low, intermediate and high risk. Patients with a
THRI score below 120 had an annualized HCC incidence of 0.3% per year. The
annualized incidence in patients with intermediate risk (120-240 points) was
1.0% whereas for patients with high risk (>240 points) was 3.2%.  The curves for
the 3 risk groups were compared using the log-rank test.

**Figure 3: Nomogram for 5- and 10-year HCC incidence based on THRI
point-score.**  The 5 and 10-year HCC incidence was calculated for each point
score allowing for individualized risk assessment. The cumulative HCC incidence
for a given THRI score equals $1 - S_0(t)^{exp(\beta*THRI)}$ , where $S_0$ is the estimated
baseline HCC-free survival at time t and $\beta=0.013227$, the estimated effect of
each additional THRI point.

**Figure 4: Observed vs. Predicted HCC incidence in the validation cohort:**
Observed and expected cumulative Incidence of HCC in the validation dataset for the
three risk groups: >240 (red), 120-240 (green) and <120 (blue). Jagged lines (with dots)
observed HCC incidence obtained by Kalpan-Meier method. Dark colored lines:
predicted HCC incidence obtained from the THRI and baseline survival function
estimated from the derivation dataset.

**Supplementary Figure 1: Patient flow**: 12,199 patients were seen in the clinic
during the study period. Using APRI and FIB-4 cutoffs, 3,064 patients had
probable cirrhosis.  Charts of all of these patients were reviewed and 2,416
(79.5%) had clinical, radiological or histological evidence of cirrhosis.  Of these,
2,079 patients were followed for >6 month and constituted the at risk cohort, of
whom 226 developed HCC during follow-up.

**Supplementary Figure 2:** Calibration plot contrasting predicted vs. observed 10-
year risk (probability) of HCC in the internal validation cohort

## References

[1]     Ferenci P, Fried M, Labrecque D, Bruix J, Sherman M, Omata M, et al. Hepatocellular Carcinoma (HCC). Journal of Clinical Gastroenterology. 2010;44:239–245.

[2]     Schuumltte K, Bornschein J, Malfertheiner P. Hepatocellular Carcinoma &ndash; Epidemiological Trends and Risk Factors. Dig Dis. 2009;27:80–92.

[3]     Torre LA, Bray F, Siegel RL, Ferlay J. Global cancer statistics, 2012. CA Cancer J Clin. 2015 Mar;65(2):87-108

[4]     Altekruse SF, McGlynn KA, Reichman ME. Hepatocellular Carcinoma Incidence, Mortality, and Survival Trends in the United States From 1975 to 2005. Journal of Clinical Oncology. 2009;27:1485–1491.

[5]     Forner A, MD AF, MD JML, Llovet JM, MD DJB, Bruix J. Hepatocellular carcinoma. The Lancet. 2012;379:1245–1255.

[6]     Bruix J, Sherman M. Management of hepatocellular carcinoma: An update. Hepatology. 2011;53:1020–1022.

[7]     EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma. J Hepatol. 2012 Apr;56(4):908-43

[8]     Sherman M, Bruix J, Porayko M, Tran T, for the AASLD Practice Guidelines Committee. Screening for hepatocellular carcinoma: The rationale for the American Association for the Study of Liver Diseases recommendations. Hepatology. 2012;56:793–796.

[9]     Tezuka M, Hayashi K, Kubota K, Sekine S, Okada Y, Ina H, et al. Growth Rate of Locally Recurrent Hepatocellular Carcinoma After Transcatheter Arterial Chemoembolization: Comparing the Growth Rate of Locally Recurrent Tumor with That of Primary Hepatocellular Carcinoma. Dig Dis Sci. 2007;52:783–788.

[10]    Kubota K, Ina H, Okada Y, Irie T. Growth rate of primary single hepatocellular carcinoma: determining optimal screening interval with contrast enhanced computed tomography. Dig Dis Sci. 2003;48:581–586.

[11]    Yuen MF. Screening for hepatocellular carcinoma: survival benefit and cost-effectiveness. Annals of Oncology. 2003;14:1463–1467.

[12]    Zhang B-H, Yang B-H, Tang Z-Y. Randomized controlled trial of screening for hepatocellular carcinoma. J Cancer Res Clin Oncol. 2004;130.

27

[13]     Ascha MS, Hanouneh IA, Lopez R, Tamimi TA-R, Feldstein AF, Zein NN. The incidence and risk factors of hepatocellular carcinoma in patients with nonalcoholic steatohepatitis. Hepatology. 2010;51:1972–1978.

[14]     Fassio E. Hepatitis C and hepatocellular carcinoma. Ann Hepatol. 2010;9 Suppl:119–122.

[15]     Caballería L, Caballeria L, Parés A, Pares A, Castells A, Castells A, et al. Hepatocellular carcinoma in primary biliary cirrhosis: similar incidence to that in hepatitis C virus-related cirrhosis. Am J Gastroenterology. 2001;96:1160–1163.

[16]     Ohira H, Abe K, Takahashi A, Zeniya M, Ichida T. Clinical features of hepatocellular carcinoma in patients with autoimmune hepatitis in Japan. J Gastroenterol. 2012;48:109–114.

[17]     Wai C, Wai C-T, Greenson JK, Fontana RJ, Kalbfleisch JD, Marrero JA, et al. A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. Hepatology. 2003;38:518–526.

[18]     Pichard AV, Mallet V, Nalpas B, Verkarre V. FIB-4: An inexpensive and accurate marker of fibrosis in HCV infection. comparison with liver biopsy and fibrotest. Hepatology. 2007 Jul;46(1):32-6.

[19]     Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med. 1996 Feb 28;15(4):361-87.

[20]     Wong VWS, Chan SL, Mo F, Chan TC, Loong HHF, Wong GLH, et al. Clinical Scoring System to Predict Hepatocellular Carcinoma in Chronic Hepatitis B Carriers. Journal of Clinical Oncology. 2010;28:1660–1665.

[21]     Uno H, Cai T, Pencina MJ, D'Agostino RB, Wei LJ. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. Stat Med. 2011 May 10;30(10):1105-17

[22]     Harrell FE, Jr, Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the Yield of Medical Tests. JAMA. 1982; 247(18):2543-2546

[23]     Royston P, Altman DG. External validation of a Cox prognostic model: principles and methods. BMC Med Res Methodol. 2013 Mar 6;13:33

[24]     Lammers WJ, van Buuren HR, Hirschfield GM. Levels of alkaline phosphatase and bilirubin are surrogate end points of outcomes of patients with primary biliary cirrhosis: an international follow-up study. Gastroenterology. 2014 Dec;147(6):1338-49

28

[25] Arends P, Sonneveld MJ, Zoutendijk R, Carey I. Entecavir treatment does not eliminate the risk of hepatocellular carcinoma in chronic hepatitis B: limited role for risk scores in Caucasians. Gut. 2015 Aug;64(8):1289-95.

[26] Brouwer WP, Meer AJ, Boonstra A, Pas SD. The impact of PNPLA3 (rs738409 C> G) polymorphisms on liver histology and long-term clinical outcome in chronic hepatitis B patients. Liver Int. 2015 Feb;35(2):438-47.

[27] van der Meer AJ, Veldt BJ, Feld JJ, Wedemeyer H. Association between sustained virological response and all-cause mortality among patients with chronic hepatitis C and advanced hepatic fibrosis. JAMA. 2012 Dec 26;308(24):2584-93.

[28] Yang HI, Sherman M, Su J, Chen PJ, Liaw YF, Iloeje UH, et al. Nomograms for Risk of Hepatocellular Carcinoma in Patients With Chronic Hepatitis B Virus Infection. Journal of Clinical Oncology. 2010;28:2437–2444.

[29] Ikeda K, Saitoh S, Koida I, Arase Y, Tsubota A, Chayama K, et al. A multivariate analysis of risk factors for hepatocellular carcinogenesis: a prospective observation of 795 patients with viral and alcoholic cirrhosis. Hepatology1993;18:47–53.

[30] Lok AS, Seeff LB, Morgan TR, Di Bisceglie AM, Sterling RK, Curto TM, et al. Incidence of Hepatocellular Carcinoma and Associated Risk Factors in Hepatitis C-Related Advanced Liver Disease. Gastroenterology. 2009;136:138–148.

[31] Sangiovanni A, Prati GM, Fasani P, Ronchi G, Romeo R, Manini M, et al. The natural history of compensated cirrhosis due to hepatitis C virus: A 17-year cohort study of 214 patients. Hepatology. 2006;43:1303–1310.

[32] Wang C, Chen V, Vu V, Le A, Nguyen L, Zhao C. Poor adherence and low persistency rates for hepatocellular carcinoma surveillance in patients with chronic hepatitis B. Medicine (Baltimore). 2016 Aug;95(35):e4744
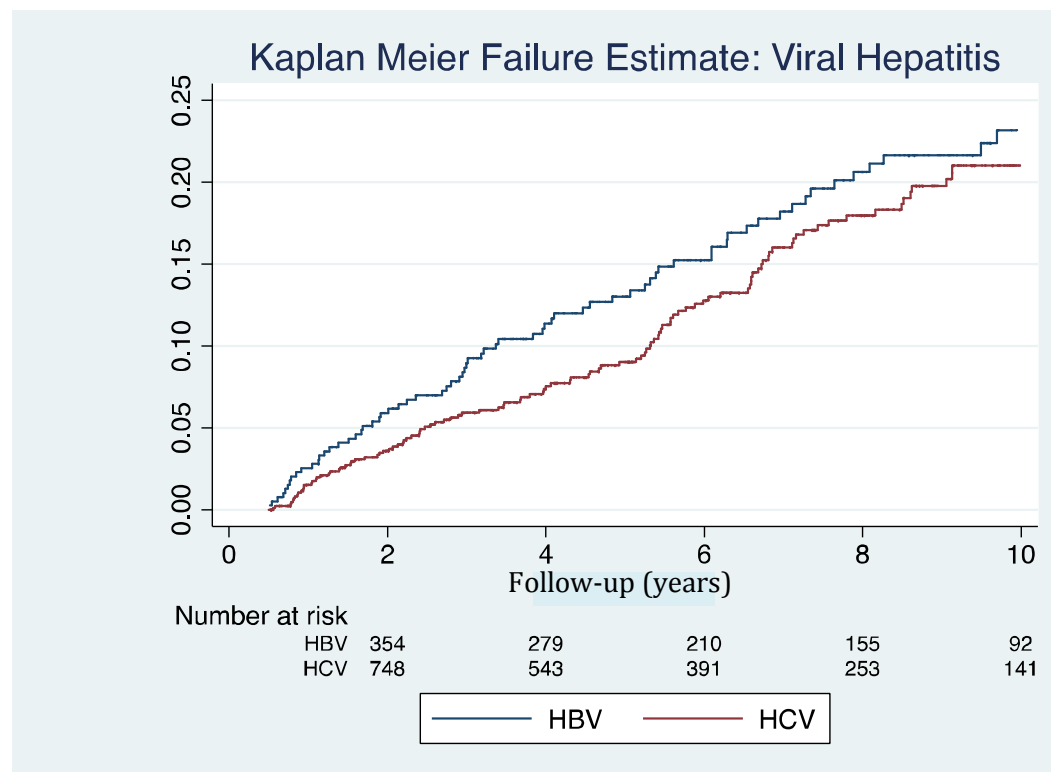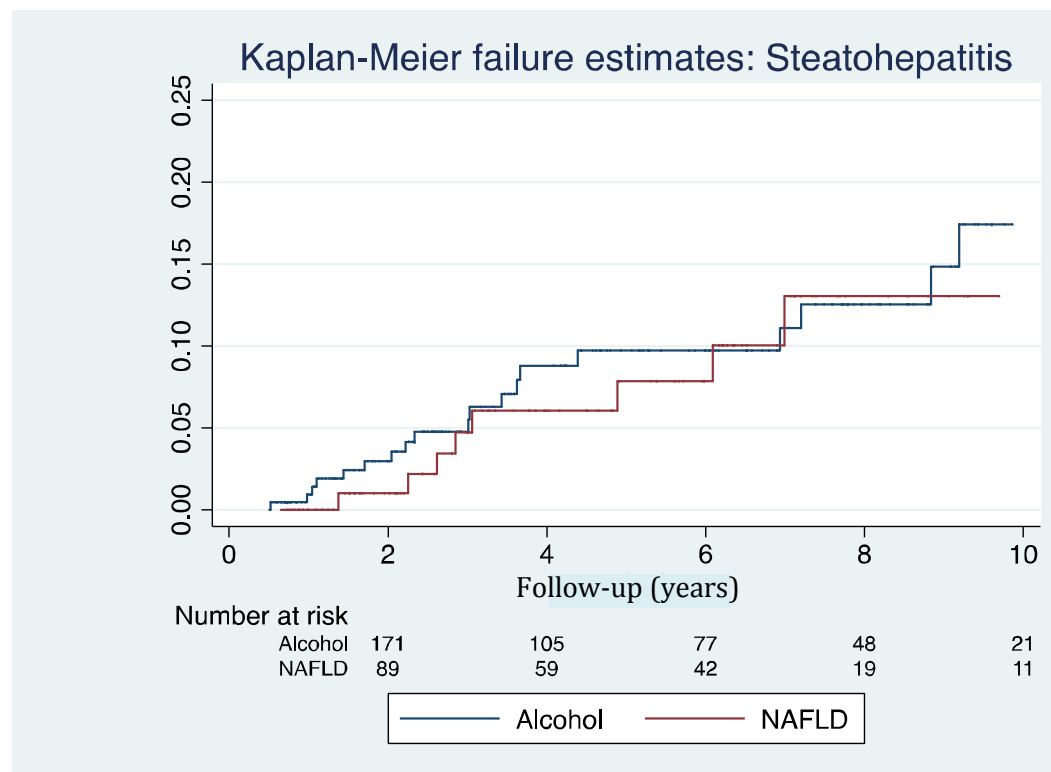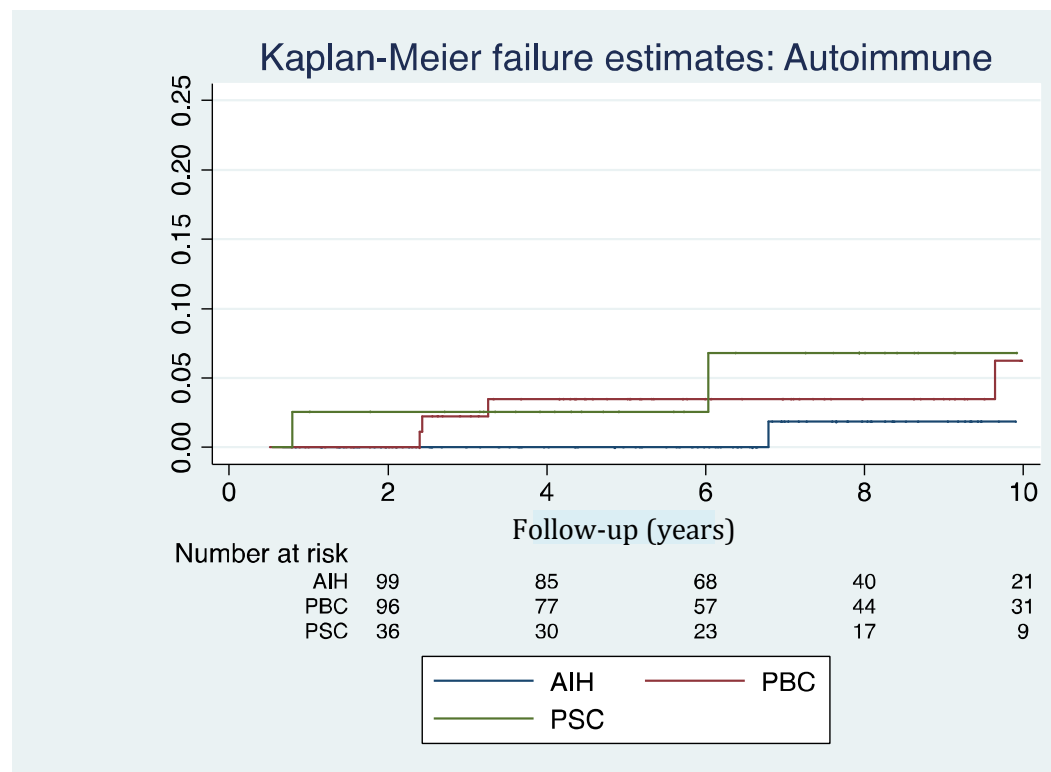
**Figure 1a**



Kaplan Meier Failure Estimate: Viral Hepatitis

Number at risk
| | | | | | |
|---|---|---|---|---|---|
| HBV | 354 | 279 | 210 | 155 | 92 |
| HCV | 748 | 543 | 391 | 253 | 141 |

**Figure 1b**



Kaplan-Meier failure estimates: Steatohepatitis

| Number at risk | | | | | |
|---|---|---|---|---|---|
| Alcohol | 171 | 105 | 77 | 48 | 21 |
| NAFLD | 89 | 59 | 42 | 19 | 11 |

**Figure 1c**



Kaplan-Meier failure estimates: Autoimmune

**Figure 1d**



Kaplan-Meier failure estimates by etiologic category
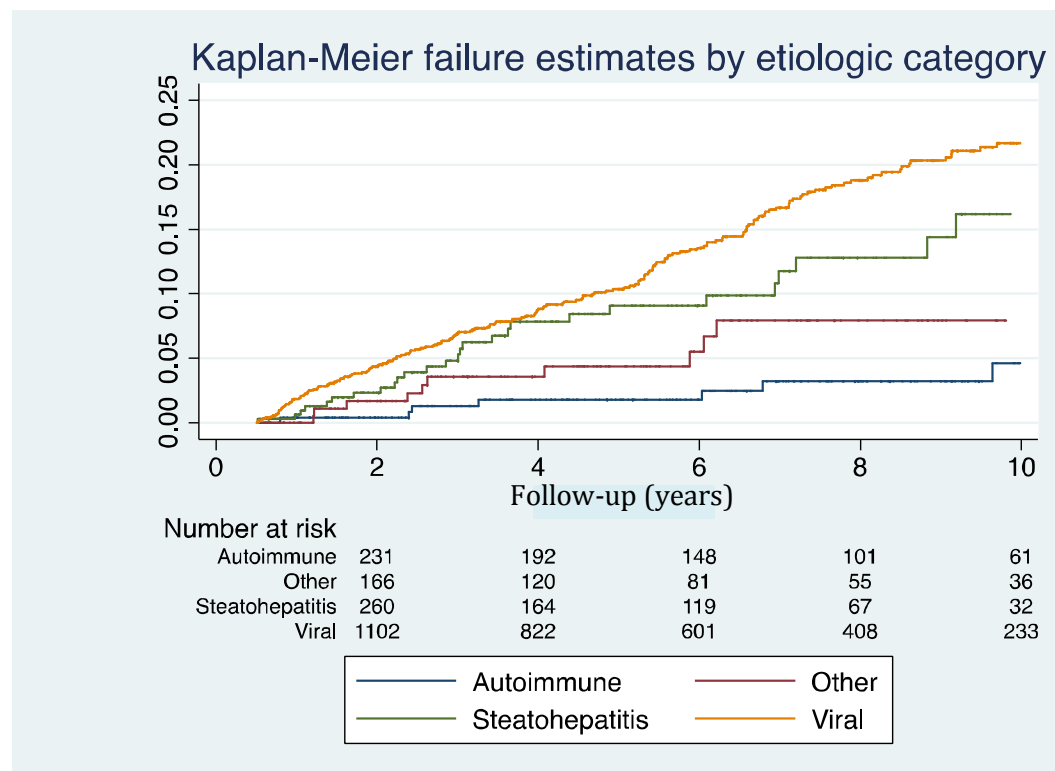
Graphical abstract



Kaplan-Meier failure estimates by etiologic category
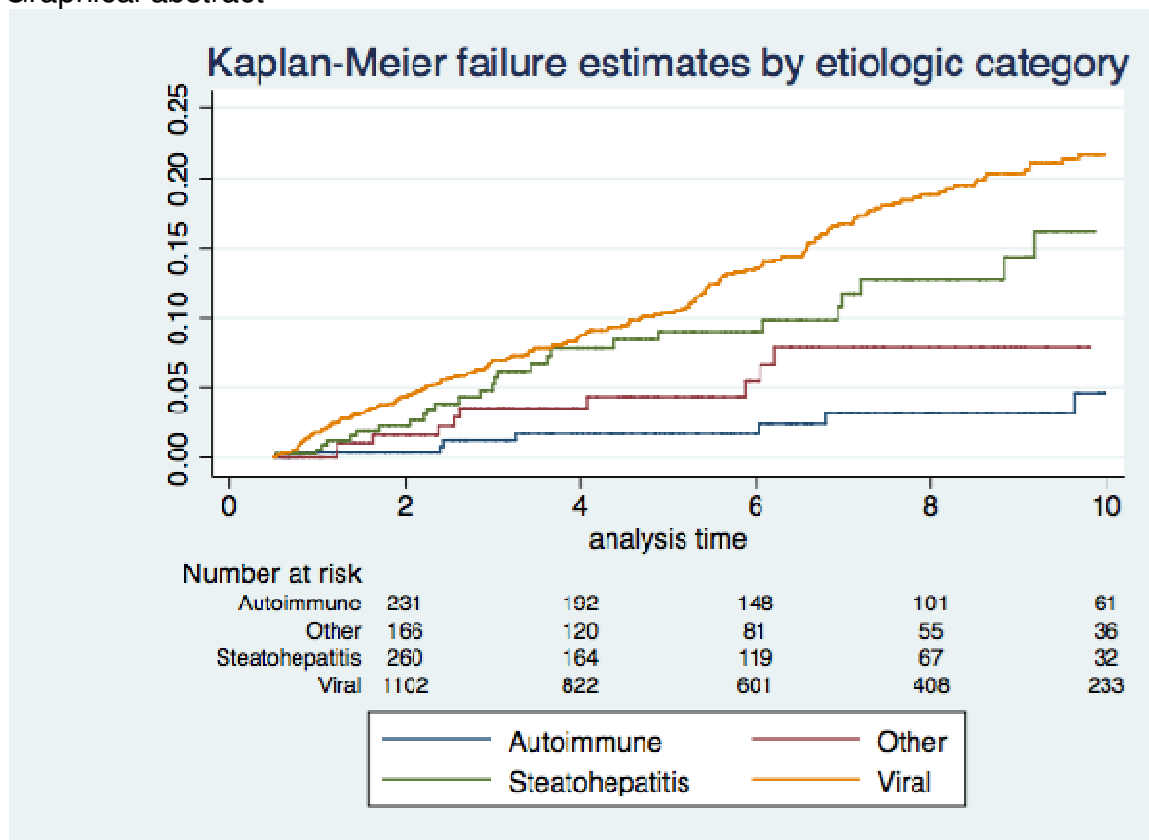
Highlights

- HCC incidence varies markedly by etiology of cirrhosis.

- THRI is simple to use, has good predictive ability, and has been externally validated

- THRI may help to refine  HCC surveillance guidelines for cirrhotic patients

31

**Figures**

**Figure 1: Cumulative incidence of HCC by etiology and disease category**:
The cumulative incidence of HCC is shown for (a) patients with viral hepatitis
(chronic HBV and chronic HCV), (b) patients with steatohepatitis (alcohol or
NASH) and (c) patients with autoimmune liver disease (AIH, PBC and PSC).
The cumulative incidence of HCC is also compared across etiologic categories
(d). Cumulative incidence is compared between groups using the log-rank test.
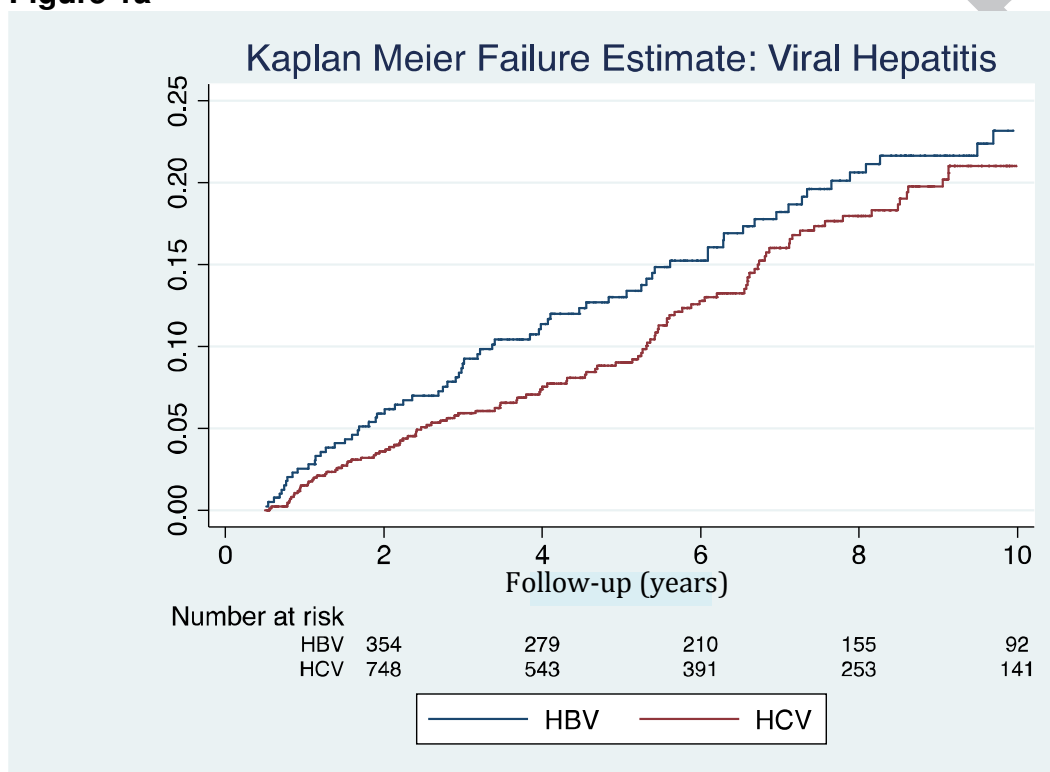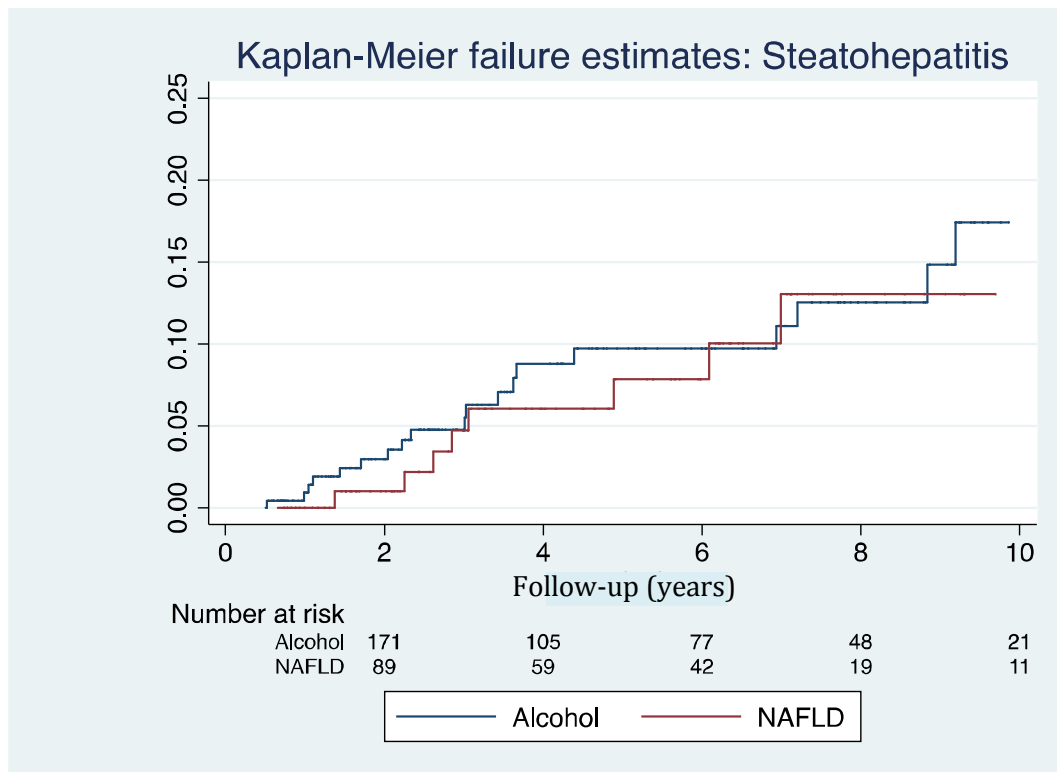
**Figure 1a**



Kaplan Meier Failure Estimate: Viral Hepatitis

Number at risk

| | | | | | |
|---|---|---|---|---|---|
| HBV | 354 | 279 | 210 | 155 | 92 |
| HCV | 748 | 543 | 391 | 253 | 141 |

HBV ——— HCV

**Figure 1b**



Kaplan-Meier failure estimates: Steatohepatitis

Number at risk

| | | | | | |
|---|---|---|---|---|---|
| Alcohol | 171 | 105 | 77 | 48 | 21 |
| NAFLD | 89 | 59 | 42 | 19 | 11 |

Alcohol — NAFLD

**Figure 1c**



Kaplan-Meier failure estimates: Autoimmune

Number at risk

| | | | | | |
|---|---|---|---|---|---|
| AIH | 99 | 85 | 68 | 40 | 21 |
| PBC | 96 | 77 | 57 | 44 | 31 |
| PSC | 36 | 30 | 23 | 17 | 9 |

AIH — PBC
PSC

**Figure 1d**



Kaplan-Meier failure estimates by etiologic category

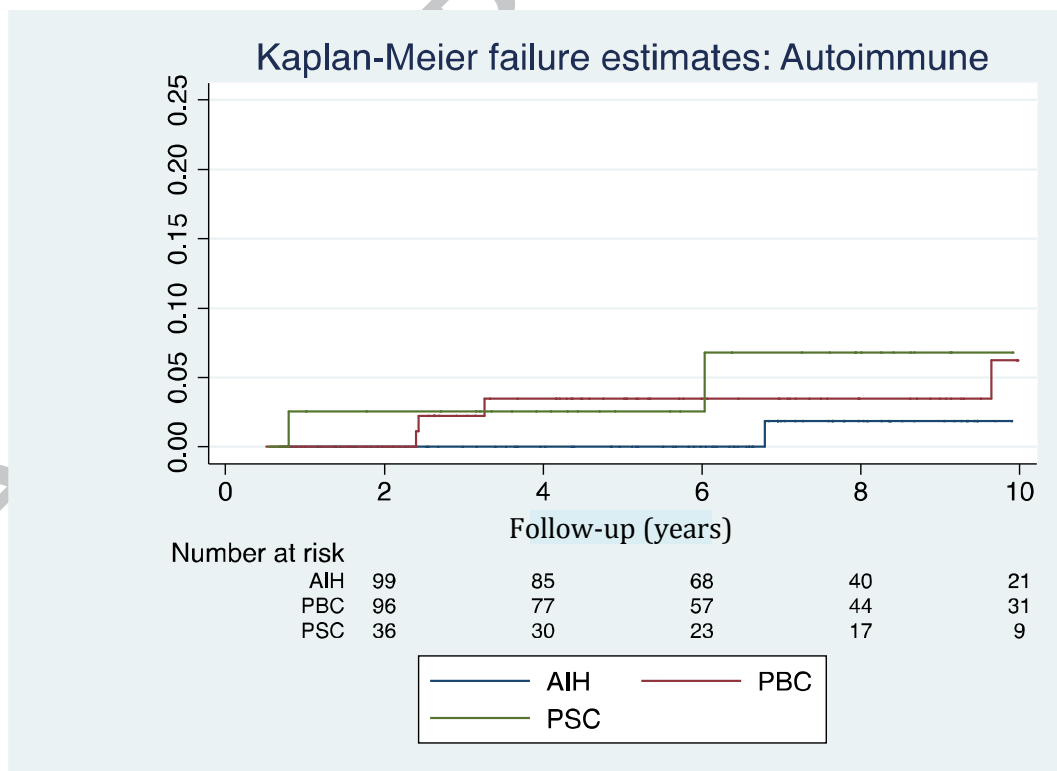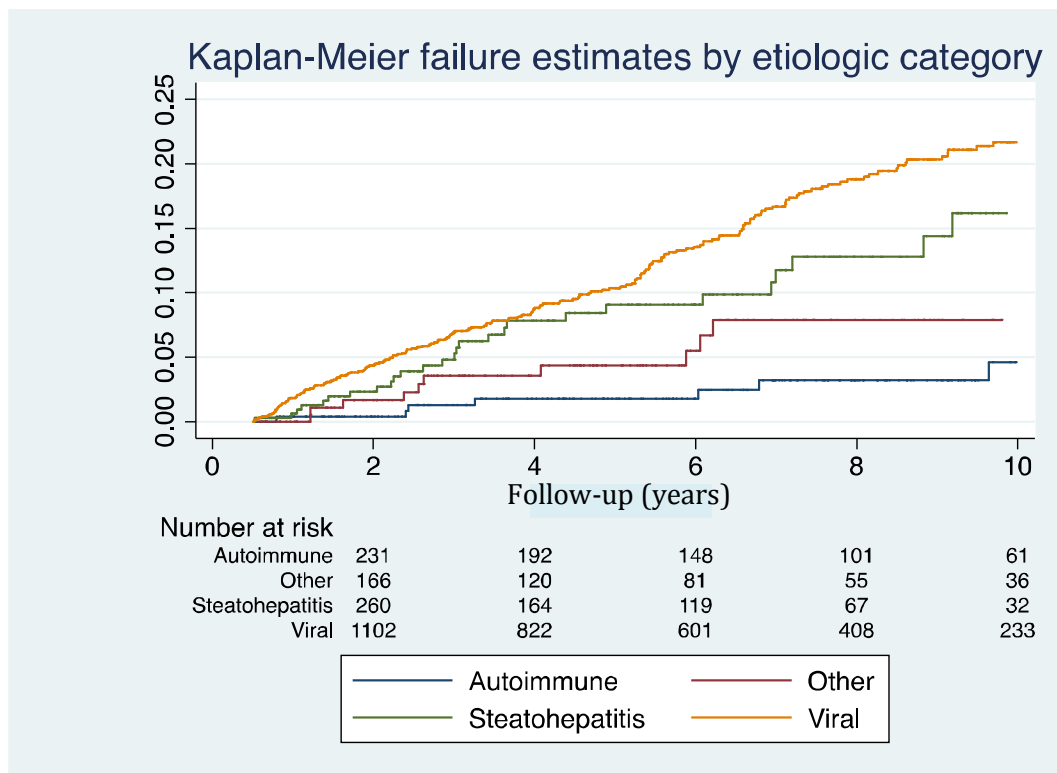| Number at risk | | | | | |
|---|---|---|---|---|---|
| Autoimmune | 231 | 192 | 148 | 101 | 61 |
| Other | 166 | 120 | 81 | 55 | 36 |
| Steatohepatitis | 260 | 164 | 119 | 67 | 32 |
| Viral | 1102 | 822 | 601 | 408 | 233 |

**Figure 2: Cumulative HCC incidence by Toronto HCC Risk index (THRI)**
The THRI stratified patients into low, intermediate and high risk. Patients with a THRI score below 120 had an annualized HCC incidence of 0.3% per year. The annualized incidence in patients with intermediate risk (120-240 points) was 1.0% whereas for patients with high risk (>240 points) was 3.2%. The curves for the 3 risk groups were compared using the log-rank test.
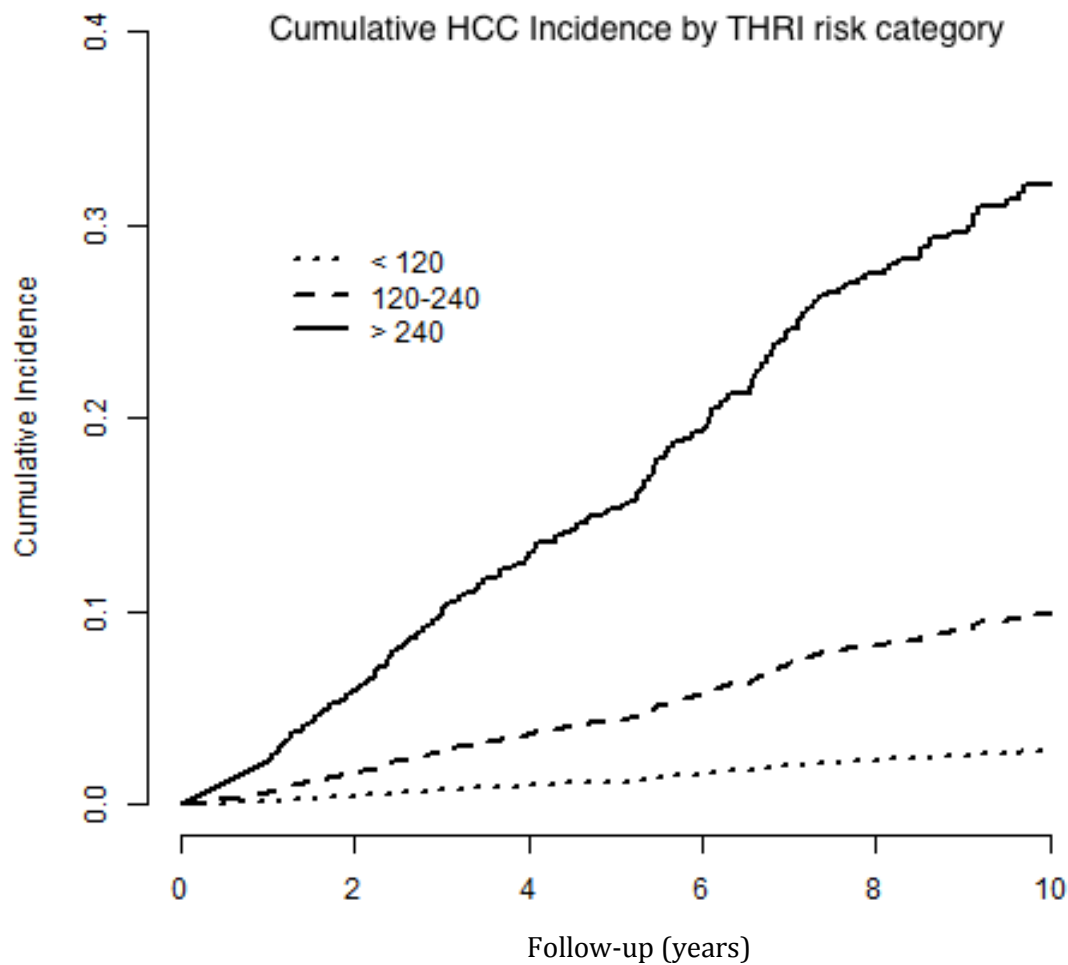
Figure 2



**Cumulative HCC Incidence by THRI risk category**

Legend:
· · · < 120
− − 120-240
— > 240

Y-axis: Cumulative Incidence (0.0, 0.1, 0.2, 0.3, 0.4)
X-axis: Follow-up (years) (0, 2, 4, 6, 8, 10)

*Figure 3***: Nomogram for 5- and 10-year HCC incidence based on THRI point-score.** The 5 and 10-year HCC incidence was calculated for each point score allowing for individualized risk assessment. The cumulative HCC incidence for a given THRI score equals $1 - S_0(t)^{exp(\beta*THRI)}$, where $S_0$ is the estimated baseline HCC-free survival at time t and $\beta=0.013227$, the estimated effect of each additional THRI point.
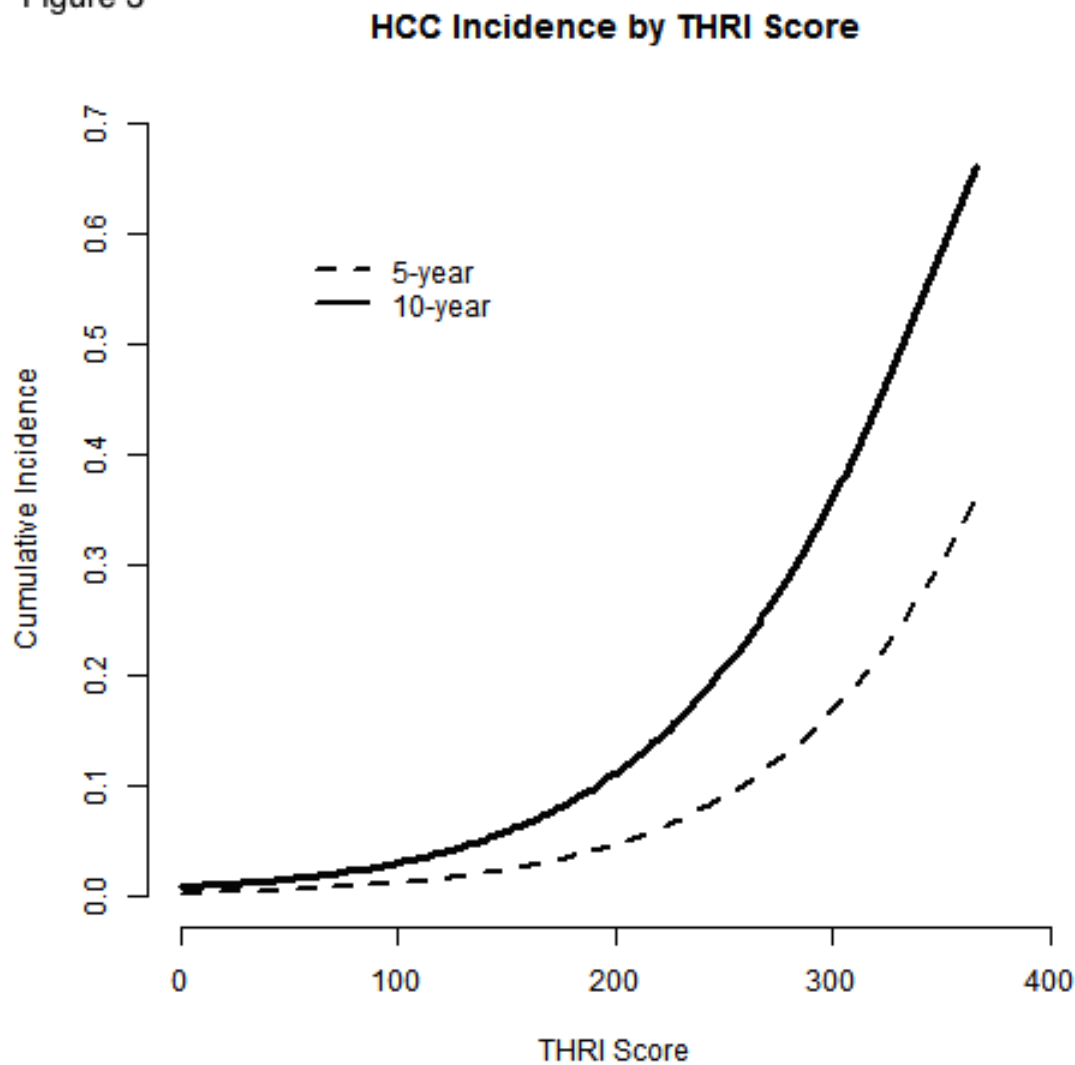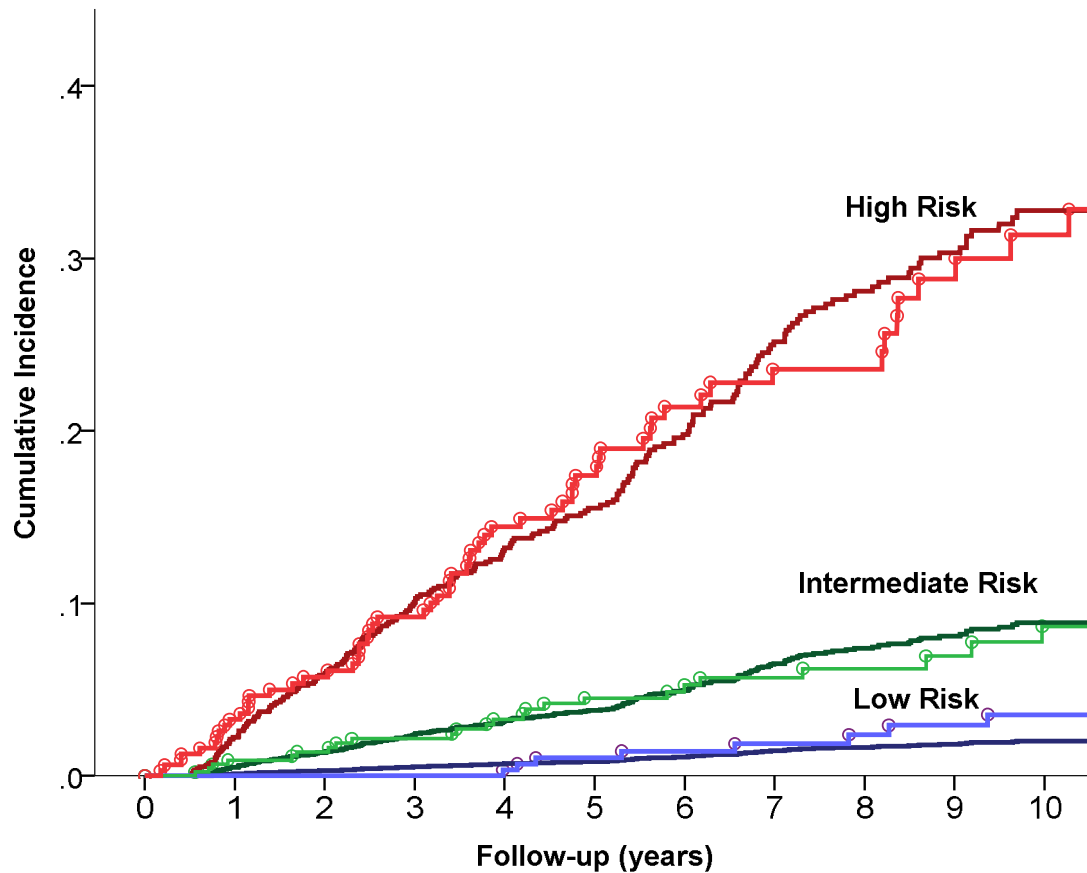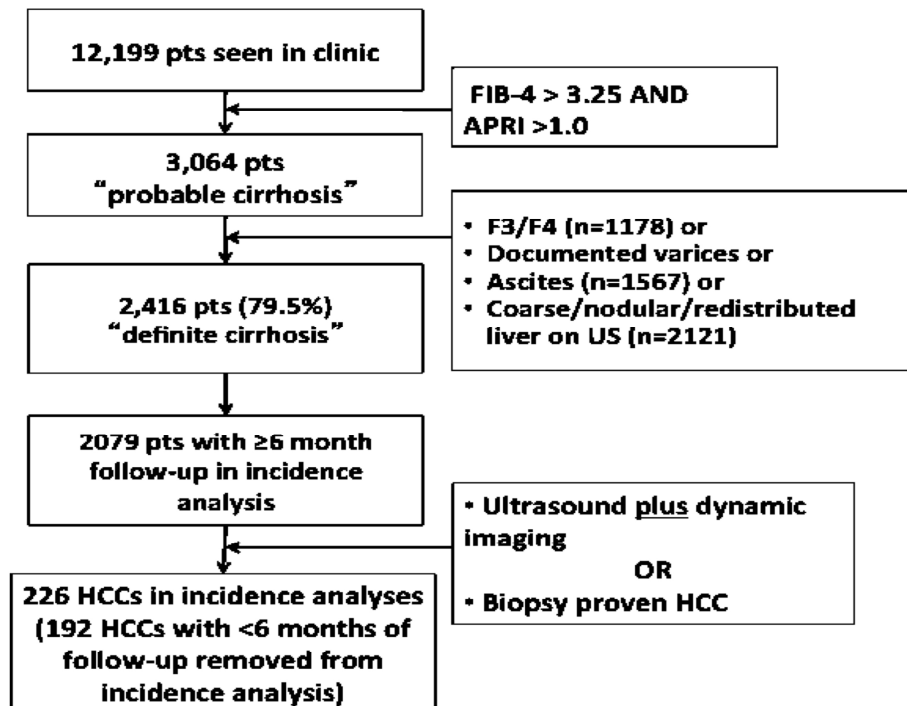
Figure 3



**Figure 4: Observed vs. Predicted HCC incidence in the validation cohort:**
Observed and expected cumulative Incidence of HCC in the validation dataset for the three risk groups: >240 (red), 120-240 (green) and <120 (blue). Jagged lines (with dots) observed HCC incidence obtained by Kalpan-Meier method. Dark colored lines: predicted HCC incidence obtained from the THRI and baseline survival function estimated from the derivation dataset.

**Supplementary Figure 1: Patient flow**: 12,199 patients were seen in the clinic during the study period. Using APRI and FIB-4 cutoffs, 3,064 patients had probable cirrhosis. Charts of all of these patients were reviewed and 2,416 (79.5%) had clinical, radiological or histological evidence of cirrhosis. Of these, 2,079 patients were followed for >6 month and constituted the at risk cohort, of whom 226 developed HCC during follow-up.

## Supplementary Figure 1



**Supplementary Figure 2:** Calibration plot contrasting predicted vs. observed 10 year risk (probability) of HCC in the internal validation cohort