



# Hepatocellular Carcinoma Risk Scores Predict Patients Under Surveillance at Low Risk of Benefit and High Risk of Harm

Chris Curran<sup>1,4</sup> · Matthew Priest<sup>1</sup> · Shouren Datta<sup>1</sup> · Ewan H. Forrest<sup>2,3</sup> · Adrian J. Stanley<sup>2,3</sup> · Stephen T. Barclay<sup>2</sup>

Received: 14 April 2022 / Accepted: 11 October 2022 / Published online: 14 November 2022  
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

## Abstract

**Aims** Surveillance for hepatocellular carcinoma (HCC) is recommended for patients with cirrhosis. Multiple risk scores aim to stratify HCC risk, potentially allowing individualized surveillance strategies. We sought to validate four risk scores and quantify the consequences of surveillance via the calculation of numbers needed to benefit (NNB) and harm (NNH) according to classification by risk score strata.

**Methods** Data were collected on 482 patients with cirrhosis during 2013–2014, with follow-up until 31/12/2019. Risk scores (aMAP, Toronto risk index, ADDRESS HCC, HCC risk score) were derived from index clinic results. The area under the receiving operating characteristic curve (AUC) was calculated for each. Additionally, per-risk strata, NNB was calculated as total surveillance ultrasounds per surveillance diagnosed early HCC (stage 0/A) and NNH as total ultrasounds performed per false positive (abnormal surveillance with normal follow-up imaging).

**Results** 22 (4.6%) patients developed HCC. 77% (17/22) were diagnosed through surveillance, of which 13/17 (76%) were early stage. There were 88 false positives and no false negatives (normal surveillance result however subsequent HCC detection). Overall NNB and NNH were 241 and 36, respectively. No score was significantly superior using AUC. Patients classified as low risk demonstrated no surveillance benefit (AMAP, THRI) or had a high NNB of > 300/900 (ADDRESS HCC, HCC risk score), with low NNH (24–38).

**Conclusion** Given the lack of benefit and increased harm through false positives in low-risk groups, a risk-based surveillance strategy may have the potential to reduce patient harm and increase benefit from HCC surveillance.

**Clinical Trials Registration** This was not a clinical trial and the study was not pre-registered.

**Keywords** Screening · Cancer · Cirrhosis · Scotland · Positive predictive value · Negative predictive value · Ultrasound · Alpha-fetoprotein · Strategy · Liver disease

## Introduction

International guidelines suggest surveillance for hepatocellular carcinoma (HCC) in patients with cirrhosis (EASL/ AASLD), is likely to be cost effective where annual risk meets or exceeds 1.5% per year [1, 2]. Despite significant observational evidence associating surveillance with receipt of curative therapy and overall survival [3–6] uncertainty persists about overall benefit given the lack of confirmatory randomized control trial evidence [7]. Patients under surveillance may experience harm from unnecessary investigations triggered by false-positive surveillance ultrasounds or raised alpha fetoprotein (AFP) levels, with recent American research suggesting over a quarter of patients are adversely affected [8]. False-positive surveillance tests may lead to anxiety while awaiting

An editorial commenting on this article is available at <https://doi.org/10.1007/s10620-022-07735-x>

✉ Chris Curran  
christopher.curran@nhs.scot

<sup>1</sup> Department of Gastroenterology, Queen Elizabeth University Hospital, 1345 Govan Rd, Glasgow G51 4TF, UK

<sup>2</sup> Department of Gastroenterology, Glasgow Royal Infirmary, 84 Castle St, Glasgow G4 0SF, UK

<sup>3</sup> College of Medical, Veterinary & Life Sciences, University of Glasgow, Wolfson Medical School Building, University Avenue, Glasgow G12 8QQ, UK

<sup>4</sup> Flat 2/2, 53 Dalnair Street, Glasgow G3 8SQ, UK

follow-up investigations and results and on rare occasions harm related to biopsy of lesions [9, 10]. Additionally patients enrolled in surveillance may not benefit due to high competing mortality from liver disease, cardiovascular disease, and non-hepatic malignancy [11]. Surveillance programs consume significant healthcare resource; however, data on their impact on quality of life is lacking [12, 13].

In this context, multiple risk scores have been created to stratify HCC risk, with the aim of allowing more individualized surveillance strategies [14–18]. European guidelines suggest it is possible to identify low-risk patients and exclude them from surveillance although this is unproven outside of research using simulated cohorts [1, 19].

To date, evaluation of scoring systems has focused on their ability to classify patients into either low/high or low/medium/high-risk groups for development of HCC. However, the aim of an HCC surveillance program should be to maximize benefit, through detection of early cancers amenable to curative treatment, while minimizing harm caused by unnecessary investigation of false-positive results. In this context we sought to examine the number needed to screen to experience HCC surveillance benefit and harm, according to risk stratification by previously published scores, in a cohort of patients with cirrhosis of mixed etiology. In addition, we sought to examine the positive and negative predictive values of screening tests within different risk strata.

## Methods

### Design, Setting, and Participants

This was a retrospective analysis of patients identified from a prospective database of patients with cirrhosis. Patients were prospectively enrolled in the database from October 2010 onwards, to aid coordination of care, including HCC surveillance, across hepatology clinics in Glasgow. During this period, standard practice was to offer HCC surveillance using 6-monthly ultrasound and measurement of AFP levels. The database captured data, including demographics, clinical data, etiology of liver disease, and laboratory results (Supplementary Table 1). Patients with a diagnosis of cirrhosis who attended at least one outpatient appointment for HCC surveillance between 1/1/2013 and 31/12/2014 were included. This allowed for a minimum of 5 years of follow-up, with patients followed until detection of HCC, transplant, death, or until 31/12/2019. Caldicott approval was sought and granted to ensure appropriate handling of patient data. Additional ethical permission for this retrospective analysis of existing data was not required.

## Variables

Data required to calculate the risk scores were extracted from the database, with missing data obtained via electronic health record review. Variables collected included age, gender, body mass index (BMI), etiology of cirrhosis, presence of diabetes, and laboratory tests. Data on outpatient radiology visits conducted throughout the study period including type of imaging and date attended were obtained from the radiology service. To account for the fact that patients may have multiple contributing etiologies a dichotomous choice was used for a range of etiologies, including, alcohol, nonalcoholic fatty liver disease (NAFLD), Hepatitis C (HCV), Hepatitis B (HBV), hemochromatosis, autoimmune, and other. Diagnosis of HCC was made according to international guidelines following review at the West of Scotland HCC Manged Care Network (MCN) multidisciplinary team meeting (MDT) [1].

Risk scores chosen were aMAP, Toronto HCC risk index, ADDRESS HCC, and the HCC risk score [14–16, 18, 20]. These scores were chosen due to demonstrating good predictive ability in large derivation cohorts of patients with cirrhosis of mixed etiology. Additionally, they are calculated using basic clinical and laboratory values from a single timepoint making them simple to calculate (Supplementary Table 2). Low-, medium-, and high-risk strata were defined as per the values set out in the original papers. For the HCC risk score, only patients with a single etiological factor had a score generated. Risk score values were derived from data at the index clinic visit or if missing, from the closest appointment within a 6-month period before or after. For each score, a receiver operated curve (ROC) was drafted, and the area under the ROC (AUROC) calculated. For each strata of risk score the proportion of the cohort selected was recorded and the negative and positive predictive values for the detection of HCC throughout the study period for values below and above this threshold within the cohort were calculated. Freedom from detection of HCC was calculated for the low-risk, intermediate-risk, and high-risk groups using the Kaplan–Meier method.

Any ultrasound or AFP result which that triggered cross-sectional imaging at the treating physician's discretion was considered a positive surveillance test. Surveillance tests not triggering cross sectional imaging were considered negatives. Positive and negative tests were evaluated as true or false according to diagnosis of HCC within 6 months of the tests at the MDT, and positive and negative predictive values of developing HCC following an abnormal surveillance result was calculated according to risk score strata.

The definitions of benefit and harm were defined in line with previous research [8, 9]. Benefit was specified

as detection of an HCC by surveillance at a potentially curative stage, i.e., BCLC stage 0 or A [1]. Harm was defined as cross-sectional imaging triggered by a positive surveillance test (ultrasound or AFP) without subsequent detection of an HCC. Cross-sectional imaging that was triggered by clinical presentation rather than surveillance findings was excluded from calculation of surveillance benefits or harms. The number of surveillance episodes was calculated according to the number of ultrasounds attended. These data were used to calculate the number of surveillance episodes needed to benefit (NNB) or harm (NNH) according to score risk strata. Compliance with surveillance was calculated as the number of ultrasounds performed over the expected number of ultrasounds (1 per 6 months of follow-up).

## Statistical Methods

Data were collected on to a Microsoft Access database and statistical analysis and graph creation was performed using R studio. Continuous data were reported as the median  $\pm$  interquartile range (IQR). Due to a low percentage of missing values these were handled by median imputation (Supplementary Table 3). The Kaplan–Meier method was used for survival analysis.

## Results

### Participants and Descriptive Data

506 patients were screened for inclusion into the study. 24 patients were excluded for reasons, including not having a clinic appointment within the study period ( $n=16$ ), patient death prior to the study period ( $n=2$ ), developing HCC prior to the study ( $n=2$ ), incorrect demographic details ( $n=2$ ), and other ( $n=2$ ), leaving 482 patients identified for analysis (Fig. 1). Median follow-up was 1930 days (IQR 1182, 2324). 98% (473/482) of patients attended for 3137 surveillance ultrasounds (mean 6.6 ( $\pm 4.2$ ) per patient). Compliance with surveillance was 68%.

Descriptive statistics and baseline biochemistry values are given in Table 1. Alcohol was the most common etiology of liver disease (262 (54.4%)). Patients were predominantly male (318 (66%)) and overweight (Median BMI 29.0 (25.5, 32.7)) with 1 in 4 having a diagnosis of diabetes.

### Outcome Data

HCC was detected in 22 (4.6%) patients. 163 (33.8%) died without a diagnosis of HCC and 17 (3.5%) underwent transplantation for non-HCC indications. Of the 22 patients in whom HCC was diagnosed, 77% (17/22) were detected

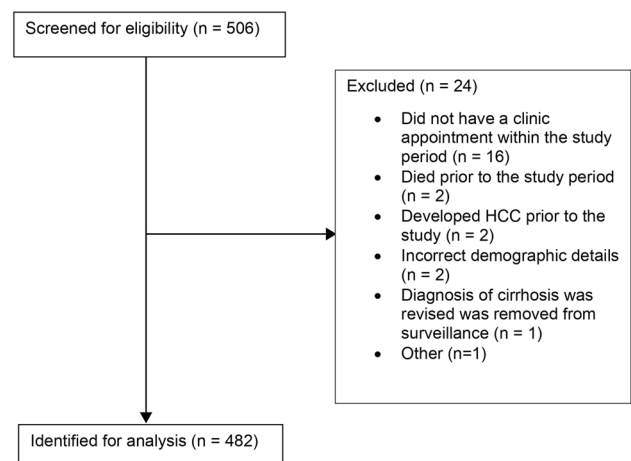


Fig. 1

Table 1 Demographic and baseline values

| Characteristics                      | (n = 482)           |
|--------------------------------------|---------------------|
| Age (median $\pm$ IQR)               | 55.5 (46.7, 63.9)   |
| BMI (median $\pm$ IQR)               | 29.0 (25.5, 32.7)   |
| Male sex—(%)                         | 318 (66.0)          |
| Diabetes—(%)                         | 122 (25.3)          |
| Etiology—n (%) *                     |                     |
| Alcohol                              | 262 (54.4%)         |
| NAFLD                                | 93 (19.3%)          |
| HCV                                  | 143 (29.7%)         |
| HBV                                  | 8 (1.7%)            |
| AIH                                  | 11 (2.3%)           |
| Hemachromatosis                      | 19 (3.9%)           |
| Other                                | 14 (2.9%)           |
| Laboratory values (median $\pm$ IQR) |                     |
| Bilirubin ( $\mu\text{mol/L}$ )      | 16.0 (10.0, 26.8)   |
| Albumin (g/L)                        | 36.0 (32.0, 39.0)   |
| Platelets ( $\times 10^9/\text{L}$ ) | 139.0 (96.0, 203.0) |
| Prothrombin time (s)                 | 12.0 (11.0, 13.0)   |

IQR interquartile range, NAFLD nonalcoholic fatty liver disease, HCV hepatitis C, HBV hepatitis B, AIH autoimmune hepatitis

\*To account for the fact that patients may have multiple contributing etiologies a dichotomous choice for each etiological factor was used

through surveillance, of which 13/17 (76%) had BCLC stage 0 or A at diagnosis. At 1, 3, and 5 years the cumulative incidence of HCC in this cohort was 0.21%, 1.24%, and 3.94%, respectively, and the cumulative risk of non-HCC mortality was 7.5%, 19.9%, and 29.1%. HCC risk rose across risk strata for all scores, as did non-HCC mortality (Table 2). Cumulative risk of HCC was highest in the high-risk strata of the HCC risk score at 8.6%, with this group also demonstrating high non-HCC mortality of 63%.

**Table 2** Cumulative risk of HCC and non-HCC mortality

| Score          | Risk Strata | No of pts (%) | Total HCC | Non-HCC mortality | Cum. risk of HCC (%) | Cum. risk of non-HCC death (%) |
|----------------|-------------|---------------|-----------|-------------------|----------------------|--------------------------------|
| AMAP           | Low         | 36 (7.6)      | 0         | 5                 | 0                    | 13.9                           |
|                | Medium      | 136 (28.8)    | 1         | 25                | 0.7                  | 18.4                           |
|                | High        | 301 (63.6)    | 21        | 133               | 7.0                  | 44.2                           |
| THRI           | Low         | 44 (9.3)      | 1         | 4                 | 2.3                  | 9.1                            |
|                | Medium      | 243 (51.4)    | 5         | 79                | 2.1                  | 32.5                           |
|                | High        | 186 (39.3)    | 16        | 80                | 8.6                  | 43.0                           |
| HCC risk score | Low         | 149 (38.0)    | 3         | 39                | 2.0                  | 26.1                           |
|                | Medium      | 185 (47.2)    | 12        | 63                | 6.5                  | 34.1                           |
|                | High        | 58 (14.8)     | 5         | 34                | 8.6                  | 63.0                           |
| ADDRESS HCC    | Low         | 158 (9.7)     | 3         | 51                | 1.9                  | 32.3                           |
|                | Medium      | 269 (56.9)    | 16        | 92                | 6.0                  | 34.2                           |
|                | High        | 46 (33.4)     | 3         | 20                | 6.5                  | 43.5                           |

**Table 3** AUROC & PPV/NPV

| Score          | AUROC               | Cut-off value | NPV (%) | PPV (%) |
|----------------|---------------------|---------------|---------|---------|
| AMAP           | 0.722 [0.636–0.808] | 50            | 100     | 5.0     |
|                |                     | 60            | 99.0    | 7.3     |
| THRI           | 0.694 [0.665–0.727] | 140           | 98.7    | 5.1     |
|                |                     | 240           | 97.9    | 8.4     |
| HCC risk score | 0.641 [0.510–0.772] | 1             | 97.4    | 5.0     |
|                |                     | 3             | 98.0    | 6.5     |
| ADDRESS HCC    | 0.654 [0.553–0.754] | 1             | 98.1    | 5.8     |
|                |                     | 3             | 95.6    | 6.4     |

AUROC area under receiver operating curve, NPV negative predictive value, PPV positive predictive value

**Performance of Risk Scores**

AUROC, PPV, and NPV of differing score strata are given in Table 3. No score was significantly better on AUROC comparison; however, the aMAP score had the highest point estimate of 0.722 (0.636–0.808). NPV was high in the low strata of all scores, ranging from 97.4% (HCC risk score) to 100% (AMAP). The low-risk strata of the HCC risk score included the highest proportion of the cohort at 38% (Table 1), with a NPV of 97.4%. Freedom from diagnosis of HCC according to each risk score’s risk strata is represented in Fig. 2.

**Predictive Value and Likelihood of Benefit or Harm from Surveillance**

PPV and NPV of surveillance tests according to HCC risk score strata are shown in Table 4. There were 17 true

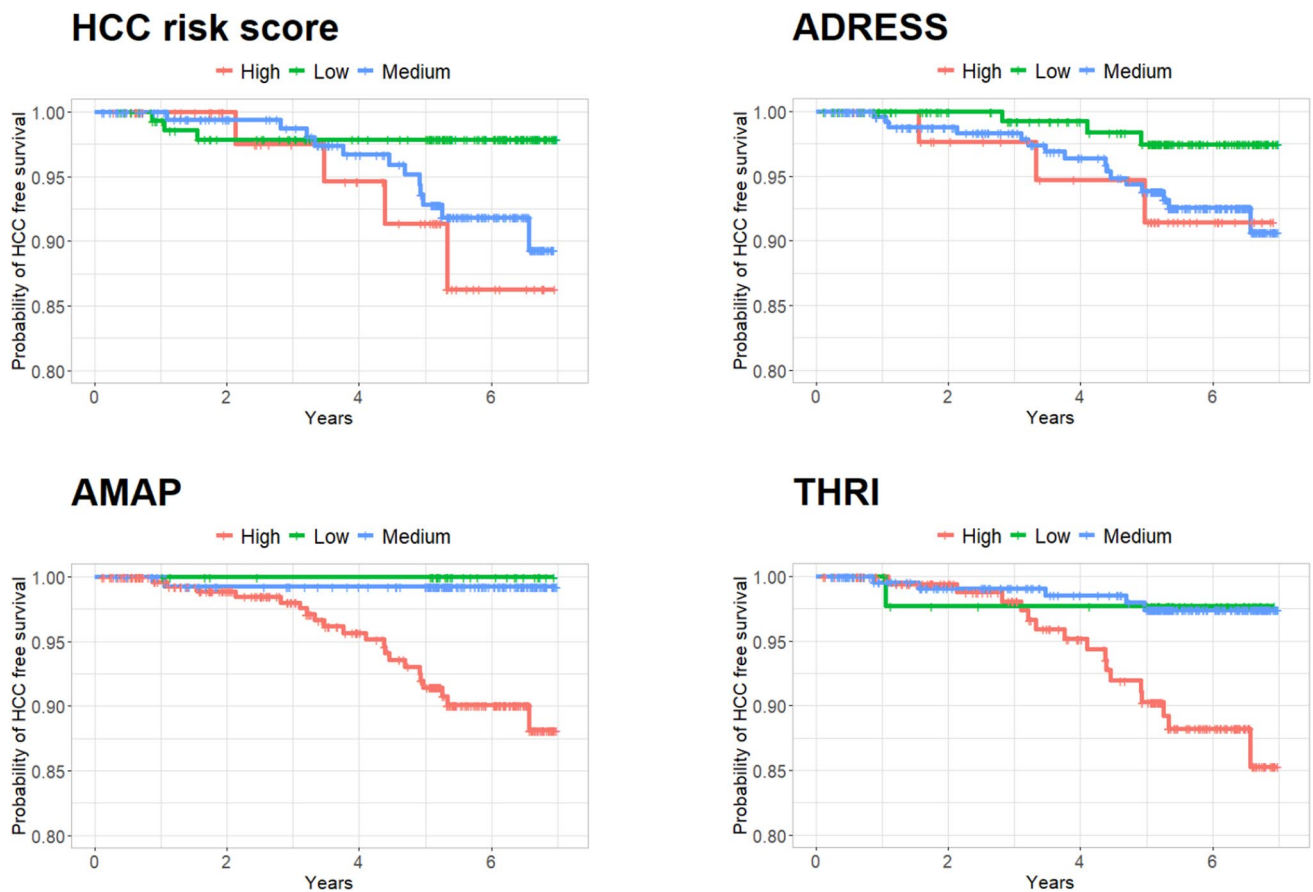
positives, 88 false positives, and no false negatives. 5 HCC were diagnosed in patients who had been non-adherent with surveillance and had presentations that triggered cross-sectional imaging directly. As there were no false negatives, the NPV of a normal screening test was 100% in all strata. The PPV of detection of HCC following an abnormal screening result increased across the strata (range 0–32.6%) and was lowest in the low-risk strata (range 0–9.4%). The PPV was 0 in the low- and medium-risk strata of the AMAP and the low-risk strata of the THRI.

**NNB & NNH**

13 patients were diagnosed at early stage via surveillance and therefore gained benefit. 88 false-positive surveillance tests triggered negative cross-sectional imaging (i.e., harm). Total NNB and NNH were 241 and 36, respectively. HCC was not diagnosed in any patients stratified as low risk by the AMAP or THRI score and therefore these groups gained no benefit from surveillance. Furthermore, patients in the low-risk categories of the ADDRESS HCC or HCC risk score had a high NNB (> 300 and > 900, respectively). NNH ranged from 24.0 to 46.8 across the strata. Additionally, all low-risk strata had a lower NNH (range 24–39) than the higher strata in their risk scores (33.9–46.8).

**Discussion**

In this cohort of mixed etiology cirrhosis, we assessed four models of HCC risk prediction aiming to quantify the benefit and harms of HCC surveillance. No score was superior in the prediction of HCC on AUROC comparison and the low-risk strata of all scores demonstrated a low PPV of an abnormal surveillance result, together with a high NNB and low NNH.



**Fig. 2** Kaplan–Meier freedom from HCC risk curves

Notably, no patients stratified as low risk by the AMAP or THRI score were diagnosed with HCC and therefore gained no benefit from surveillance. The overall NNB and NNH were 241 and 36 suggesting that around 6–7 patients are exposed to harm by way of further cross-sectional imaging for every early-stage HCC detected through surveillance.

Notably the rate of HCC diagnosis across all risk strata in this cohort is lower than that in the risk score development cohorts, creating the high NNB, low NNH, and low PPV seen here. Furthermore, this incidence rate is substantially lower than the threshold for entering surveillance recommended in international guidelines [1]. This may be explained partially by the THRI, ADRESS, and aMAP scores having a significantly smaller proportion of patients with non-viral hepatitis (and associated lower risk of HCC development) in their development cohorts than is present in ours [14, 15, 18]. However, it is likely our finding of a high competing mortality from non-HCC causes across all risk strata also plays a role. In this regard our data compare similarly to a recent large observational study of patients with cirrhosis in Denmark where a lower than expected incidence of HCC and a significantly higher risk of non-HCC death was observed [21].

The strengths of this study are in the assessment of four risk scores through several clinically relevant metrics in a cohort of patients with cirrhosis from broad range of etiologies. There are minimal missing data and a long period of follow-up. Building on research based on simulated cohorts by empirically calculating NNB and NNH from long-term follow-up, we sought to quantify more accurately the benefits and harms of HCC surveillance [9, 22, 23]. While simulated cohorts can provide an estimate of these, assumptions in such models such as a high incidence rate of HCC [24], very low rates of non-HCC competing mortality [9] and assuming 100% compliance with surveillance [22] limit interpretation.

Our cohort of patients derives from Glasgow, where almost half (47%) of residents are part of the most deprived quintile in Scotland [25]. There are high rates of alcohol use amid high hepatitis C prevalence creating a significant burden of chronic liver disease [26, 27]. The compliance of < 70% seen here reflects the real-world nature of our data and allows us to draw more accurate conclusions on the application and efficacy of HCC surveillance. We believe that this is the first time PPV of abnormal surveillance in

**Table 4** NNH, NNB, and PPV of surveillance

| Score          | Risk strata | No of US scans | Early-stage screen detected | Total screen detected (True positives) | False positives | NNB   | NNH  | PPV (%) |
|----------------|-------------|----------------|-----------------------------|--|-----------------|-------|------|---------|
| AMAP           | Low         | 220            | 0                           | 0                                      | 9               | –     | 24.4 | 0       |
|                | Medium      | 909            | 0                           | 0                                      | 20              | –     | 45.5 | 0       |
|                | High        | 2003           | 13                          | 17                                     | 59              | 154.1 | 33.9 | 22.3    |
| THRI           | Low         | 309            | 0                           | 0                                      | 10              | –     | 30.9 | 0       |
|                | Medium      | 1607           | 2                           | 3                                      | 49              | 803.5 | 32.8 | 5.8     |
|                | High        | 1221           | 11                          | 14                                     | 29              | 111.0 | 42.1 | 32.6    |
| HCC risk score | Low         | 963            | 1                           | 1                                      | 25              | 963   | 38.5 | 3.8     |
|                | Medium      | 1214           | 10                          | 11                                     | 35              | 121.4 | 35.7 | 23.9    |
|                | High        | 374            | 0                           | 3                                      | 8               | –     | 46.8 | 27.3    |
| ADDRESS - HCC  | Low         | 917            | 3                           | 3                                      | 29              | 323.7 | 31.6 | 9.4     |
|                | Medium      | 1860           | 9                           | 12                                     | 52              | 206.7 | 35.8 | 18.8    |
|                | High        | 306            | 1                           | 2                                      | 7               | 306.0 | 43.7 | 22.2    |

US ultrasound, NNB number needed to benefit, NNH number needed to harm, PPV positive predictive value

the context of HCC surveillance has been calculated from observational data.

Limitations of this study include its retrospective nature, albeit of a prospective database, and the small number of patients developing HCC in the cohort. This generates a wide range of NNB across the risk groups and signifies that the study may be underpowered to calculate this metric across the risk strata. Furthermore, this may explain the lack of a significant difference in the AUROC of the risk scores as a possible type II error. Despite this we believe our results provide a valuable framework for planning prospective larger-scale studies in this area. For scores such as the AMAP and THRI score the proportion of patients classified as low risk was very low (7–10%) which may limit their usefulness in clinical practice. In comparison the HCC risk score classifies a significantly larger proportion as low risk, although at the expense of missing several surveillance detected cancers. Recently published research suggests that most US providers would still offer surveillance if HCC risk was less than 0.5% per annum [28]. Further research is required as to what degree of risk is acceptable to patients considering surveillance. Our research gives context of harms experienced according to risk strata, which may aid informed decision-making.

In this regard, we defined harm as undergoing further investigation with cross-sectional imaging for false-positive surveillance tests, one of many harms identified in a recent review [29]. It is recognized that false-positive tests generate anxiety [30] and cause financial harm. This may be direct, in healthcare settings where the patient pays directly for investigations, or indirect via increased insurance premiums or loss of earnings while attending for investigations. Specific physical harms related to investigation such as

contrast nephropathy, extravasation of contrast, and biopsy-specific complications such as bleeding have been described [29], however, were not captured in our study. A prospective trial of HCC surveillance-related harms [31] is underway and should provide further clarity on the frequency of such harms.

Additionally, data were not available on the nature of the AFP levels or USS appearances which triggered further investigation, therefore patients progressed to cross-sectional imaging at the discretion of the treating physician. This introduces judgment and generates variation based on the physician’s tolerance to risk and attitude to borderline abnormal ultrasound and AFP results. A physician’s decision to investigate further is made on analysis of AFP over time rather than discrete values, with mathematical analysis of AFP trends confirming that such an approach is potentially beneficial [32]. This would further complicate any attempt to analyze AFP levels leading to investigation. Compared with other research we found a lower risk of harm of 18.3% compared with the 27.5% estimated in the USA in 2017 [7]. This could be explained by a lower threshold to proceed to further investigation in the US, due to cultural variation in practice related to factors that may include the perceived threat of litigation [23].

Adherence with surveillance in our cohort was relatively high at 68% compared to the 52% reported in a previous meta-analysis [33]. It may be expected that poorer adherence may result in reduced benefit, as well as reduced harm. We would not expect the reduction in harms vs benefits to be disproportionate; however, further studies in cohorts with different levels of adherence to surveillance would be beneficial.

It is recognized that the benefit attributed to detection of early-stage HCC through surveillance may be subject to length and lead time bias [6]. Length time bias may be introduced where HCCs detected are more indolent and less aggressive and would potentially be detected and treated without surveillance [34]. Furthermore, diagnosis of early-stage HCC is a surrogate outcome susceptible to lead time bias in which early detection of cancer through surveillance is erroneously associated with a longer life expectancy. Recent research adjusting for lead time bias in HCC surveillance has demonstrated a reduced effect on mortality than previously described [6]. Considering this, we could potentially be overestimating the benefit estimated by surveillance and generating a misleadingly low NNB.

## Conclusion

Using multiple clinically relevant methods of validation we have demonstrated the value of risk scores as tools to individualize the decision to offer patients HCC surveillance. The high competing mortality rate and low incident rate of HCC drive the higher NNB, lower NNH, and lower PPV, particularly in the lower-risk strata. This factor is likely to have been underappreciated in previous simulation work estimating the benefits and harms of HCC surveillance. We therefore believe large scale, prospective research focusing on risk-based surveillance strategies such as intensive HCC surveillance for higher-risk groups and exclusion of lower-risk patients from surveillance is warranted.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s10620-022-07731-1>.

**Author's contribution** CC and SB designed the study. Material preparation, data collection, and analysis were performed by CC and SB. The first draft of the manuscript was written by CC and SB and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

**Funding** None.

**Data availability** Due to the nature of this research, participants of this study did not agree for their data to be shared publicly, so supporting data are not available.

## Declarations

**Conflict of interest** Chris Curran, Matthew Priest, Shouren Datta, Adrian Stanley, and Ewan Forrest have no relevant financial or non-financial interests to disclose.

**Research involving human participants or animals** This article does not contain any studies with animal subjects.

**Ethical approval** Caldicott approval was sought from the health board and was granted to ensure appropriate handling of patient data. Additional ethical permission for this retrospective analysis of existing data

was not required. The study was performed in accordance with the ethical standards as laid down in the 1964 Declaration of Helsinki.

**Consent to participate** Not applicable.

**Consent to publication** All authors have consented to the publication of this manuscript.

## References

1. EASL–EORTC Clinical Practice Guidelines: Management of hepatocellular carcinoma. *Eur J Cancer*. 2012 Mar 1;48(5):599–641.
2. Marrero JA, Kulik LM, Sirlin CB, Zhu AX, Finn RS, Abecassis MM et al. Diagnosis, Staging, and Management of Hepatocellular Carcinoma: 2018 Practice Guidance by the American Association for the Study of Liver Diseases. *Hepatology*. 2018;68:723–750.
3. Singal AG, Pillai A, Tiro J. Early detection, curative treatment, and survival rates for hepatocellular carcinoma surveillance in patients with cirrhosis: a meta-analysis. *PLoS Med*. 2014;11:e1001624.
4. Chen VL, Singal AG, Tapper EB, Parikh ND. Hepatocellular carcinoma surveillance, early detection and survival in a privately insured US cohort. *Liver Int*. 2020;40:947–955.
5. 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;101:422–434.
6. Haq MI, Drake TM, Goh TL, Ahmed A, Forrest E, Barclay S et al. Effect of Hepatocellular Carcinoma Surveillance Programmes on Overall Survival in a Mixed Cirrhotic UK Population: A Prospective, Longitudinal Cohort Study. *J Clin Med Res*. 2021. <https://doi.org/10.3390/jcm10132770>.
7. Jepsen P, West J. We need stronger evidence for (or against) hepatocellular carcinoma surveillance. *J. Hepatol*. 2021. <https://doi.org/10.1016/j.jhep.2020.12.029>.
8. Atiq O, Tiro J, Yopp AC, Muffer A, Marrero JA, Parikh ND et al. An assessment of benefits and harms of hepatocellular carcinoma surveillance in patients with cirrhosis. *Hepatology* 2017;65:1196–1205.
9. Taylor EJ, Jones RL, Guthrie JA, Rowe IA. Modeling the benefits and harms of surveillance for hepatocellular carcinoma: information to support informed choices. *Hepatology* 2017;66:1546–1555.
10. Konerman MA, Verma A, Zhao B, Singal AG, Lok AS, Parikh ND. Frequency and outcomes of abnormal imaging in patients with cirrhosis enrolled in a hepatocellular carcinoma surveillance program. *Liver Transpl*. 2019;25:369–379.
11. Vilar-Gomez E, Calzadilla-Bertot L, Wai-Sun Wong V, Castellanos M, Aller-de la Fuente R, Metwally M et al. Fibrosis severity as a determinant of cause-specific mortality in patients with advanced nonalcoholic fatty liver disease: a Multi-National Cohort Study. *Gastroenterology* 2018;155:443–457.e17.
12. Bolondi L, Gaiani S, Casali A, Serra C, Piscaglia F. Screening for the early diagnosis of hepatocellular carcinoma: cost-effectiveness analysis. *Radiol. Med*. 1997;94:4–7.
13. NICE. Cirrhosis in the over 16s - NICE Guideline 50 [Internet]. NICE; 2016 Jul. Report No.: NG50. Available from: <https://www.nice.org.uk/guidance/ng50/evidence/full-guideline-2546537581>
14. Fan R, Papatheodoridis G, Sun J, Innes H, Toyoda H, Xie Q et al. aMAP risk score predicts hepatocellular carcinoma development in patients with chronic hepatitis. *J. Hepatol*. 2020;73:1368–1378.

15. Flemming JA, Yang JD, Vittinghoff E, Kim WR, Terrault NA. Risk prediction of hepatocellular carcinoma in patients with cirrhosis: the ADDRESS-HCC risk model. *Cancer* 2014;120:3485–3493.
16. Ioannou GN, Green P, Kerr KF, Berry K. Models estimating risk of hepatocellular carcinoma in patients with alcohol or NAFLD-related cirrhosis for risk stratification. *J. Hepatol.* 2019;71:523–533.
17. Ioannou GN, Tang W, Beste LA, Tincopa MA, Su GL, Van T et al. Assessment of a deep learning model to predict hepatocellular carcinoma in patients with Hepatitis C cirrhosis. *JAMA Netw. Open.* 2020;3:e2015626.
18. Sharma SA, Kowgier M, Hansen BE, Brouwer WP, Maan R, Wong D et al. Toronto HCC risk index: a validated scoring system to predict 10-year risk of HCC in patients with cirrhosis. *J. Hepatol.* 2017. <https://doi.org/10.1016/j.jhep.2017.07.033>.
19. Goossens N, Singal AG, King LY, Andersson KL, Fuchs BC, Besa C et al. Cost-effectiveness of risk score-stratified hepatocellular carcinoma screening in patients with cirrhosis. *Clin. Transl. Gastroenterol.* 2017;8:e101.
20. Ioannou GN, Green PK, Beste LA, Mun EJ, Kerr KF, Berry K. Development of models estimating the risk of hepatocellular carcinoma after antiviral treatment for hepatitis C. *J. Hepatol.* 2018;69:1088–1098.
21. Jepsen P, Ott P, Andersen PK, Sørensen HT, Vilstrup H. Clinical course of alcoholic liver cirrhosis: a Danish population-based cohort study. *Hepatology* 2010;51:1675–1682.
22. Roskilly A, Rowe IA. Surveillance for hepatocellular cancer. *Clin. Med.* 2018;18:s66–s69.
23. Cucchetti A, Trevisani F, Pecorelli A, Erroi V, Farinati F, Ciccarese F et al. Estimation of lead-time bias and its impact on the outcome of surveillance for the early diagnosis of hepatocellular carcinoma. *J. Hepatol.* 2014;61:333–341.
24. Thompson Coon J, Rogers G, Hewson P, Wright D, Anderson R, Jackson S et al. Surveillance of cirrhosis for hepatocellular carcinoma: a cost-utility analysis. *Br. J. Cancer* 2008;98:1166–1175.
25. Scottish Index of Multiple Deprivation 2016 [Internet]. [cited 2021 Apr 23]. Available from: <https://www.gov.scot/collections/scottish-index-of-multiple-deprivation-2020/>
26. Hutchinson SJ, Roy KM, Wadd S, Bird SM, Taylor A, Anderson E, Shaw L, Codere G, Goldberg DJ. Hepatitis C Virus infection in Scotland: epidemiological review and public health challenges. *Scot. Med. J.* 2006;51:8–15.
27. Giles L RE. Monitoring and Evaluating Scotland’s Alcohol Strategy: Monitoring Report 2020 [Internet]. Public Health Scotland; 2020. Available from: <https://www.careknowledge.com/media/47431/mesas-monitoring-report-2020.pdf>
28. Kim NJ, Rozenberg-Ben-Dror K, Jacob DA, Rich N, Singal AG, Aby ES et al. Provider attitudes toward risk-based hepatocellular carcinoma surveillance in patients with cirrhosis in the United States. *Clin. Gastroenterol. Hepatol.* 2020;20:183–193.
29. Petrasek J, Singal AG, Rich NE. Harms of hepatocellular carcinoma surveillance. *Curr. Hepatol. Rep.* 2019;18:383–389.
30. Vermeer NCA, Snijders HS, Holman FA, Liefers GJ, Bastiaannet E, van de Velde CJH et al. Colorectal cancer screening: systematic review of screen-related morbidity and mortality. *Cancer Treat Rev.* 2017;54:87–98.
31. ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). Identifier NCT03756051, Harms of Hepatocellular Carcinoma Surveillance, 2018 Nov 28 [cited 2022 Jun 21]. Available from: <https://clinicaltrials.gov/ct2/show/NCT03756051>
32. Bird TG, Dimitropoulou P, Turner RM, Jenks SJ, Cusack P, Hey S et al. Alpha-Fetoprotein detection of hepatocellular carcinoma leads to a standardized analysis of dynamic AFP to improve screening based detection. *PLoS ONE* 2016;11:e0156801.
33. Zhao C, Jin M, Le RH, Le MH, Chen VL, Jin M et al. Poor adherence to hepatocellular carcinoma surveillance: a systematic review and meta-analysis of a complex issue. *Liver Int.* 2018;38:503–514.
34. Cucchetti A, Garuti F, Pinna AD, Trevisani F, Italian Liver Cancer (ITA.LI.CA) Group. Length time bias in surveillance for hepatocellular carcinoma and how to avoid it. *Hepatol. Res.* 2016;46:1275–80.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.