Chien-Hung Chen[2]
Benjamin Maasoumy[3], for the CREATE study group
[1]*Department of Gastroenterology and Hepatology, Erasmus MC University Medical Center, Rotterdam, The Netherlands*
[2]*Department of Internal Medicine, Koahsiung Chang Gung Memorial Hospital, Kaohsiung, Taiwan*
[3]*Department of Gastroenterology and Hepatology, Hannover Medical School, Hannover, Germany*
[*]Corresponding author. Address: Department of Gastroenterology and Hepatology, Erasmus MC University Medical Center, Rotterdam, the Netherlands.
E-mail address: m.j.sonneveld@erasmusmc.nl (M.J. Sonneveld)

HCC prediction post SVR: Many tools yet limited generalizability!

To the Editor:

Despite attaining a sustained virological response (SVR), the risk of hepatocellular carcinoma (HCC) remains a significant concern in patients with chronic hepatitis C (CHC). The EASL guidelines advise HCC screening in a population with a high incidence of HCC, considering cost, expertise, treatment options, and rate of tumor growth.[1] Accordingly, HCC screening is recommended in patients with CHC and >F3 fibrosis. Several prediction tools have been applied in various studies for HCC prediction; however, none is generalizable to the global population to date. Recently, novel HCC prediction models based on artificial intelligence have shown the superiority of deep learning models over traditional statistical models. However, these models are not restricted to patients who attain SVR and are yet to be validated[2,3](Table 1). Semmler *et al.*, in a recent study, developed a simple bedside score incorporating post SVR variables such as age, albumin, liver stiffness measurement (LSM), alpha-fetoprotein, and alcohol consumption to stratify future risk of HCC among patients with CHC and compensated advanced chronic liver disease.[4] They suggested avoiding HCC screening in low-risk (<1% person-year) individuals; however, certain issues in the study merit further discussion.

Firstly, the selection of cohorts – the derivation cohort included 527 patients from 3 institutions in Europe in whom hepatic venous pressure gradient, LSM, and histopathology were assessed. These patients were derived from subgroups of patients from 6 different studies where HCC detection was not a target outcome. Such a retrospective inclusion may not accurately represent the real-world population of patients with CHC. Further, the validation cohort included 1,500 patients across multiple European centers. However, it is worth noting that the largest share of patients with CHC reside in Asia and Africa (about 78%).[5] Whether the findings generalize to such populations or if the score needs further fine-tuning is still questionable.

Secondly, patients with CHC are known to have a progressive reduction in inflammation and fibrosis, including LSM, till 24-96 weeks post-SVR.[6] In the present study, the authors included LSM at 12 weeks post-treatment in the derivation cohort and 48 weeks post-treatment in the validation cohort. Indeed, they represent 2 different time frames with varying values of LSM and relate to dynamic changes in liver inflammation, fibrosis, and normalization of liver functions.

Thirdly, the authors provided a simple score-based approach to stratify the risk of HCC as low or high. However, to generalize this score to a large population, the performance characteristics of the final score in terms of discrimination (c-index), sensitivity, specificity, and accuracy are highly desired for policy decisions. It is also unclear whether the score is biased toward better categorizing patients without HCC due to class imbalance problems in the derivation cohort.[7]

Fourthly, the authors based their research discussion on the proposition of <50,000 USD per life-year saved as a cost-effective approach.[1] However, we feel that the cost of testing and surveillance varies significantly across the world. The cost-effectiveness thresholds derived from developed countries for a disease that is much more prevalent in developing and underdeveloped countries with different economic standards, remain biased and merit further research.

Fifthly, the points given for variables in the final score were derived from sub-distribution hazard ratios (SHRs) of a competing risk model with death and liver transplant as competing events and the development of HCC as a target event. These individual points were added to obtain the final risk prediction score. However, adding such SHRs has been identified as mathematically incorrect and misleading in prediction models. Literature suggests a mathematical addition of the coefficients of variables rather than exponential of coefficients (*i.e.*, SHR) in the model as an appropriate approach.[8]

Received 3 March 2022; received in revised form 14 April 2022; accepted 15 April 2022; available online 05 May 2022
https://doi.org/10.1016/j.jhep.2022.04.028

Table 1. Studies reporting predictive scores for HCC.[2,3]

Study	Patient population, fibrosis stage	Score name	Variables included	Algorithm	Risk classes	HCC rate according to risk classes
Abe, 2020	CHC, F4, Post SVR	NA	ALBI score, platelet, diabetes status	0 or 1 points for each: ALBI score ≤ or >2-3, platelet ≥ or < 8.2×10^4/μL, absence or presence of diabetes	0–1: Low-score 2–3: High-score	Low- *vs.* high-score group 0.7% *vs.* 12.5% at 1 year 2.2% *vs.* 15.2% at 2 years 3.1% *vs.* 33.9% at 3 years 3.1% *vs.* 41.2% at 4 years
Fan, 2020	CHC, F4, Post SVR	aMAP	Age, gender bilirubin albumin, platelet	Mathematical formula	<50: Low-risk 50–60: Intermediate risk >60: High-risk	Low *vs.* intermediate *vs.* high-risk 0–0.8% *vs.* 1.5–4.8% *vs.* 8.1–17.8% at 3–5 year
Shiha, 2020	CHC, F3-F4, Post SVR	GES	Age, male gender, AFP, albumin, fibrosis	0 points to 3.5 points: Female or Male, Age ≤ or > 54 years, Albumin ≥ or <3.8 g/dl, AFP ≤ or >20 ng/mL F3 or F4	GES ≤6: Low Risk GES 6–7.5: Intermediate Risk GES >7.5: High Risk	Low *vs.* intermediate *vs.* high-risk 0.1% *vs.* 0.7% *vs.* 1.2% at 1 year 1.2% *vs.* 3.3% *vs.* 7.1% at 2 years 1.9% *vs.* 5.8% *vs.* 9.5% at 3 years
Alonso-Lopez, 2020	CHC, F3-F4, Post SVR	NA	**LSM model:** Albumin, LSM, SVR48 ΔLSM **FIB-4 model:** Albumin, FIB-4, SVR48 FIB-4, SVR48 GGT	**LSM model (0 or 1 points):** Albumin ≥ or < 4.2 g/dl, LSM ≤ or > 17.3 kPa, ΔLSM ≥ or <25.5% **FIB-4 model (0 to 2 points):** Albumin≥ or <4.2 g/dl, FIB-4 ≤ or >3.7, SVR48 FIB-4 ≤ or >3.3, SVR48 GGT ≤ or >42 U/l	**LSM model:** Score 0-1-2-3 **FIB-4 model:** Score 1–2 *vs.* 3–4 *vs.* 5–6	**LSM model** Score 0 *vs.* 1 *vs.* 2 *vs.* 3 0% *vs.* 2.1% *vs.* 5.8% *vs.* 16.3% at 3 years **FIB-4 model** Score 1–2 *vs.* 3–4 *vs.* 5–6 0.4% *vs.* 1.7% *vs.* 6.5 *vs.* 19% at 3 years
Watanabe, 2019	CHC, Post SVR	NA	**Pre-DAA model:** FIB-4, albumin, gender **Post-DAA model:** end of treatment FIB-4, AFP	**Pre-DAA model:** 0 or 1 points for FIB-4 < or ≥4.0, albumin > or ≤3.8 g/dl, female or male **Post-DAA model:** 0 or 1 points for FIB-4 < or ≥4.0, AFP< or ≥6.0 ng/mL	**Pre-DAA model:** 0: Low Risk, 1–2: Intermediate Risk, 3: High Risk **Post-DAA model:** 0: Low, 1: Intermediate, 2: High	**Pre-DAA model** Low *vs.* intermediate *vs.* high risk 0.4% *vs.* 2.1% *vs.* 9.5% at 1 year 0.4% *vs.* 4.4% *vs.* 16.4% at 2 years **Post-DAA model** Low *vs.* intermediate *vs.* high risk 0.4% *vs.* 1.4% *vs.* 6.1% at 1 year 0.4% *vs.* 3.2% *vs.* 14.4% at 2 years
Hiraoka, 2019	CHC, Post SVR	ADRES	Gender, SVR24, FIB-4, SVR24-AFP	1 point to each variable: male, FIB-4 >3.25, AFP >5 ng/mL	ADRES 0-1-2-3	ADRES 0 *vs.* 1 *vs.* 2 *vs.* 3 0% *vs.* 0.5% *vs.* 8.4% *vs.* 18% at 1 year 0% *vs.* 1.6% *vs.* 13.4% *vs.* 32.8% at 2 years
Iio, 2019	CHC, Post SVR	NA	SVR24-AFP, SVR24-FIB-4, TLL1 AA/TT	1 point to each variable: AFP >4.6 ng/mL, FIB-4 >2.67, TLL1 AA/TT	0: Low Risk, 1–2: Intermediate Risk, 3: High Risk	Low *vs.* intermediate *vs.* high risk 0% *vs.* 2.2% *vs.* 10.4% at 1 year 0% *vs.* 3.0% *vs.* 13.6% at 2 years
Tani, 2020	CHC, Post SVR+	NA	End of treatment: Age, AFP	0 to 1 point: Age < or ≥75 years, AFP < or ≥6 ng/mL	Score 0-1-2	Score 0 *vs.* 1 *vs.* 2 0.3% *vs.* 1.05% *vs.* 4.92% at 1 year 0.3% *vs.* 6.27% *vs.* 18.37% at 2 years 1.26% *vs.* 10.45% *vs.* 18.37% at 3 years
Ioannou, 2020*	CHC Cirrhosis, Both SVR and Non-SVR	RNN	Clinical variables	Recurrent neural network	-	RNN predicted HCC development with AUC of 0.759, and AUC of 0.806 among SVR achieved patients
Phan, 2020*	CHB and CHC	CNN	Disease history	Convolutional neural network;	-	CNN achieved an accuracy of 0.980 and AUC of 0.886 for predicting the development of HCC among viral hepatitis patients
Nam, 2020*	CHB and CHC cirrhosis	DL	Clinical variables	Deep Learning	-	DL model achieved accuracy of 0.763 and AUC of 0.782 in validation cohort and outperformed the previous models

AFP, alpha-fetoprotein; CHB, chronic hepatitis B; CHC, chronic hepatitis C; DAA, direct-acting antiviral; HCC, hepatocellular carcinoma; LSM, liver stiffness measurement; SVR, sustained virological response.
*Denotes deep learning models using artificial intelligence.

Finally, the study provides an easy single-time prediction score for risk stratification of HCC in a limited setting. Further multi-ethnic studies are warranted for HCC prediction with better generalizability.

Financial support
The authors received no financial support to produce this manuscript.

Conflict of interest
The authors declare no conflicts of interest.

Please refer to the accompanying ICMJE disclosure forms for further details.

Authors' contributions
NB – Writing – original draft, NV – Conceptualization, Writing – review and editing, VS – Conceptualization, Writing – review and editing.

Supplementary data
Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhep.2022.04.028.

References
[1] EASL Clinical Practice Guidelines. Management of hepatocellular carcinoma. J Hepatol 2018;69:182–236.
[2] D'Ambrosio R, Degasperi E, Lampertico P. Predicting hepatocellular carcinoma risk in patients with chronic HCV infection and a sustained virological response to direct-acting antivirals. J Hepatocell Carcinoma 2021;8:713–739.
[3] Ahn JC, Qureshi TA, Singal AG, Li D, Yang JD. Deep learning in hepatocellular carcinoma: current status and future perspectives. World J Hepatol 2021 Dec 27;13(12):2039–2051.
[4] Semmler G, Meyer EL, Kozbial K, Schwabl P, Hametner-Schreil S, Zanetto A, et al. HCC risk stratification after cure of hepatitis C in patients with compensated advanced chronic liver disease. J Hepatol 2022;76(4):812–821.
[5] Petruzziello A, Marigliano S, Loquercio G, Cozzolino A, Cacciapuoti C. Global epidemiology of hepatitis C virus infection: an up-date of the distribution and circulation of hepatitis C virus genotypes. World J Gastroenterol 2016;22:7824–7840.
[6] Pietsch V, Deterding K, Attia D, Ringe KI, Heidrich B, Cornberg M, et al. Long-term changes in liver elasticity in hepatitis C virus-infected patients with sustained virologic response after treatment with direct-acting antivirals. United European Gastroenterol J 2018;6:1188–1198.
[7] Ling CX, Sheng VS. Class imbalance problem. In: Sammut C, Webb GI, editors. Encyclopedia of machine learning. Boston, MA: Springer US; 2010. 171-171.
[8] Moons KGM, Harrell FE, Steyerberg EW. Should scoring rules be based on odds ratios or regression coefficients? J Clin Epidemiol 2002;55:1054–1055.

Naveen Bhagat
Nipun Verma
Virendra Singh*
Department of Hepatology, Postgraduate Institute of Medical Education and Research, Chandigarh, India
*Corresponding author. Address: Department of Hepatology, Postgraduate Institute of Medical Education and Research, Chandigarh, 160 012 India; Tel.: 911722756338, fax: 91-172-2744401.
E-mail address: virendrasingh100@hotmail.com (V. Singh)

Reply to: "HCC prediction post SVR: many tools yet limited generalizability!"

De novo HCC risk stratification after HCV cure: All roads lead to Rome?

To the Editor:

We would like to thank Dr. Bhagat and colleagues[1] for their interest in our study on *de novo* hepatocellular carcinoma (HCC) risk stratification in compensated advanced chronic liver disease (cACLD) patients who achieved sustained virologic response (SVR).[2] The authors provided a summary of 11 selected risk stratification approaches and highlighted 5 points related to our study but also HCC risk stratification/screening in general.

The multiplicity of available risk stratification tools underlines the high scientific interest in this clinically relevant issue. While some approaches follow beaten tracks (simple algorithms based on conventional statistical methods; *e.g.*, Pons *et al.*[3]), complex artificial intelligence-based methods discovered alternative routes that are less obvious. However, several factors/variables are shared between models, indicating their key role in regard to the outcome *de novo* HCC: age and alpha-fetoprotein, next to surrogates of hepatic dysfunction (serum albumin) and liver fibrosis/portal hypertension (*i.e.*, liver stiffness measurement [LSM] and platelet count as well as its derivatives, such as the fibrosis-4 score). It is important to note that variable selection is determined by their availability in retrospective datasets, indicating that their intersections – *i.e.*, broadly available parameters – are likely overrepresented throughout the different risk stratification tools. Importantly, our derivation cohort of comprehensively characterized patients allowed for the selection of the best, rather than the best available predictors.

As addressed by Bhagat *et al.*[1] in their first point, ACLD was diagnosed by either hepatic venous pressure gradient measurement, LSM, or histology in our derivation cohort. However, this does not induce selection bias, as liver disease severity was staged by one or more of these methods in all patients undergoing antiviral therapy, due to implications for reimbursement and regimen selection. Moreover, the generalizability of our risk stratification approach to other predominately Caucasian patient cohorts was confirmed by extensive validation – which is of paramount importance in the context of predictive modeling, but has often been omitted

Received 7 July 2022; accepted 9 July 2022; available online 20 July 2022
https://doi.org/10.1016/j.jhep.2022.07.011