We applied the documented observations from the datasets to examine the prevalence of ACLF and their 30-day mortality rates. When 'death' is used as an end-point, incorporation bias may not be applicable. Our critics are perhaps referring to a broad terminology when the gold standard and diagnostic test are not entirely independent, but it is unavoidable for our study.

Another suggestion was to compare a continuous numeric scoring system (CLIF-C ACLF scores) with a binary NACSELD model. We do not believe it is logical and identifying the "best" prediction model for 30-day mortality was not our primary objective. To assert that 51.2% of patients who had grade-3 ACLF by EASL-CLIF criteria did not have ACLF by NACSELD because of 'inappropriate exclusions or presence of liver or coagulation failure among study population' is mere speculation and not based on objective evidence. Verma *et al.* also suggest that "it was strange to note that 18.6% patients with grade-III EASL-ACLF did not have any OF by NACSELD criteria". Precisely, that was our point. We do not believe that a competing risk analysis would have altered our conclusions regarding the poor sensitivity in detecting ACLF by NACSELD. The all-cause mortality rates for EASL-CLIF ACLF grade 3 and NACSELD ACLF (≥2 OF) were similar (25.8% *vs.* 28.2%) in our study, but when patients with no OF by NACSELD were stratified by EASL-CLIF grades 0-3, the transplant-free mortality rates ranged from 1.5% to 86.0%.

By citing one of their recent publications, Verma *et al.* suggest that "the sensitivity of the NACSELD-definition is better than EASL-ACLF for mortality prediction".[5] In that study, the sensitivity, specificity, positive predictive value, and negative predictive value for NACSELD-ACLF-binary were reported as 100%, 0.0%, 65.1%, and 0.0% respectively. The highest sensitivity of the NACSELD-ACLF binary was reached by keeping specificity at 0%. This observation was perhaps overlooked by Verma and colleagues.

Indeed, we did not compare APASL-ACLF with either EASL-CLIF or NACSELD. We had stated the reasons for that in our manuscript. More importantly, we had cited a comparative study of EASL-CLIF and APASL using a Veteran Affairs administrative dataset where they found that 76.1% of patients with EASL-ACLF did not fulfil APASL criteria.[6] Moreover, the 30-day mortality was 37.6% for those who met the EASL-CLIF criteria suggesting that APASL criteria probably missed 75% of patients with ACLF with a very high short-term mortality.

Finally, we believe that a serious debate on ACLF is fundamental to identifying the best model and understanding its role in managing our patients.

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

Conflict of interest
The authors declare no conflicts of interest that pertain to this work.

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

Authors' contributions
Paul Thuluvath and Feng LI drafted the response.

Supplementary data
Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhep.2021.10.007.

References
[1] Verma N, Mehtani R, Duseja A. Universal definition and prognostication in acute-on-chronic liver failure – an unmet need! J Hepatol 2022;76(1):241–242.
[2] Thuluvath PJ, Thuluvath AJ, Hanish S, Savva Y. Liver transplantation in patients with multiple organ failures: feasibility and outcomes. J Hepatol 2018;69(5):1047–1056.
[3] Karch A, Koch A, Zapf A, Zerr I, Karch A. Partial verification bias and incorporation bias affected accuracy estimates of diagnostic studies for biomarkers that were part of an existing composite gold standard. J Clin Epidemiol 2016 Oct;78:73–82.
[4] Worster A, Carpenter C. Incorporation bias in studies of diagnostic tests: how to avoid being biased about bias. CJEM 2008;10(2):174–175.
[5] Verma N, Dhiman RK, Singh V, Duseja A, Taneja S, Choudhury A, et al. Comparative accuracy of prognostic models for short-term mortality in acute-on-chronic liver failure patients: CAP-ACLF. Hepatol Int 2021;15:753–765.
[6] Mahmud N, Kaplan DE, Taddei TH, Goldberg DS. Incidence and mortality of acute on chronic liver failure using two definitions in patients with compensated cirrhosis. Hepatology 2019;90:2150–2163.

Paul J. Thuluvath[1,2,*]
Feng Li[1]
[1]*Institute of Digestive Health & Liver Diseases, Mercy Medical Center, Baltimore, USA*
[2]*Department of Medicine, University of Maryland School of Medicine, Baltimore, MD, 21202, USA*
[*]Corresponding author. Address: Institute of Digestive Health & Liver Diseases, Mercy Medical Center, Baltimore, MD 21202, USA; Tel.: 1 410 332 9308
E-mail address: thuluvath@gmail.com (P.J. Thuluvath)

FIPS score: Solid but we still need even better!

To the Editor:
Bettinger and colleagues recently described a new score, the Freiburg index of post-TIPS survival (FIPS), which they developed to stratify mortality risk and identify good candidates for transjugular intrahepatic portosystemic shunt (TIPS) implantation.[1] The FIPS score, a linear predictor composed of albumin, creatinine, bilirubin and age, showed good prognostic performances to predict 3 month (M3) and 6-month (M6) post-TIPS survival and the threshold of ≥0.92 (85th percentile of their derivation cohort) was defined as the cut-off for high-risk patients before TIPS implantation.

Received 27 August 2021; accepted 30 August 2021; available online 23 October 2021
https://doi.org/10.1016/j.jhep.2021.08.026

We aimed to validate this linear predictor in the Toulouse cohort as a prognostic score of post-TIPS survival. All the patients with cirrhosis in whom a TIPS had been implanted between 2006 and 2020, except those treated for acute refractory bleeding or before abdominal surgery, were included. Clinical and biological data were collected before TIPS and patients were followed for at least 1 year. Statistical analyses were performed with STATA® 17.0 statistical software (Statacorp College Station, TX, USA).

Two hundred and seventy-seven patients were included (mean age 57 ± 9 years, 73% males, alcoholic etiology 76%). TIPS indication was refractory ascites (n = 160), preemptive TIPS (n = 26), secondary prophylaxis of variceal bleeding (n = 81), refractory hydrothorax (n = 10). Mean model for end-stage liver disease (MELD) and Child-Pugh scores were 12 ± 4 and 8 ± 2, respectively (main characteristics of the patients are listed in Table S1). The mean score (linear predictor) was -0.37 in our cohort compared to -0.12 in the German one, which implies our population was at lower risk, mostly because patients had already been selected as good candidates for TIPS in our tertiary care center, using Child-Pugh, MELD and bilirubin-platelet scores.[2]

In our cohort, we estimated that the regression coefficient on the linear predictor from the Cox model (calibration slope) was 1.104 (95% CI 0.596-1.612) at M6, the Harrell's C index was 0.736 (95% CI 0.649- 0.823) and the Gönen and Heller's K index 0.713. Those 3 analyses argue for a fair prognostic performance of the score. Nevertheless, the R-squared value (R^2_D from Royston and Sauerbrei) was 0.306, which means only 31% of the observations' variability was explained by the model's variables.[3]

Two hundred and sixty five of our patients were considered at low risk (FIPS score <0.92) and 12 at high risk (FIPS score ≥0.92). Three high-risk patients died within 3 months, and 4 within 6 months. Using the threshold of 0.92, sensitivity, specificity, positive (PPV) and negative (NPV) predictive values were 17%, 97%, 25% and 94% at 3 months, and 12%, 97%, 33% and 89% at 6 months, respectively. Survival in our cohort was significantly different between high-risk patients and the others (p = 0.004 at 3 months, p = 0.010 for the Log-Rank test comparing Kaplan-Meier survival curves at 6 months).

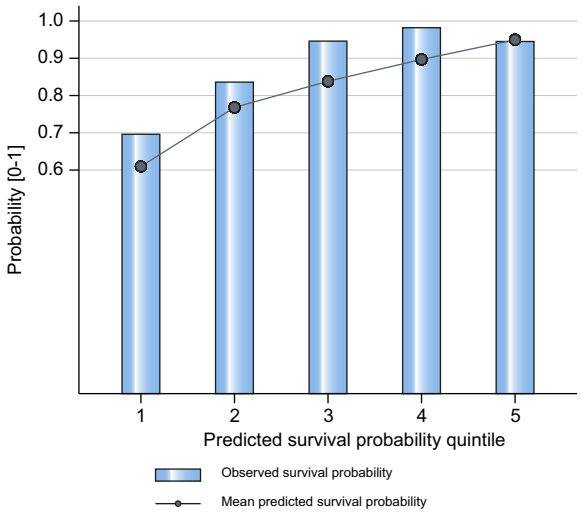

Fig 1. Comparison between mean predicted survival probability by FIPS score and observed survival probability in each quintile of predicted probability. Blue rectangles represent observed survival probability in each quintile of predicted probability. Greypoints represent mean predicted survival probability in each quintile of predicted probability. FIPS, Freiburg index of post-TIPS survival; TIPS, transjugular intrahepatic portosystemic shunt.

Besides, for any specific patient evaluated for TIPS, survival probability would be calculated with the formula given by Bettinger *et al.*: overall survival probability at time t = $S0(t)^{exp(FIPS+0.12)}$ with S0(t) = 0.87 at M3 and = 0.81 at M6. Fig. 1 shows mean predicted survival probability and observed survival probability in each quintile of predicted probability. Curves appear quite superimposable. Accordingly, the Freiburg score appears to be useful (with satisfying discrimination), but its sensitivity for identifying patients who will die remains poor. Research on new predictive markers that may further improve the discrimination of this prediction tool are still welcome. Besides, differences in absolute mortality rate between the German cohort in which the score was developed and other populations in which the score could be applied could have affected the performance of the model. Calibration of the score to different absolute mortality rates may address this problem and enable the effective use of the score in any group of potential candidates for TIPS.

Financial support

The authors received no financial support to produce this manuscript.

Conflict of interest

Christophe Bureau declares honoraria from Gore Inc (support for attending congress and speaker).

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

Authors' contributions

Hélène Larrue: collection of data and writing; Clara Brusq statistical analysis; Vanina Bongard: statistical analysis and critical review of the manuscript; Jean Pierre Vinel: statistical analysis and critical review of the manuscript; Christophe Bureau: scientific content and critical review of the manuscript.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhep.2021.08.026.

References

[1] Bettinger D, Sturm L, Pfaff L, Hahn F, Kloeckner R, Volkwein L, et al. Refining prediction of survival after TIPS with the novel Freiburg index of post-TIPS survival. J Hepatol 2021 Jun;74(6):1362–1372.
[2] Bureau C, Métivier S, D'Amico M, Péron JM, Otal P, Pagan JC, et al. Serum bilirubin and platelet count: a simple predictive model for survival in patients with refractory ascites treated by TIPS. J Hepatol 2011 May;54(5):901–907.
[3] Royston P, Altman DG. External validation of a Cox regression model: principles and methods. BMC Med Res Methodol 2013;13(1):33.

Hélène Larrue[1]
Clara Brusq[2]
Vanina Bongard[2,3]
Jean-Pierre Vinel[1]
Christophe Bureau[1,*]
[1]*Service d'Hépatologie, CHU de Toulouse, Université Toulouse 3 – Paul Sabatier, France*

JOURNAL
OF HEPATOLOGY

[2]USMR, Service d'épidémiologie clinique et santé publique, CHU de Toulouse, France
[3]CERPOP – UMR 1295 Inserm, Université Toulouse 3 – Paul Sabatier, France

*Corresponding author. Address: Service d'Hépatologie, CHU de Toulouse, 1 avenue Jean Poulhes 31059 Toulouse Cedex, France; Tel.: 033561323686, fax 033561323683.
E-mail address: bureau.c@chu-toulouse.fr (C. Bureau)

Poor performance of FIB-4 in elderly individuals at risk for chronic liver disease – implications for the clinical utility of the EASL NIT guideline

To the Editor:

With the identification of novel risk factors for chronic liver disease, the number of patients potentially eligible for referral for hepatologist consultation has expanded rapidly. Hence, non-invasive tools for risk stratification and identification of patients at highest risk are essential. We, therefore, read with great interest the 2021 update of the "*EASL Clinical Practice Guidelines on non-invasive tests for evaluation of liver disease severity and prognosis*".[1] Through the synthesis of available evidence and expert opinion, an algorithm was constructed to aid practitioners in identifying patients at highest risk of significant liver disease. In the algorithm, a major role has been allocated to fibrosis-4 (FIB-4) as an early stratification tool, despite uncertain diagnostic accuracy.[2,3] We assessed the performance of the FIB-4-based risk stratification in participants enrolled in the Rotterdam Study,[4] a large, population-based cohort with available data on liver biochemistry, metabolic syndrome, alcohol consumption, and liver stiffness. Participants at risk for chronic liver disease were defined according to the EASL guideline as having either metabolic syndrome (based on the ATP-III criteria[5]) or excessive alcohol consumption (≥ 20 grams or ≥30 grams daily for females/males) or viral hepatitis. Participants with viral hepatitis were excluded from this analysis since they require referral regardless of FIB-4 outcomes. Next, we applied the first step of the algorithm, which selects participants with FIB-4 ≥1.3 for liver stiffness assessment, and investigated its performance to detect fibrosis based on liver stiffness ≥8.0 kPa. In sensitivity analysis, we investigated the algorithm in several subgroups and applied a validated age-dependent cut-off for FIB-4, raising the cut-off for participants ≥65 years old from 1.3 to 2.0.[6]

We included 3.891 participants (aged 67.3±8.2, 44.2% male), of whom 6.0% had significant liver fibrosis based on liver stiffness. Among those considered at increased risk for chronic liver disease based on the presence of the metabolic syndrome and/or excessive alcohol consumption (n = 1,875, 8.6% liver stiffness

Table 1. Diagnostic performance of FIB-4 in elderly individuals at risk of chronic liver disease.

	n	LSM ≥8.0 kPa	Sens	Spec	NPV	PPV	Accuracy	DOR	Referral
Overall	1,849	159	74.8%	41.7%	94.6%	10.8%	44.6%	2.13	59.7%
Inclusion criteria									
Metabolic only	1,263	116	72.4%	40.9%	93.6%	11.0%	43.8%	1.82	60.3%
Alcohol only	320	16	81.3%	44.1%	97.8%	7.1%	45.9%	3.42	57.2%
Both	266	27	81.5%	42.7%	95.3%	13.8%	46.6%	3.28	59.8%
Liver enzymes									
Normal	1,340	65	69.2%	41.3%	96.3%	5.7%	42.7%	1.59	59.2%
Elevated	509	94	78.7%	42.9%	89.9%	23.8%	49.5%	2.78	61.1%
Age									
<65 years	669	40	47.5%	62.3%	94.9%	7.4%	61.4%	1.50	38.3%
≥65 years	1,180	119	84.0%	29.5%	94.3%	11.8%	35.0%	2.20	71.9%
≥65 years*	1,180	119	36.1%	76.2%	91.4%	14.5%	72.1%	1.81	25.1%
Diabetes									
Yes	395	62	75.8%	44.7%	90.9%	20.3%	49.6%	2.54	58.5%
No	1,435	95	73.7%	41.4%	95.7%	8.2%	43.6%	1.98	59.6%
Steatosis									
Yes	921	114	73.7%	46.1%	92.5%	16.2%	49.5%	2.39	56.4%
No	928	45	77.8%	37.7%	97.1%	6.0%	39.7%	2.12	63.0%

DOR, diagnostic odds ratio; FIB-4, fibrosis-4; LSM, liver stiffness measurement; NPV, negative predictive value; PPV, positive predictive value; sens, sensitivity; spec, specificity.
*An age-dependent cut-off of FIB-4 (2.0 instead of 1.3) was applied for individuals ≥65 years old.

Keywords: chronic liver disease; non-invasive tests; FIB-4; fibrosis; liver stiffness; epidemiology.
Received 16 July 2021; received in revised form 13 August 2021; accepted 16 August 2021; available online 28 August 2021
https://doi.org/10.1016/j.jhep.2021.08.017