External validation of the Freiburg index of post-TIPS survival

To the Editor:

We read with great interest the article "Refining prediction of survival after TIPS with the novel Freiburg index of post-TIPS survival" by Bettinger *et al.*[1] We congratulate the authors on developing a novel model (the FIPS score) to predict 3- and 6-month survival after planned transjugular intrahepatic portosystemic shunt (TIPS) implantation using 4 easily obtainable prognostic factors: age, creatinine, bilirubin, and albumin.

For a prediction model to be considered useful, it must be validated in an external, independent cohort.[2] In the Freiburg study, both the training set used to develop the model and the validation set were random samples from the same cohort with an equal allocation from each contributing study centre. As a result, the training and validation cohorts were identical except for sampling variation, meaning that there may still be substantial optimism bias in the estimates of model performance.[3] This potential bias was not corrected for during model development by statistical shrinkage techniques, nor during assessment of its performance, *e.g.* by bootstrapping techniques.[4] The lack of optimism correction might explain why the FIPS score outperformed the established scores in the Freiburg study. A related issue is the lack of external validation. The authors assessed the discriminatory performance of the FIPS score in an external, Chinese cohort of patients with preemptive TIPS implantation, but in that cohort the score had a poor discriminatory performance with c indices of 0.576 (95% CI 0.462–0.691) for 3-month survival and 0.574 (95% CI 0.462–0.639) for 6-month survival.

Moreover, only 5 of 290 patients (1.7%) had a FIPS score ≥0.92, which was the limit set by the authors to indicate a high risk of mortality. Such a small proportion suggests that the FIPS score may have limited clinical utility.

We evaluated the FIPS score in a cohort of 104 patients who received TIPS implantation at our centre between 2013 and 2018. This cohort, which was similar to the Freiburg cohort with regards to indication and liver disease aetiology, is described in detail elsewhere.[5] Only 5 of our 104 patients (4.8%) had a FIPS score ≥0.92 before the implantation. None of these patients died less than 6 months after TIPS insertion. We computed the discriminatory performance of the FIPS score measured by Harrell's c index using Stata's 'somersd' package. We similarly computed the discriminatory performance of the current model for end-stage liver disease (MELD)[6] and Child-Pugh scores in our cohort. The FIPS score had the poorest discrimination with c indices of 0.59 (95% CI 0.44–0.74) for 3-month survival and 0.57 (95% CI 0.43–0.72) for 6-month survival. The MELD score performed slightly better, while the Child-Pugh score reached c indices of 0.75 (95% CI 0.56–0.94) for 3-month survival and 0.72 (95% CI 0.54–0.90) for 6-month survival (Table 1). Thus, our findings indicate that the Child-Pugh score is superior to the FIPS score for prediction of 3- and 6-month mortality after TIPS insertion.

Our analyses do not invalidate the FIPS score, but they emphasize the importance of external validation of prediction models. Although our cohort is similar to the Freiburg cohort in the most relevant aspects (TIPS indication and liver disease aetiology), it differs in some ways, and these differences might explain the lower discriminatory performance of the FIPS score. In particular, creatinine was lower in our cohort and in the Chinese preemptive TIPS cohort than it was in the FIPS development cohort. However, the potential usefulness of the FIPS score depends on its ability to maintain discriminatory performance across different study populations. Of note, our cohort is small, and we await with great anticipation future validation studies assessing the utility of the novel FIPS score.

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.

Table 1. The discriminatory performance of the FIPS, MELD, and Child-Pugh scores in our external TIPS cohort measured by Harrell's c index.

	FIPS°c index (95% CI)	MELD°c index (95% CI)	Child-Pugh score c index (95% CI)
3-months survival	0.59 (0.44–0.74)	0.63 (0.43–0.84)	0.75 (0.56–0.94)
p values *vs.* FIPS	–	0.74	0.20
6-months survival	0.57 (0.43–0.72)	0.60 (0.41–0.79)	0.72 (0.54–0.90)
p values *vs.* FIPS	–	0.82	0.21

FIPS, Freiburg index of post-TIPS; MELD, model for end-stage liver disease; TIPS, transjugular intrahepatic portosystemic shunt.

Keywords: Transjugular intrahepatic portosystemic shunt; Liver cirrhosis; External score validation; Survival.
Received 15 April 2021; accepted 21 April 2021; available online 26 April 2021
https://doi.org/10.1016/j.jhep.2021.04.027

Authors' contributions

All authors contributed to the conception of the study and the final manuscript. FK carried out the formal analyses and wrote the original draft.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhep.2021.04.027.

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] Moons KG, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: II. External validation, model updating, and impact assessment. Heart 2012;98:691–698.
[3] Steyerberg E. Overfitting and optimism in prediction models. 2009. p. 83–100.
[4] Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. Eur Heart J 2014;35:1925–1931.
[5] Kraglund F, Jepsen P, Amanavicius N, Aagaard NK. Long-term effects and complications of the transjugular intrahepatic portosystemic shunt: a single-centre experience. Scand J Gastroenterol 2019;54:899–904.
[6] MELD policy changes 2016—UNOS. [cited 13/04/2021]; Available from: https://optn.transplant.hrsa.gov/news/meld-serum-sodium-policy-changes/.

Frederik Kraglund*
Rasmus Hvidbjerg Gantzel
Peter Jepsen
Niels Kristian Aagaard
Department of Hepatology and Gastroenterology, Aarhus University Hospital, Aarhus, Denmark
*Corresponding author. Address: Department of Hepatology & Gastroenterology, Aarhus University Hospital, DK-8200 Aarhus N, Denmark.
E-mail address: frekra@clin.au.dk (F. Kraglund)

Freiburg index of post-TIPS survival: The first score for individual prediction and a complementary tool for risk stratification

To the Editor:

We have read with great interest the impressive study conducted by Bettinger *et al.* about the development of the Freiburg index of post-TIPS survival (FIPS) score.[1] The authors established a prognostic model to achieve both individual outcome prediction (with the formula and the online calculator) and risk stratification (with a cut-off point), which had superior performance compared to previous prognostic models. Moreover, its ability to stratify patients at high and low risk remained robust in subgroups with different indications for transjugular intrahepatic portosystemic shunt (TIPS), different types of TIPS stents, well-preserved liver function, and impaired renal function.

Since the FIPS score is developed on the basis of a training cohort which consisted of German patients, we hoped to externally validate it and to assess its prognostic value within a Chinese TIPS cohort of 536 patients. In this cohort, the indication of TIPS included secondary prophylaxis of variceal rebleeding (n = 468, 87.3%), refractory ascites (n = 60, 11.2%) and early TIPS (n = 8, 1.5%). Among these patients, 109 (20.3%), 350 (65.3%), and 77 (14.4%) patients were graded as Child-Pugh A, B, and C, respectively, and the median model for end-stage liver disease (MELD) score was 11.5 (IQR 9.7–141). Major etiologies of chronic liver disease included HBV infection (292, 54.5%), autoimmune liver disease (53, 9.9%), and HCV infection (7.8%). Median follow-up was 23.7 (IQR 16.8–36.2) months.

We first tested the performance of FIPS score, Child-Pugh score,[2] MELD score,[3] and CLIF C-AD score.[4] Regarding the time-dependent AUROC for 3-month, 6-month, 1-year, 2-year, and 3-year death, Child-Pugh score had the best performance at all time points, whereas corresponding AUROCs of FIPS score were slightly lower (Fig. 1A, Table S1). Similarly, the C-indices for Child-Pugh score and FIPS score were 0.68 and 0.66, respectively (Table S1).

Subsequently, we validated the discriminative ability of FIPS score and Child-Pugh score. According to the original cut-off point of the FIPS score (0.92), only 7 patients were identified as high-risk patients, with significantly higher mortality than those with low risk (log-rank p <0.01, Fig. S1). Given the extremely low number of high-risk patients, we replaced the original cut-off with -0.006, the 85th percentile in the current cohort. For ease of use, we further modified it to 0 since only 1 patient had a FIPS score falling between -0.006 and 0, and the capability of risk stratification remained robust (log-rank p <0.01, Fig. 1B). Similarly, Child-Pugh score could also achieve adequate risk stratification with its grading system (log-rank p <0.01, Fig. 1C).

Finally, we investigated the risk stratification effect of the FIPS score in different Child-Pugh grades and *vice versa*. The FIPS score could significantly discriminate patients with high and low risk in Child-Pugh A and B subgroups (both with log-rank p <0.01, Fig. 1D, E), but not in the Child-Pugh C subgroup (log-rank p = 0.10, Fig. 1F). Inversely, in the low-risk subgroup defined by FIPS score, Child-Pugh score could stratify patients into grades A, B, and C (log-rank p <0.01, Fig. 1G). While in high-risk subgroup, Child-Pugh score did not reach statistical significance for risk stratification (log-rank p = 0.56, Fig. 1H).

In the original study, Bettinger *et al.* focused on the performance of FIPS score in predicting 3-month and 6-month survival, which proved satisfactory. Accordingly, we confirmed that FIPS score has similar AUROC for predicting 3-month and 6-month survival in this external validation cohort, and that patients with high and low risk

Keywords: transjugular intrahepatic portosystemic shunt; variceal bleeding; refractory ascites; individual outcome prediction; risk stratification.
Received 12 April 2021; received in revised form 26 April 2021; accepted 4 May 2021; available online 21 June 2021
https://doi.org/10.1016/j.jhep.2021.05.042