Improving MELD for use in acute liver failure

To the Editor:
Bechmann et al. [1] presented an interesting modification to the Model for End-Stage Liver Disease (MELD) score using M-65, a marker of hepatic necrosis, as a substitute for bilirubin in the MELD equation which they subsequently applied to a series of patients with acute liver failure (ALF). The statistical justification and reporting of the subsequent results is useful to discuss.

MELD is constructed (Eq. (1)) from serum creatinine, bilirubin, and internationalized normalized ratio (INR), and is a highly successful risk stratification marker for death in patients with cirrhosis. While it is used for prognostication in ALF in the USA, there is considerable doubt as to its applicability in this setting partly because of the lack of utility of bilirubin as an early marker of disease severity in ALF. MELD was constructed [2] from a Cox-proportional regression analysis of 231 patients with cirrhosis at the time of transjugular intrahepatic portosystemic shunt (TIPSS), where the quantitative predictors were logarithmically transformed to reduce the effect of influential outliers. The coefficients seen in the MELD equation are the regression coefficients derived from the computation using these transformed predictors.

\[
\text{MELD score} = 9.57 \ln(\text{Creatinine}) + 3.78 \ln(\text{Bilirubin}) + 11.2 \ln(\text{INR}) + 6.43
\]  

Creatinine – mg/L, Bilirubin – mg/dL, INR – unitless.

Bechmann et al.’s implantation of M-65 into MELD in this ALF cohort is concerning for a number of reasons. Their analysis is based on outcome alone and not the length of survival and therefore uses logistic regression (via Receiver Operating Characteristic (ROC) curves) not proportional hazards. In any multivariate analysis, the independence of predictors must be assessed and M-65, while having a clear univariate correlation with outcome, was not tested against INR and creatinine in a logistic or proportional hazards multivariate model. Furthermore, the effect of INR and creatinine on survival in ALF is likely to be different from cirrhosis; in fact, de novo regression modeling may give different regression coefficients that lead to a more useful MELD improvement rather than the assumption that the regression coefficients from patients with a different disease in a different country are immediately transferable.

Improvements in modeling when assessed by ROC curve analysis can be assessed quantitatively as was lacking in the Bechman et al. analysis. ROC curve comparisons via the Delong method [3] are non-parametric, well validated, and produce a p value for the comparison. While sensitivity and specificity improvements were discussed (again without quantitative comparators) and ROC curves were produced, there was no head to head graphical or quantitative ROC curve comparison between MELD and its modification or comparison among new multivariate models. These would have strengthened the analysis.

Statistical models are not necessarily interchangeable between diseases or populations, and repeat modeling is necessary to provide evidence of validity of the model. Statistical assessment of improvements of multivariate models is valuable and should be reported to allow the quantitation of the significance of any improvement.

Financial support
Mark McPhail is supported by the Wellcome Trust, UK.

Conflict of interest
Mark McPhail declared that he does not have anything to disclose regarding conflict of interest with respect to this letter.

References

Mark J.W. McPhail
Department of Hepatology & Gastroenterology, Imperial College London, Hammersmith Hospital Campus, Du Cane Road, London W12 0HS, United Kingdom
E-mail address: m.mcphail@imperial.ac.uk

Reply to the Letter to the Editor ‘Improving MELD for use in acute liver failure’

This is a reply to the Letter to the Editor by Dr. McPhail:
The authors thank Dr. McPhail for initiating this valuable discussion on our manuscript. We agree with Dr. McPhail that the MELD score was initially developed to stratify the risk of death in a cohort of patients with end-stage chronic liver disease to predict death after a TIPS procedure, as we pointed out in the introduction of the original manuscript [1,2]. We thus agree with Dr. McPhail that the MELD score was neither primarily designed...