

PRECISION OF APP-BASED MODEL FOR END-STAGE LIVER DISEASE SCORE CALCULATORS

Dr Simon Hews MBBS¹, Dr Perri Chambers MBBS¹

¹Department of Gastroenterology, Austin Health, Australia

Corresponding Author: simonhew@hotmail.com

Background: The prioritisation of patients with end-stage liver disease for liver transplantation requires a quantification of clinical disease severity. The Model for End-Stage Liver Disease (MELD) score is used to prognosticate survival for these patients and is therefore useful to prioritise for transplantation. The MELD score utilises a complex equation, which is now available for calculation using a range of smartphone applications ('apps'). There is however no published data on the precision of these app-based calculators in calculating a MELD score.

Methods: In a cohort of 46 adults patients awaiting liver transplantation, the precision of 14 free and pay-for-use Apple iPhone app-based MELD score calculators in calculating the MELD score was compared with the actual MELD equation using kappa statistics.

Results: Kappa statistics demonstrated agreement of 0.70 to 0.97 (mean of 0.78; 95% CI, 0.6 – 0.95) between the app-based calculators and the MELD equation.

Conclusion: This study showed substantial but not perfect precision of app-based MELD score calculators compared to the actual MELD equation. This is an important finding in assessing the validity of app-based MELD score calculators and further studies evaluating the growing number and availability of app-based medical calculators are required.

Journal MTM 1:4:11-15, 2012

DOI: 10.7309/jmtm.72

www.journalmtm.com

Introduction

End-stage liver disease is associated with a high morbidity and mortality¹. For selected patients with end-stage liver disease, liver transplantation is a suitable management option. With advances in surgical technique and post-transplant care, the survival rates after liver transplant continue to improve². Unfortunately, the scarcity of donor livers in Australia limits the availability of transplantation for patients with end-stage liver disease.

The allocation of livers to potential recipients requires a balance between clinical urgency and the likelihood of improved survival post-transplantation. A robust prognostic model is therefore necessary to assess disease severity and quantify the risk of

mortality without transplantation. The Model for End-Stage Liver Disease (MELD) score is a prospectively validated index used in end-stage liver disease for this purpose^{3, 4}. It has been adopted globally by many transplant programmes, including The Transplantation Society of Australia and New Zealand (TSANZ)⁵. The MELD score is determined from three laboratory values: the serum bilirubin, the serum creatinine and the international normalised ratio for prothrombin time (INR). These values form part of a logarithmic equation to produce a numerical score (see Figure 1). The resultant score estimates three-month survival, with a higher score reflective of a lower survival rate. The MELD score is therefore used in the prioritisation patients on the liver transplant waiting list and is calculated

periodically as donor livers become available or with change in the clinical state of potential recipients.

$$\text{MELD} = 10 \times [0.957 \times \text{Log}_e(\text{creatinine mg/dL}) + 0.378 \times \text{Log}_e(\text{bilirubin mg/dL}) + 1.12 \times \text{Log}_e(\text{INR}) + 0.643]$$

Figure 1. MELD score equation³. Note that the maximum serum creatinine for use in the MELD score equation is 4 mg/dL.

Information technology plays an integral role in the modern clinical environment. It is specifically used at transplant centres to coordinate pre-and post-transplant care, digitise radiology and maintain a database of patient records. With the increasingly widespread use of smartphones, such as the Apple iPhone (Apple Corporation, Cupertino, USA), in the medical workplace, clinicians can gain timely and portable access to reference materials, clinical tools and even patients' results⁶⁻⁸. There is an emerging volume of smartphone applications ('apps') that offer medical calculators that can be used at the bedside, however there is a paucity of studies assessing the accuracy of app-based medical calculators. Specifically, there is no published literature evaluating apps with MELD score calculators and their adherence to the actual MELD score equation. Our study aimed to assess the precision of the app-based MELD score calculators available for the iPhone by comparing the app calculated MELD scores with the actual, formula derived scores of a cohort of adult patients awaiting liver transplantation.

Methods

Subjects

A cohort of patients awaiting liver transplantation was selected in May 2012 from the transplant database of a tertiary hospital offering the sole statewide service for liver transplantation. All patients over the age of 18 were included in this study (n = 46). The most recent laboratory values for creatinine (mg/dL), bilirubin (mg/dL) and INR as of the 17th May 2012 were obtained. Clinical details for each patient were recorded and included the patient's sex, age and aetiology of liver disease.

Applications

App-based MELD score calculators (n=14) were obtained from the Apple iTunes Australia App Store (Version 10.6.1). Search terms for the apps were 'MELD score calculator', 'MELD score', 'liver

transplantation', 'liver disease', 'medical calculator', 'med calc' and 'medical application'. Both free and pay-for-use apps were included. Non-English apps were excluded from the study. Apps were downloaded and installed on an iPhone 4.

MELD Scores

A MELD score was calculated for each patient using the equation shown in Figure 1. This calculation was performed using Microsoft Excel 2008 for Mac (Version 12.3.4). MELD scores were subsequently calculated for each patient using each installed app on the iPhone.

Statistics

Statistical analyses were performed using a statistical package (Stata, Version 11, StataCorp LP, College Station, USA). A kappa statistic was used to evaluate the magnitude of agreement between the MELD scores of the cohort as calculated by the equation in Figure 1 ('gold standard') and the MELD scores as calculated by each app. A kappa statistic quantifies the agreement of two observations, adjusted for agreement by chance alone. Kappa is standardised on a scale from -1 to 1. A kappa of 1 indicates perfect agreement, a kappa of 0 indicates agreement expected by chance alone and a kappa of -1 indicates agreement occurring even less often than chance⁹. Therefore in this study, a higher kappa statistic reflected a higher degree of agreement between the scores derived from the actual MELD equation and the scores obtained from a particular app, thereby conferring a higher degree of precision to the app. It has been proposed that the strength of agreement quantified by a kappa statistic can be graded as < 0 = poor; 0.01 – 0.2 = slight, 0.21 – 0.4 = fair; 0.41-0.6 = moderate; 0.61 – 0.8 = substantial and 0.81 – 0.99 = almost perfect⁹.

Results

Patient Characteristics

There were 46 adult patients on the liver transplant waitlist database at the time of data collection (see Table 1). The patients had a mean age (± SD) of 53 ± 9.90 years and were comprised primarily of males (76%). The most common aetiology amongst the patients was hepatitis C (39%), followed by hepatocellular carcinoma (30%) and hepatitis B (15%).

Patient Characteristics (n = 46)	Values
Age (years, mean ± SD)	53 ± 9.90
Males (n)	35
Females (n)	11
Aetiology of liver disease (%)	
- Hepatitis C	39.13
- Hepatitis B	15.22
-Hepatocellular Carcinoma	30.43
- Alcohol	13.04
- Primary Biliary Cirrhosis	4.35
- Primary Sclerosing Cholangitis	10.87
- Non-Alcoholic Steato-Hepatitis	10.87
- Other†	15.22

Table 1. Patient Characteristics. † polycystic liver disease, haemangioendothelioma, cryptogenic, familial amyloidotic polyneuropathy

Apps

14 app-based MELD score calculators were downloaded and installed from iTunes (see Table 2). Five were published by individuals. The apps ranged in price from \$0 to \$5.49, with a mean price of \$2.93. Only one app's function was solely to calculate a MELD score (MELD Calculator); the other apps had an ability to calculate a variety of medical formulas, scores or classifications across a range of medical specialties. 11 of the 14 apps listed the MELD score equation utilised in the app.

MELD App	Publisher	Version	Cost (\$)
Advanced Medical Calculator	RapidStar LLC	1.0.0	2.99
ClinicCalc Medical Calculator	Medicon Apps	1.7	Free
Calculate (Medical Calculator) by QxMD	QxMD Medical Software	2.9.3	Free
ER-ICU Calculator	RapidStar LLC	1.0.0	0.99
GI Calculator	QxMD Medical Software	1.7	Free
Gastrointestinal Calc	RapidStar LLC	1.0.0	0.99
MedCalc	Mathias Tschopp and Pascal Pfiffner	2.6	Free
MedCalc Pro	Mathias Tschopp and Pascal Pfiffner	2.6	5.49
Medical Calculator	Marketwall.com	1.9	0.99
Medical Observer	UBM Medica	1.1.2	Free
MediMath Medical Calculator	Evan Schoenberg	4.1	5.49
MedSolve Medical Calculator	Charles Vu	1.2.1	0.99
Mediquations Medical Calculator	Mediquations	34.1	5.49
MELD Calculator	Marc L. Melcher	1.03	Free

Table 2. App-based MELD calculator

MELD Scores

The mean MELD score (\pm SD) using the equation in Figure 1 was 13.31 ± 7.69 (see Table 3).

MELD	Values (mean \pm SD)
Br (mg/dL)	3.76 ± 5.85
Cr (mg/dL)	1.11 ± 0.55
INR	1.47 ± 0.41
MELD score	13.31 ± 7.69

Table 3. MELD characteristics and score.

The range of MELD scores obtained using the MELD apps was 13.32 to 15.39 (see Table 4).

Kappa statistics for the apps ranged from 0.70 to 0.97, with a mean of 0.78 (95% CI, 0.6 – 0.95). Three apps, Medical Observer, MediMath Medical Calculator and Mediquations Medical Calculator, achieved a kappa of 0.97 and the app, MediSolve, achieved a kappa of 0.94. The remaining apps had a kappa score close to 0.70. The kappa statistics can be represented in Bland-Altman plots. Figure 2 demonstrates a Bland-Altman plot for the Medical Observer app. The close agreement between the app-based calculator and the MELD equation is represented by the narrow spread of points, yielding a high kappa statistic (0.97).

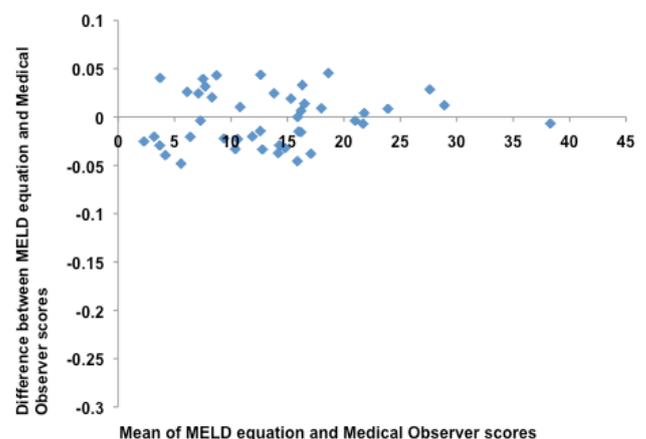


Figure 2. Bland-Altman plot of MELD equation and Medical Observer app MELD scores.

Figure 3 demonstrates a Bland-Altman plot for the GI Calculator app. The wider distribution of points reflects reduced agreement between the app-based calculator and the MELD equation. The corresponding kappa statistic (0.70) is therefore lower than the app represented in Figure 2.

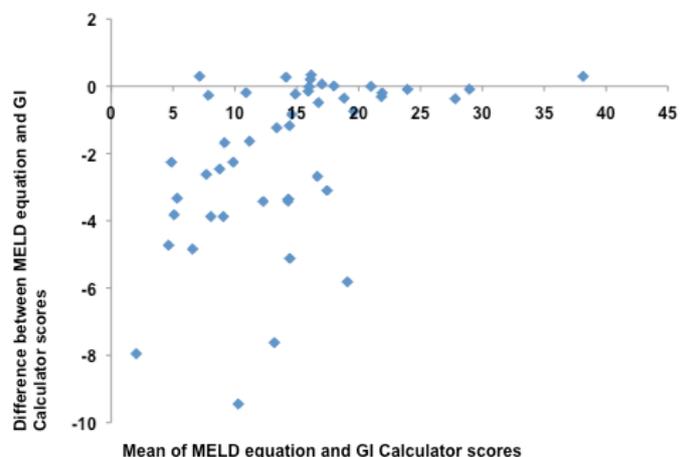


Figure 3. Bland-Altman plot of MELD equation and GI Calculator app MELD scores.

MELD App	Values (mean ± SD)	Kappa (95% CI)
Advanced Medical Calculator	15.38 ± 6.43	0.71
ClinicCalc Medical Calculator	15.35 ± 6.54	0.70
Calculate (Medical Calculator) by QxMD	15.37 ± 6.54	0.70
ER-ICU Calculator	15.38 ± 6.43	0.71
GI Calculator	15.37 ± 6.54	0.70
Gastrointestinal Calc	15.39 ± 6.45	0.71
MedCalc	15.36 ± 6.43	0.71
MedCalc Pro	15.35 ± 6.53	0.70
Medical Calculator	15.35 ± 6.53	0.70
Medical Observer	13.32 ± 7.70	0.97
MediMath Medical Calculator	13.32 ± 7.70	0.97
MedSolve Medical Calculator	13.33 ± 7.70	0.94
Mediquations Medical Calculator	13.35 ± 7.78	0.97
MELD Calculator	15.37 ± 6.54	0.70

Table 4. App-based MELD calculator scores and kappa statistics.

Discussion

This is the first study to evaluate the precision of app-based MELD score calculators available for the iPhone smartphone. Our key finding is that there exists a degree of disagreement between the scores derived from the apps studied and those generated from the actual MELD equation. All apps had a kappa statistics less than 1, signifying less than perfect agreement between the scores from the MELD equation and the scores from the apps in this cohort of patients. The difference in agreement may be accounted for by a discrepancy in the MELD equations employed by the app-based calculators, compared to the actual MELD equation. The MELD equation was listed for only 11 of the 14 app-based calculators but appeared identical to the correct equation. It is therefore possible that the mechanics or programming of the calculations performed by the app-based calculators was incorrect. Whilst the equation is complex, the calculation of the MELD score should be mathematically straightforward. Given these apps are publicly available through iTunes and eight required purchasing, this difference in scores between the actual MELD equation and the app-based calculators is an important finding not previously delineated.

Whilst not perfect, the strengths of the agreements of the app-based calculators could be subjectively classified as substantial to almost perfect¹⁰. The three apps with kappa statistics of 0.97 differed only slightly in their agreement with the MELD equation. As a collective group, the app-based calculators were in substantial agreement with the MELD

equation. Furthermore, the difference in the MELD scores as obtained from the actual equation and that from the app-based calculators is only a couple of points. Clinically, this minimal difference would unlikely have a significant effect in the prognostication of end-stage liver disease, with the corresponding decrease in three-month survival in the order of a few percent. Additionally, prioritising of patients with end-stage liver disease on the transplant waiting list takes into account a variety of factors. Whilst the MELD score is the predominant clinical component of prioritisation, other factors such as ABO blood type, time waitlisted and clinician discretion also contribute to the final ranking¹¹. This may temper the small difference in MELD scores demonstrated in the app-based calculators.

There are few studies assessing app-based medical calculators in the clinical environment. One study in the literature by Flannigan and McAloon evaluated the use of the PICU Calculator app in a paediatric resuscitation setting¹². It was found the app-based calculator was more accurate, faster and gave prescribers more confidence in prescribing than using the paper based British National Formulary for Children. The stated accuracy of the app was 100%. The ease of use and portability of app-based medical calculators makes their use in the day-to-day clinical environment of a liver transplant unit appealing. The rapid derivation of a clinical index, such as a MELD score, from a complex mathematical formula could facilitate bedside management and transplantation decisions. Additional studies evaluating app-based medical calculators are required.

The kappa statistics generated for the app-based MELD calculators is unique to this cohort of patients. Kappa values are therefore difficult to generalise across populations and comparison with kappa statistics from other medical calculators in differing settings cannot be performed. The kappa statistic also does not identify the specific point of disagreement between the MELD equation and the app-based calculators. It nevertheless is a powerful tool to assess precision.

Conclusion

The precision of apps which calculate a MELD score in patients with end-stage liver disease is substantial, though not perfect, when compared to the official MELD equation. This study provides a validation of the precision of app-based MELD calculators which have been thus far been available without assessment. Of the 14 apps evaluated, the apps Medical Observer, MediMath Medical Calculator and MediQuations Medical Calculator yielded the highest kappa score for the studied population of patients waitlisted for liver transplantation. The use of smartphone medical calculators offers a useful tool for the liver transplant physician and clinicians in general and further studies examining other clinical calculators are required.

Acknowledgements

The authors would like to acknowledge Anastasia Hutchinson (PhD) for her statistical consultancy.

References

1. D'Amico G, Garcia-Tsao G, Pagliaro, L. Natural history and prognostic indicators of survival in cirrhosis a systematic review of 118 studies. *Journal of Hepatology* 2006; 44:217-231.
2. Yoo HY, Pratt CH, Geschwind JF, et al. The outcome of liver transplantation in patients with hepatocellular carcinoma in the United States between 1988 and 2001: 5-year survival has improved significantly with time. *J Clin Oncol* 2003; 21(23): 4329.
3. Kamath PS, Wiesner RH, Malinchoc M, et al. A model to predict survival in patients with end-stage liver disease. *Hepatology* 2001; 33:464-470.
4. Leise MD, Kim WR, Kremers WK, et al. A revised model of end-stage liver disease optimizes prediction of mortality among patients awaiting liver transplantation. *Gastroenterology* 2011; 140:1952-1960.
5. The Transplantation Society of Australia and New Zealand. Organ transplantation from deceased donors: consensus statement on eligibility criteria and allocation protocols version 1.1, <http://www.tsanz.com.au/downloads/201123June-TSANZConsensusStatementVs1.1.pdf> (accessed September 2012)
6. León SA, Fontelo P, Green L, et al. Evidence based medicine among internal medicine residents in a community hospital program using smart phones. *BMC Med Informatics and Decision Making* published online February 2007. doi: 10.1186/1472-6947-7-5.
7. Franko OI, Tirrell TF. Smartphone app use among medical providers in ACGME training programs. *J Med Syst* 2011; doi:10.1007/s10916-011-9798-7.
8. Payne KF, Wharrad H, Watts K. Smartphone and medical related app use among medical students and junior doctors in the United Kingdom: a regional survey. *BMC Medical Informatics and Decision Making* published online Oct 2012. doi:10.1186/1472-6947-12-1212.
9. Cohen J. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 1960; 20: 37-46.
10. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33:159-174.
11. US Department of Health and Human Services, Organ Procurement and Transplant Network. Organ distribution: allocation of livers, http://www.optn.transplant.hrsa.gov/PoliciesandBylaws2/policies/pdf/policy_8.pdf (accessed October 2012)
12. Flannigan C, McAloon J. Students prescribing emergency drug infusions utilising smartphone outperform consultants using BNFCs. *Resuscitation* 2011;82: 1424-1427