Published in final edited form as:

Gastroenterology. 2012 November ; 143(5): 1237–1243. doi:10.1053/j.gastro.2012.07.113.

Development of an Accurate Index for Predicting Outcomes of Patients with Acute Liver Failure

Anna Rutherford^{1,3,*}, Lindsay Y. King^{1,2,*}, Linda S. Hynan⁴, Chetan Vedvyas⁵, Wenyu Lin², William M. Lee⁶, and Raymond T. Chung² for the ALF Study Group

¹Department of Internal Medicine, Brigham & Women's Hospital, Boston, Massachusetts

²Gastrointestinal Unit, Massachusetts General Hospital, Boston, Massachusetts

³Division of Gastroenterology, Hepatology & Endoscopy, Brigham & Women's Hospital, Boston, Massachusetts

⁴Departments of Clinical Sciences (Division of Biostatistics) and Psychiatry, University of Texas Southwestern Medical Center, Dallas, Texas

^{2,5}Harvard Medical School, Boston, Massachusetts

⁶Division of Digestive and Liver Diseases, University of Texas Southwestern Medical Center, Dallas, Texas

Abstract

Background & Aims—Patients with acute liver failure (ALF) have high mortality and frequently require liver transplantation (LT); few reliable prognostic markers are available. Levels of M30, a cleavage product of cytokeratin-18 caspase, are significantly increased in serum samples from patients with ALF who die or undergo LT. We developed a prognostic index for ALF based on level of M30 and commonly measured clinical variables (called the ALFSG index), and compared its accuracy with that of the King's College criteria (KCC) and model for end stage liver disease (MELD). We also validated our model in an independent group of patients with ALF.

Methods—Serum levels of M30 and M65 antigen (the total cytokeratin-18 fragment, a marker of apoptosis and necrosis) were measured on 3 of the first 4 days following admission of 250 patients with ALF. Logistic regression was used to determine if the following factors, measured on day 1, were associated with LT or death: age, etiology; coma grade; international normalized ratio (INR); serum pH; body mass index; levels of creatinine, bilirubin, phosphorus, arterial ammonia, and

© 2012 The American Gastroenterological Association. Published by Elsevier Inc. All rights reserved.

Corresponding Author: Raymond T. Chung, MD, Gastrointestinal Unit, Massachusetts General Hospital, Warren 10, Boston, MA 02114, RTChung@partners.org, 617-724-7562.

*Drs. Lindsay King and Anna Rutherford are co-first authors

None of the authors have any conflicts of interest to declare

Author involvement:

Study Concept & Design: Anna Rutherford, Linda Hynan, Raymond Chung, William Lee

Acquisition of Data: Lindsay King, Chetan Vedyvas, Wenyu Lin

Analysis & Interpretation of Data: Anna Rutherford, Lindsay King, Linda Hynan, Raymond Chung

Technical Support: Chetan Vedyvas, Wenyu Lin

Drafting of Manuscript: Anna Rutherford, Lindsay King, Linda Hynan, William Lee, Raymond Chung

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

lactate; and $\log_{10}$ M30 and $\log_{10}$ M65. The area under the receiver operating characteristic (AUROC) was calculated for the ALFSG and other indices.

Results—Coma grade, INR, levels of bilirubin and phosphorus, and $\log_{10}$ M30 value at study entry most accurately identified patients that would require LT or die. The ALFSG index identified these patients with 85.6% sensitivity and 64.7% specificity. Based on comparison of AUROC values, the ALFSG Index (AUROC 0.822) better identified patients most likely to require LT or die than the KCC (AUROC 0.654) or MELD (AUROC 0.704) ($P=.0002$ and $P=.0010$, respectively). We validated these findings in a separate group of 250 patients with ALF.

Conclusions—The ALFSG Index, a combination of clinical markers and measurements of the apoptosis biomarker M30, better predicts outcomes of patients with ALF than the KCC or MELD.

ClinicalTrials.gov NCT00518440

Keywords

prognostic factor; cell death marker; liver disease progression; risk factor

INTRODUCTION

Acute liver failure (ALF) is characterized by sudden loss of hepatic function in people without underlying liver disease. ALF affects 2,000 people/year in the United States from a variety of etiologies including acetaminophen toxicity, viral hepatitis, drug induced liver injury, and indeterminate causes. Only 45% of patients with ALF survive without liver transplantation (1). Because of the rapid progression of ALF and organ shortage, there is a need for improved predictive models for outcome.

The King's College criteria (KCC) is the most widely used prediction model for outcome in ALF (2). However, recent studies have shown that despite the high positive predictive value for poor outcome, the KCC has reduced sensitivity and negative predictive value, with a significant number of patients dying without meeting KCC (3,4). Other prognostic markers, including MELD, coma grade, bilirubin, etiology of acute liver failure (1), systemic inflammatory response syndrome (5), serum Gc-globulin (6), arterial blood lactate (7), phosphorous (8), arterial blood ammonia(9), alpha fetoprotein (10), Factor V levels (11) and body mass index (12) have been identified, but are not used routinely in clinical practice.

ALF is characterized by widespread hepatocyte death in excess of regeneration. Hepatocyte death typically occurs through either apoptosis or necrosis. There is increasing evidence that apoptotic cell death plays a significant role in acute liver failure. In apoptotic cell death, stimuli such as Fas ligand, TNF alpha and DNA damage activate caspases, cysteine proteases that cleave structural proteins and proteins involved in DNA synthesis and repair (13). Animal studies have shown that apoptosis plays a key role in acetaminophen induced liver injury (14), viral liver disease(15), alcoholic hepatitis (16) and Wilson disease (17).

Bantel et al described M30, which selectively recognizes a caspase cleaved neopeptide of cytokeratin 18, indicative of apoptotic hepatocyte cell death (18). Serum caspase activity was found to be a more sensitive method of detecting early liver injury than measurement of ALT (18). In a pilot study, Rutherford et al measured levels of apoptotic markers in the serum of 67 patients with ALF (19). They found that serum M30 levels were ten-fold greater in patients with acute liver failure than in chronic HCV or normal controls. It was also noted that median M30 levels were significantly higher in patients who underwent liver transplantation or died compared to transplant free survivors, suggesting that measurement of serum M30 levels may be able to predict outcome in acute liver failure.

There has been further investigation of apoptosis and necrosis markers as predictors of outcome in ALF. Volkmann et al found that those patients with ALF who survived without liver transplantation had higher levels of apoptosis and caspase activation as compared to greater levels of levels of DNA fragmentation and non-apoptotic cell death in those who died or required transplantation (20). Bechmann et al found that a MELD score supplemented with M-65 antigen (total cytokeratin 18 fragment, a marker of all-cause cell-death including apoptosis and necrosis) more sensitively and specifically predicted outcome in ALF patients than did MELD or KCC (21). M30 (apoptosis) and M65 (apoptosis + necrosis) levels have been reproducibly shown to quantitatively reflect functional and therefore damaged liver cell mass in ALF (22). It has not been determined whether some ratio, quantity, inverse relationship or trajectory of apoptotic and necrotic cell death in ALF is more predictive of outcome or hepatic regenerative capacity.

In this study, we sought to expand our pilot study (19) to a larger group of 250 patients with ALF to confirm the predictive value of markers of cell death in ALF patients. We measured serum levels of M30 and M65 antigen in 125 ALF patients who survived without transplant and 125 ALF patients who died or underwent liver transplantation. We hypothesized that M30 and M65 would be significantly elevated in patients who died or underwent liver transplantation compared to transplant free survivors, as suggested by the degree of cell death these markers help to quantify. We then combined these cell death markers with clinical predictors of outcome in acute liver failure to generate a clinical prediction rule for outcome in ALF. Finally, we validated our clinical prediction rule in an independent set of 250 patients with ALF.

PATIENTS AND METHODS

Patients

In creating our derivation set, we investigated sera from 250 ALF patients, randomly selected to include equal numbers of those who survived without liver transplant (125) or those who were transplanted or died (125) on the basis of availability of serum samples from the Acute Liver Failure Study Group (ALFSG) data and serum bank. Our validation set was a different 250 ALFSG patients, again selected to include approximately equal groups (122 patients who survived without transplant and 128 who were transplanted or died). All investigators were blinded to diagnosis and outcome.

All samples were stored at -80° within 2 hours of collection. Etiology of ALF was determined by the site investigator using standard criteria. Since 1998 the ALFSG has prospectively collected demographic, clinical, laboratory, and outcome data and serum on all subjects meeting entry criteria for ALF at 23 U.S. centers.

Eligible patients had an international normalized ratio (INR) >1.5 , hepatic encephalopathy, and presented within 26 weeks of illness onset without apparent chronic liver disease. Because subjects were encephalopathic, written informed consent was obtained from next of kin. Outcomes were defined as liver transplantation, discharge, or death three weeks after admission (1). The study was performed according to the institutional review board guidelines of each of the 23 original centers of the ALFSG and under approval of the ALFSG ancillary studies committee.

Serum Cell Death Marker Assays

For quantitative detection of M-30 and M-65, we used M-30 Apoptosense ELISA and M-65 ELISA (PEVIVA; Alexis, Gruwald, Germany). The minimal detectable level for M-30 was 25 U/L, and for M-65 11 U/L.

For the derivation set, serum levels of M-30 and M-65 were measured on 3 of the first 4 days of enrollment into ALFSG. Assays were performed on 166 day 1 samples, 248 day 2 samples, 242 day 3 samples and 93 day 4 samples. Sample availability determined the days of samples. Assays were performed in duplicate. Initial dilutions for M-30 were 7ul sample to 75ul solution (or 1:11.7). Further dilutions of 1:19.75 and 1:26 were performed as necessary. Initial dilution for M-65 was 1:19.75 with further dilutions of 1:26 and 1:38.5 as necessary. A mean value was calculated for each sample, with excellent reproducibility of assays. The ICC for the M-30 assay was 0.997 ($p < 0.0001$), with a value of 1.0 representing a perfect score. The ICC for the M-65 assay was 0.998 ($p < 0.0001$).

For the validation set, M-30 was measured on one of the first two days of enrollment into ALFSG in 124 transplant free survivors and 126 transplant/died patients. We had 194 day 1 and 56 day 2 samples available. Assays were performed in duplicate. Initial dilutions for M-30 were 7ul sample to 75ul solution (or 1:11.7). Further dilutions were performed as necessary.

Statistical Analysis

Statistical analyses were conducted using IBM® SPSS Statistics V19 (SPSS, Inc., IBM® SPSS Statistics V19, Chicago, IL, 2010) and SAS V9.2 (SAS Institute, Inc. SAS V9.2, Cary, NC, 2008). Intraclass correlation (ICC) was calculated for both M-30 and M-65 assays. Scores for KCC and MELD were calculated for all 500 participants after performing the analysis for M-30 and M-65. Baseline demographic characteristics for categorical measures were described using N (%) and groups (derivation and validation sets) were compared using χ^2 ; continuous measures were described using median [range] and groups were compared using the Mann-Whitney U tests. Stepwise logistic regression models predicting the need for liver transplant or death (versus transplant free survival) were fit to the derivation set; potential measures included in this prediction model were entry M-30 level, entry $\log_{10}$ M-30 level, entry M-65 level, entry $\log_{10}$ M-65 level, ratio entry M-65/M-30 levels, entry M65 level – entry M30 level, $\log_{10}$ entry M65 level – entry M30 level, and changes in day 1, day 2 and day 3 M-30 level or M-65 level. Other potential clinical measures used as predictors in these models included age (both continuous and dichotomized: <40 versus > 40 years), gender, etiology groupings (acetaminophen versus others; acetaminophen + drug induced hepatitis versus others; acetaminophen + hepatitis A + Shock versus others), coma grade (both as 4 levels and dichotomized: I–II versus III–IV), creatinine, bilirubin, INR, phosphorus (< 3.7 versus > 3.7), ALT, pH, arterial ammonia, lactate, BMI, and sodium. Several measures were eliminated from consideration in the model because of a large number of missing values (arterial ammonia, lactate, pH, and BMI). The criteria for entry into the model was set to $p = 0.05$ and for leaving the model was $p = 0.10$. The fit of the model to the data was examined using the Hosmer-Lemeshow p value with the criteria of a good model defined as $p > 0.40$. Longitudinal logistic regression was also used to predict liver transplant (versus no transplant) or death (versus alive) when using M-30 or M-65 measures over time and initial clinical characteristics; however, none of these models were predictive of liver transplant or death when using the M-30 or M65 measures over time. The best fitting logistic regression model from the derivation data set (ALFSG Index), KCC, and MELD were analyzed using ROC analysis and comparisons of area under the curve between these three measures were performed (22). Using the results of the ROC analyses, the best threshold for group predictions was determined using a combination of criteria: sensitivity, specificity, accuracy, and maximum perpendicular distance above the 45° line of equality (23).

The validity of the ALFSG Index was evaluated in the validation data set. The weights from the best logistic regression model from the derivation data set were applied to the validation data set. ROC analysis was again used to compare predictions of liver transplant or death for

the ALFSG Index, KCC, and MELD and statistical comparisons of AUC for each of these models was again performed.

Assumptions of all statistical tests were reviewed and transformation of variables were performed where necessary. Statistical significance was set at $p < 0.05$ unless otherwise specified. All authors had access to the study data and reviewed and approved the final manuscript.

RESULTS

Patient Population

Table 1 depicts the entry demographic characteristics of the 250 ALF patients in the derivation set in comparison to the 250 ALF patients in the validation set. The groups were selected only to ensure equal numbers of patients who were transplanted and/or died from ALF (125) versus those who survived without a liver transplant (125). Other characteristics including age, gender, etiology of ALF, entry laboratory values, entry coma grade, entry KCC score and MELD score were not revealed to the investigators until the analysis of M-30 and M-65 samples was completed. The groups differed significantly when comparing etiologies of ALF ($p = 0.0003$): most notably the derivation set had 30% acetaminophen versus 48.4% in the validation set. In addition, the derivation set had significantly higher bilirubin levels than the validation set (11.8 vs. 6.8, $P = 0.0028$). This difference is likely to be explained by the greater numbers of drug-induced hepatitis and indeterminate cases, which tend to have higher admission bilirubin levels than acetaminophen cases (1), randomly included in the derivation set. The groups also differed in entry coma grade (derivation set (30% with grade II, 20.4% grade IV) versus validation set (19.7% grade II, 30.9% grade IV, $p = 0.0132$)), and median M-30 level (2726 versus 2139, $p < 0.0001$). The groups were similar in terms of age, gender, median entry creatinine, INR, phosphorous, ALT, sodium, MELD and in the proportion of patients predicted to need liver transplantation based on KCC.

The ALFSG Index

Table 2 represents the final model created using stepwise logistic regression which predicted need for liver transplantation or death in ALF using coma grade, bilirubin, INR, phosphorus and $\log_{10}$ M-30; the Hosmer-Lemeshow $p = 0.633$ indicated a good fit of the model to the data. The area under the ROC curve was 0.822. The odds ratio for transplant or death for a one-log increase in entry M-30 was 3.331. Using a threshold value of 0.4285, the model has a sensitivity of 85.6%, specificity of 64.7% and an accuracy of 75.7%.

Patients without complete information for all covariables included in the multivariable models were excluded from these analyses: a maximum of 38 patients in the derivation set (15.2%) and 37 patients in the validation set (14.8%).

The ALFSG Index Outperforms KCC and MELD

After entry KCC and MELD scores were calculated, the ROC curves for the ALFSG Index versus KCC versus MELD score in predicting liver transplant or death in the same patients were compared. Figure 1 displays the ROC curves for the ALFSG Index (AUROC 0.822) against KCC (AUROC 0.654) and MELD (AUROC 0.704). This comparison showed a statistically significant advantage for the ALFSG Index ($p = 0.002$, $p = 0.0010$, respectively).

To confirm the validity of this model, we repeated this analysis in our validation set. Figure 2 displays the ROC curves for the ALFSG Index (AUROC 0.839) against KCC (AUROC 0.684) and MELD (AUROC 0.717) in predicting liver transplant or death for patients in the

validation set. We again confirmed a statistically significant advantage for the ALFSG Index in its ability to predict outcome in ALF compared to KCC ($p=0.003$) or MELD ($p=0.0005$). Using a threshold value of 0.4285, the model has a sensitivity of 81.13%, specificity of 72.04% and accuracy of 76.5%.

DISCUSSION

Although many clinical and investigational variables have been shown to predict outcome or need for liver transplantation in ALF, few have succeeded in fulfilling the basic requirements of a prognostic model in ALF, which requires objectivity, early applicability and accuracy. The KCC and MELD score have been repeatedly shown to be reasonably sensitive, but limited in terms of specificity in predicting early outcome (25). Other studies have looked at the addition of further clinical variables to KCC or MELD to strengthen the potential to predict outcome in ALF, such as M-65 incorporated into MELD (21) in 68 patients with ALF due to acute viral hepatitis B, drug-induced liver failure and congestive heart failure, with AUROC 0.870 (95% CI 0.722–0.933, $p<0.001$). Another study in patients with ALF due solely to acute viral hepatitis, found that within this patient population and etiology the combination of any 3/6 clinical parameters: age >50 , jaundice to encephalopathy time greater than 7 days, grade 3–4 encephalopathy, cerebral edema, prothrombin time >35 seconds and serum creatinine >1.5 ng/dL, was more predictive of outcome in ALF than KCC or MELD, with comparative AUROC of 0.821 versus MELD (AUROC 0.717) and KCC (AUROC 0.676) (26). Both of these studies were limited in terms of number of patients studied, variability of etiologies and differences in outcome.

The current study analyzed the largest number of patients with ALF ($n=500$), with variable etiologies and outcomes, using clinical entry variables in addition to quantitative measurements of the degree of hepatocyte apoptosis and necrosis previously proven to be independently predictive of outcome in ALF (19). The resulting amalgam is the ALFSG Index: a combination of entry level coma grade, bilirubin, INR, phosphorus (all common, easily and objectively measured variables) and entry level $\log_{10}$ M-30 (objectively measured with ELISA). This model has a sensitivity of 85.6% and a specificity of 64.7% in predicting likelihood of death or need for liver transplantation in ALF, and AUROC of 0.822 in comparison to MELD (0.704) and KCC (0.654). Even with some variability between our randomly selected derivation set and validation sets, including significant differences in etiologies, entry bilirubin, entry coma grade and entry $\log_{10}$ M-30, the superiority of the ALFSG Index was convincingly validated in a further 250 patients. We feel that the superior sensitivity of the ALFSG-Index is particularly useful, because it is conceived as an early screening test to determine who is likely to die from ALF and should therefore undertake early liver transplant evaluation.

In keeping with the bipartite nature of KCC, with separate prediction rules for need for liver transplantation based on acetaminophen or non-acetaminophen etiology, we examined the ALFSG Index performance within the distinct groups of study patients with ALF due to acetaminophen and non-acetaminophen. The sensitivity, specificity and accuracy of the ALFSG Index in acetaminophen ALF (80.6% sensitivity, 78.1% specificity, 79.1% accuracy), non-acetaminophen ALF (84.7% sensitivity, 59.2% specificity, 74.6% accuracy) and in the group overall (83.4% sensitivity, 69% specificity, 76.4% accuracy) was very similar, unlike the more substantial variability in sensitivity, specificity and accuracy of KCC in acetaminophen ALF (73.8% sensitivity, 65.5% specificity, 68.9% accuracy), non-acetaminophen ALF (57.8% sensitivity, 81.7% specificity, 68.1% accuracy) and overall (62.8% sensitivity, 74.1% specificity 62.8%, 68.4% accuracy). To simplify our model, and because distinction among acetaminophen and non-acetaminophen etiologies of ALF was statistically insignificant for the ALFSG Index, we have shown that the ALFSG Index has

improved sensitivity, specificity and accuracy for predicting need for liver transplantation or death in all etiologies of ALF.

With the ability to serially measure both an apoptosis marker (M-30) and a marker of total hepatocyte death (M-65) in 250 ALF patients on 3 consecutive days, early in their illness, we were able to closely examine the dynamics of these measurements as they relate to the contributions of apoptosis and necrosis in the cell death of ALF. It remains unclear whether apoptosis or necrosis plays a more pivotal role in the cell death of ALF, whether the mode of cell death differs by etiology of ALF, or whether there is a balance between the two forms of cell death that might ultimately determine outcome (25). In initial analysis of our derivation set we examined entry M-30 level, entry $\log_{10}$ M-30 level, entry M-65 level, entry $\log_{10}$ M-65 level, ratio entry M-65/M-30 levels, entry M65 level – entry M30 level, and changes in day 1, 2 and 3 M-30 level or M-65 level to see if any of these configurations were more effective in predicting outcome in ALF. As noted in our final model, the single most useful cell death marker was entry $\log_{10}$ M-30 level, further suggesting that the absolute quantity of hepatocyte apoptosis in ALF determines whether one survives ALF without need for LT. The relevance of M-30 antigen provides further confirmation of our pilot data (19), and affirms the potential therapeutic relevance of caspase inhibitors in ALF. The variability of the other configurations of M-30 and M-65 within different etiologies and outcomes, even within a large a group of patients with ALF, suggests that hepatocyte necrosis plays a less important role in dictating outcome.

Limitations in this study include imperfect matching in our randomized derivation and validation sets, in terms of etiologies and some entry-level criteria. This was attributable in part due to our pre-established requirement for equal numbers of patients with each outcome (death/transplant and transplant-free survival), in order to guarantee that our model could predict outcome. Despite the differences in etiology noted above, the groups were well-matched in terms of MELD score and KCC criteria. Also, despite the imperfect matching, the ALFSG Index worked equally well in predicting outcome in both derivation and validation sets (Figures 1, 2). There was also variability in the days of serum collection. As noted above, in both the derivation and validation sets, 66% and 77% respectively, had Day 1 samples available.

The ALFSG Index meets criteria for outcome prediction in ALF because of its early applicability and accuracy. Unlike KCC and MELD score, the ALFSG Index includes coma grade, which is a more subjective variable than each aspect of the MELD or KCC. Coma grade is based on simple, standardized definitions of the 4 stages of hepatic encephalopathy used uniformly in our centers; however, we acknowledge that the variability of this grading system may be somewhat greater in clinical practice.

We have shown that log M30 is a significant addition to the logistic regression model ($p=0.019$), and that its inclusion in the ALFSG Index improves our AUROC, thereby enhancing the predictive ability of our model. While measurement of the M-30 antigen is a non-standard assay, it is an ELISA-based test that can be readily established at transplant centers given its demonstrated enhancement of current prediction rules in ALF. As we have discussed, serial measurements of M-30 were not found to be helpful in predicting outcome in ALF, so the ALFSG-Index would require a simple one-time M-30 measurement early in any admission of a patient with ALF. Given its simplicity and reproducibility, this assay could be performed within the same time frame as other more standard serologic tests in critically ill patients.

Prediction rules such as KCC and MELD, relying strictly on early clinical variables, have reached a ceiling for accuracy. In contrast, the ALFSG Index has shown that the inclusion of

a marker of hepatocyte apoptotic cell death, a biomarker reflecting the underlying pathophysiology, extends this ceiling. The improvement in accuracy of predicting outcome in ALF (and its impact on organ allocation) would appear to outweigh the inconvenience of an additional ELISA assay. Further study is warranted to quantify the impact of this improved accuracy on cost effectiveness.

In summary, the ALFSG Index, a combination of 4 easily obtainable clinical markers, including admission coma grade, INR, bilirubin, and phosphorus, together with the novel serum-based biomarker M30, more accurately predicts outcome early in the course of ALF than previous well established criteria. The ALFSG Index is broadly applicable across etiologies of ALF. The contribution of a marker of hepatocyte apoptosis to outcome supports consideration of caspase inhibitors to treat ALF. Strong consideration should be given to using M-30 antigen to evaluate patients with ALF.

Acknowledgments

The work was supported by U01DK58369 and R21DK077716

References

1. Lee WM, Squires RH, Nyberg SL, et al. Acute Liver Failure: Summary of a Workshop. *Hepatology*. 2008; 47:1401–1415. [PubMed: 18318440]
2. O'Grady JG, Alexander GJ, Hayllar KM, et al. Early indicators of prognosis in fulminant hepatic failure. *Gastroenterology*. 1989; 97(2):439–45. [PubMed: 2490426]
3. Simpson KJ, Bates CM, Henderson NC, et al. The utilization of liver transplantation in the management of acute liver failure: comparison between acetaminophen and non-acetaminophen etiologies. *Liver Transplantation*. 2009; 15(6):600–9. [PubMed: 19479803]
4. Shakil AO, Kramer D, Mazariegos GV, et al. Acute liver failure: clinical features, outcome analysis, and applicability of prognostic criteria. *Liver Transplantation*. 2000; 6(2):163–9. [PubMed: 10719014]
5. Miyake Y, Sakaguchi K, Iwasaki Y, et al. New prognostic scoring model for liver transplantation in patients with non-acetaminophen-related fulminant hepatic failure. *Transplantation*. 2005; 80(7):930–6. [PubMed: 16249741]
6. Schiodt FV, Rossaro L, Stravitz RT, et al. Gc-globulin and prognosis in acute liver failure. *Liver Transpl*. 2005; 11(10):1223–7. [PubMed: 16184570]
7. Bernal W, Donaldson N, Wyncoll D, et al. Blood lactate as an early predictor of outcome in paracetamol-induced acute liver failure: a cohort study. *Lancet*. 2002; 359(9306):558–63. [PubMed: 11867109]
8. Baquerizo A, Anselmo D, Shackleton C, et al. Phosphorus as an early predictive factor in patients with acute liver failure. *Transplantation*. 2003; 75(12):2007–14. [PubMed: 12829902]
9. Bhatia V, Singh R, Acharya SK. Predictive value of arterial ammonia for complications and outcome in acute liver failure. *Gut*. 2006; 55(1):98–104. [PubMed: 16024550]
10. Schiodt FV, Ostapowicz G, Murray N, et al. Alpha-fetoprotein and prognosis in acute liver failure. *Liver Transpl*. 2006; 12(12):1776–81. [PubMed: 17133565]
11. Elinav E, Ben-Dov I, Hai-Am E, et al. The predictive value of admission and follow up factor V and Vii levels in patients with acute hepatitis and coagulopathy. *J Hepatology*. 2005; 42:82–86.
12. Rutherford A, Davern T, Hay JE, et al. Influence of high body mass index on outcome in acute liver failure. *Clin Gastroenterol Hepatol*. 2006; 4(12):1544–9. [PubMed: 16996806]
13. Malhi H, Gores GJ, Lemasters JJ. Apoptosis and necrosis in the liver: a tale of two deaths? *Hepatology*. 2006; 43(2 Suppl 1):S31–44. [PubMed: 16447272]
14. El-Hassan H, Anwar K, Macanas-Pirard P, et al. Involvement of mitochondria in acetaminophen-induced apoptosis and hepatic injury: roles of cytochrome c, Bax, Bid, and caspases. *Toxicol Appl Pharmacol*. 2003; 191(2):118–29. [PubMed: 12946648]

15. Song le H, Binh VQ, Duy DN, et al. Variations in the serum concentrations of soluble Fas and soluble Fas ligand in Vietnamese patients infected with hepatitis B virus. *J Med Virol.* 2004; 73(2):244–9. [PubMed: 15122799]
16. Tagami A, Ohnishi H, Moriwaki H, et al. Fas-mediated apoptosis in acute alcoholic hepatitis. *Hepatogastroenterology.* 2003; 50(50):443–8. [PubMed: 12749243]
17. Strand S, Hofmann WJ, Grambihler A, et al. Hepatic failure and liver cell damage in acute Wilson's disease involve CD95 (APO-1/Fas) mediated apoptosis. *Nat Med.* 1998; 4(5):588–93. [PubMed: 9585233]
18. Bantel H, Luger A, Heidemann J, et al. Detection of apoptotic caspase activation in sera from patients with chronic HCV infection is associated with fibrotic liver injury. *Hepatology.* 2004; 40(5):1078–87. [PubMed: 15486927]
19. Rutherford AE, Hynan LS, Borges CB, et al. Serum apoptosis markers in acute liver failure: a pilot study. *Clin Gastroenterol Hepatol.* 2007; 5(12):1477–83. [PubMed: 17967565]
20. Volkmann X, Anstaett M, Haden J, et al. Caspase Activation is Associated with spontaneous recovery from ALF. *Hepatology.* 2008; 47 (5):1624–33. [PubMed: 18393389]
21. Bechmann LP, Jochum C, Kocabayaglo P, et al. Cytokeratin-18 based modification of MELD score improves prediction of spontaneous survival after acute liver injury. *J Hepatology.* 2010; 53 (4):593–5.
22. Bechmann LP, Marquitan G, Jochum C, et al. Apoptosis versus necrosis rate as predictor for ALF following ACM intoxication compared to acute-on-chronic liver failure. *Liver International.* 2008; 28 (5):713–716. [PubMed: 18433398]
23. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operation characteristic curved derived from the same cases. *Radiology.* 1983; 148:839–843. [PubMed: 6878708]
24. Riffenburgh, RH. *Statistics in Medicine.* 2. Elsevier; New York, NY: 2006.
25. Polson J. Assessment of Prognosis in Acute Liver Failure. *Seminars of Liver Disease.* 2008; 28:218–225.
26. Dhiman RK, Jain S, Maheshwari, et al. Early Indicators of Prognosis in Fulminant Hepatic Failure: An Assessment of the Model for End-Stage Liver Disease (MELD) and King's College Hospital Criteria. *Liver Transplantation.* 2007; 13:814–821. [PubMed: 17370333]

Figure 1.

ROC curves for our model, the ALSFG index, MELD and King's College Criteria in Derivation Set

Comparison of ROC curves for our model, the ALFSG Index (AUROC 0.822), against KCC (AUROC 0.654) and MELD (AUROC 0.704) in predicting LT or death in the same patients showed a statistically significant benefit to the ALFSG Index ($p=0.0002$, $p=0.0010$, respectively)

Figure 2.

ROC curves for our model, the ALSFG index, MELD and King's College Criteria in Validation Set

Comparison of ROC curves for our model, the ALFSG Index (AUROC 0.839), against KCC (AUROC 0.684) and MELD (AUROC 0.717) in predicting LT or death in the same patients showed a statistically significant benefit to the ALFSG Index ($p=0.0003$, $p=0.0005$, respectively)

Table 1

Baseline Characteristics of Derivation Set versus Validation Set

Measure	N=250	N=250	p-value
Age [median (range)]	40.0 (17–78)	42.2 (18–87)	0.1356
Gender - Female n(%)	156(62.45%)	170 (68.0%)	0.1887
Etiology n(%)			0.0003
Acetaminophen	75 (30.0%)	121 (48.4%)	
Drug Induced Hepatitis	56(22.4%)	43 (17.2%)	
Indeterminate	33 (13.2%)	23 (9.2%)	
Hepatitis B	20 (8.0%)	9 (3.6%)	
Autoimmune hepatitis	26 (10.4%)	22 (4.4%)	
Shock	16(6.4%)	17 (6.8%)	
Other	24 (9.6%)	26 (10.4%)	
Outcome n(%)			
Transplanted	54 (21.6%)	53 (21.2%)	0.9132
Death	74 (29.6%)	82 (32.8%)	0.4400
Transplant Free Survival	125 (50%)	124 (48.4%)	*
Creatinine [median (range)]	1.6 (0.3–9.2)	2.0 (0.0–9.6)	0.1473
Bilirubin [median (range)]	11.8 (0.5–49.8)	6.8 (0.7–53)	0.0028
INR [median (range)]	2.6 (1.2–24.8)	2.7 (0.9–14.0)	0.9977
Coma Grade n(%)			0.0132
I	65(26.0%)	64 (25.7%)	
II	75 (30.0%)	49 (19.7%)	
III	59 (23.6%)	59 (23.7%)	
IV	51 (20.4%)	77 (30.9%)	
Phosphorus [median (range)]	3.2(0.3–244.0)	3.1 (0.4–89.0)	0.5414
ALT [median (range)]	1665 (22 – 15680)	1961(7– 19826)	0.1848
pH [median (range)]	7.44 (6.94 – 7.65)	7.43 (6.83–7.81)	0.3057
Sodium [median (range)]	139 (120 – 158)	139 (122 – 161)	0.5448
M-30 level [median (range)]	2725.7 (475–40005)	2138.7 (288–43869)	<0.0001
Log ₁₀ M-30 level [median (range)] ^{**}	3.44 (2.68–4.60)	3.33 (2.46–4.64)	
KCC Predicted to have transplant n(%)	102 (40.8%)	121 (48.4%)	0.0874
MELD [median (range)]	32 (11 – 40)	33 (11–40)	0.9578

Missing Data: Patients without complete information for all covariables included in the multivariable models were excluded from those analyses. A maximum of 38 of patients in the derivation set (15.2%) and 37 patients in the validation set (14.8%) were excluded for missing covariate information.

* A p-value for comparison of these two groups was not provided, since the derivation set and the validation set were selected to include a balanced number of patients who survived without transplantation versus those who were transplanted or died.

** log₁₀ M-30 is provided for reference purposes.

Table 2

The ALFSG Index: A model for prediction of need for liver transplantation or death in ALF using entry variables

Variable	Odds Ratio	95% C.I. for Odds Ratio	p value
Coma Grade			<0.001
II	1.335	0.546–3.262	0.526
III	3.425	1.342–8.744	0.01
IV	9.402	3.329–26.554	<0.001
Bilirubin	1.105	1.066–1.146	<0.001
INR	1.137	1.003–1.289	0.045
Phosphorus (≥ 3.7)	1.831	0.934–3.591	0.078
$\log_{10}$ M-30	3.331	1.214–9.143	0.019

Note: Entry level bilirubin, INR and $\log_{10}$ M30 are all continuous variables. Coma grade was a categorical variable with Coma Grade I the reference group. Phosphorus was used as a categorical variable at a threshold of ≥ 3.7 versus <3.7 .