Published in final edited form as:

J Pediatr. 2013 May ; 162(5): 1010–6.e1-4. doi:10.1016/j.jpeds.2012.11.021.

Evaluation of the Liver Injury Unit Scoring System to Predict Survival in a Multinational Study of Pediatric Acute Liver Failure

Brandy R. Lu, MD¹, Song Zhang, MS², Michael R. Narkewicz, MD^{1,3}, Steven H. Belle, MD², Robert H. Squires, MD⁴, Ronald J. Sokol, MD^{1,3}, and on behalf of the Pediatric Acute Liver Failure (PALF) Study Group*

¹Section of Pediatric Gastroenterology, Hepatology, and Nutrition, Department of Pediatrics, University of Colorado School of Medicine, Aurora, CO

²Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA

³Children's Hospital Colorado, Aurora, CO

⁴Department of Pediatrics, University of Pittsburgh and the Children's Hospital of UPMC, Pittsburgh, PA

Abstract

Objective—To examine the predictive value of the Liver Injury Units (LIU) and admission values (aLIU) of bilirubin and prothrombin time and international normalized ratio scores in a large cohort from the Pediatric Acute Liver Failure (PALF) Study Group, a multinational prospective study.

Study design—LIU and aLIU scores were calculated for 461 and 579 individuals, respectively, enrolled in the PALF study from 1999 to 2008. Receiver operator characteristic curves were used to evaluate the scores with respect to survival without liver transplantation (LT), death, or LT by 21 days after enrollment.

Results—At 21 days, 50.3% of participants were alive without LT, 36.2% underwent LT, and 13.4% died. The c-indices for transplant-free survival were 0.81 based on the LIU score with the international normalized ratio (95% CI, 0.78-0.85) and 0.76 based on the aLIU score (95% CI, 0.72-0.79). The LIU score predicted LT better than it predicted death (c-index for LT 0.84, c-index for death 0.76).

Conclusion—Based on data from a large, multicenter cohort of patients with PALF, the LIU score was a better predictor of transplant-free survival than was the aLIU score. The LIU score might be a helpful, dynamic tool to predict clinical outcomes in patients with PALF.

Pediatric acute liver failure (PALF) is a life-threatening illness in which a previously healthy child rapidly progresses to severe hepatic dysfunction and synthetic liver failure within 8 weeks of onset of symptoms.¹ Although the diagnosis of acute liver failure (ALF) in adults requires the presence of hepatic encephalopathy, this is difficult to assess in young children and is not always present and thus it is not included in the diagnostic criteria for PALF.² For

Copyright © 2013 Mosby Inc. All rights reserved.

Reprint requests: Ronald J. Sokol, MD, Children's Hospital Colorado, Box B290, 13123 E 16th Ave, Aurora, CO 80045.

ronald.sokol@childrenscolorado.org.

*List of members of the PALF Study Group is available at www.jpeds.com (Appendix).

Portions of this study were presented as a poster at the American Association of Liver Diseases Annual Meeting in October 30-November 3, 2009.

The authors declare no conflicts of interest.

many years, the characterization of PALF was based on the experience of single centers, which may introduce bias if generalized to all cases of PALF because of different center populations, varying definitions of PALF, and different time periods. In 1999, the PALF Study Group was created to overcome these deficiencies by collecting demographic, clinical, laboratory, and short-term outcome data in a uniform manner and with standardized nomenclature for pediatric cases of PALF from 24 pediatric centers (21 in the US, 2 in the United Kingdom, and 1 in Canada).² Despite current therapeutic approaches, PALF results in death or liver transplantation (LT) in up to 45% of pediatric patients. The ability to predict PALF clinical outcomes and the need for LT, to stratify patients for clinical trials, and to determine the severity of illness is currently limited² and is an unmet need in the clinical care and research priorities of PALF.

A promising scoring system for predicting LT-free survival in PALF was derived using objective laboratory data based on a single-center experience³ and validated in a second, independent cohort of patients with PALF from the same center.⁴ This system was named the Liver Injury Units (LIU) Scoring System, with $LIU = [3.584 \times \text{peak total bilirubin (mg/dL)}] + [1.809 \times \text{peak prothrombin time (PT) (seconds)}] + [0.307 \times \text{peak ammonia } (\mu\text{mol/L})]$. Alternatively, substituting international normalized ratio (INR) for PT, the score was calculated as $LIU = (3.507 \times \text{peak total bilirubin}) + (45.51 \times \text{peak INR}) + (0.254 \times \text{peak ammonia})$. An attempt to develop a scoring system using admission laboratory values demonstrated that only serum bilirubin and PT/INR were significantly associated with outcome; however, the admission LIU (aLIU) score did not have strong predictive ability.⁴ A number of other studies have identified the degree of cholestasis, coagulopathy, hepatic encephalopathy, or blood ammonia as predictors of death or LT; however, none have been validated for clinical decision making.^{2,5-9} The LIU score was validated at a single center,⁴ so there is a need to examine its applicability across multiple centers and in a larger sample. The objective of this study was to determine the predictive accuracy of the LIU and aLIU scores in participants enrolled in the PALF Study Group database, which includes >700 participants from 24 sites. Secondary aims included determining if the LIU score could be used on a daily basis and determining the relative predictive strength of the LIU score for LT versus death.

Methods

Enrollment in the PALF study cohort began in December 1999. The PALF study protocol has been described in detail.² Institutional review board approval was secured at each of the 24 clinical sites. Briefly, after informed consent was provided, demographic, clinical, and laboratory information was recorded daily on case report forms for up to 7 days after enrollment, and outcome was assessed at 21 days.² Diagnostic evaluation and medical management were consistent with the standard of care at each site. Data were collected until LT, death, or 21 days after enrollment. The LIU score was not used in treatment decisions, although components of the LIU score may have been used in clinical decisions at individual centers. Treatment recommendations were not part of the PALF study protocol. Completed data forms were coded and forwarded to the Data Coordinating Center at the University of Pittsburgh. Participants from birth through 17 years of age were eligible for enrollment if they met the following criteria: (1) no known evidence of chronic liver disease; (2) biochemical evidence of acute liver injury; and (3) hepatic-based coagulopathy defined as PT ≥ 15 seconds or INR ≥ 1.5 not corrected with vitamin K in the presence of hepatic encephalopathy or PT ≥ 20 seconds or INR ≥ 2.0 regardless of the presence or absence of clinical hepatic encephalopathy. The underlying cause of PALF, based on standard laboratory tests obtained clinically at each center and investigator judgment, included indeterminate, acetaminophen toxicity, autoimmune liver disease, infectious, non acetaminophen drug-induced liver disease, metabolic liver disease, shock, and other (Budd-

Chiari, hemophagocytic syndrome, leukemia, neonatal iron storage disease, and veno-occlusive disease).

Calculation of the LIU Score

For this study, both the LIU and aLIU scores were calculated for all PALF Study Group participants with available data, excluding those from the University of Colorado/Children's Hospital Colorado, which provided the study population for the original and replicate analysis of the LIU score.^{3,4} Although the criteria to define the PALF causes were the same at the University of Colorado and for the PALF Study, there were differences in data collection between the original derivation study and the current study. The original LIU score derivation was based on the peak laboratory values (total bilirubin, blood ammonia, PT/INR), defined as the highest recorded values during the entire PALF hospitalization (ending at transplantation if relevant), whereas the PALF study data were limited to laboratory values for the first 7 days after enrollment into the study. For each case, the highest values were obtained from available laboratory data, which were not available every day for all cases. In the original derivation, the aLIU score was calculated from admission values of total serum bilirubin and PT/INR (ammonia was not independently predictive and was not included in the aLIU calculation) obtained at the time the participant first met study criteria for PALF. In the PALF Study, the admission values were obtained on the day of enrollment into the PALF Study, which may have occurred after the day of admission or after entry criteria for the PALF Study were met. Outcomes of death before transplantation, transplantation, and alive with native organ were assessed at 3 weeks after enrollment in the PALF Study Group, compared with 16 weeks in the original derivation of the LIU and aLIU scores.^{3,4} In the original derivation, individual receiver operating characteristic (ROC) curves of the LIU/aLIU scores were generated for the combined end point of death or LT at 4 weeks after admission to hospital, as opposed to outcomes recorded at 3 weeks after enrollment in the PALF Study Group.⁴ Additional subcohort post hoc analyses included examining the LIU score based on cause (indeterminate vs known cause) and in participants at least 6 months old. Additional predictive analysis of the LIU score included using laboratory values obtained on the same day as opposed to during a 7-day period and they were called the highest daily laboratory values (highest daily LIU [hdLIU] score).

Statistical Analyses

SAS (SAS Institute, Cary, North Carolina), Stata (StataCorp, College Station, Texas), and R (CRAN) (<http://cran.r-project.org/>) were used for the analyses. Pearson ² test for association (without continuity correction) or Fisher exact test was used to compare categorical demographic and clinical characteristics between inclusion and exclusion groups and by outcomes, and the nonparametric Wilcoxon rank sum test was used to compare continuous lab values. ROC curves for the LIU score (and for the aLIU score) for 21-day outcomes were generated and c-indices (areas under the ROC curve) were compared using a nonparametric test.¹⁰ The product-limit method¹¹ was used to estimate survival curves (alive without LT vs death or LT) for each quartile of LIU (or aLIU) score, and a log-rank test¹² was used to determine whether they differed significantly. Because death without LT and LT are not equivalent events, they are preclusive and related to each other, the cumulative incidences of LT and of death without LT were also estimated and compared in a competing risk model.¹³ A value of $P < .05$ was considered to be statistically significant.

Results

From December 1999 through October 2008, the PALF Study Group registry included 709 participants, of whom 461 had data available for calculating the LIU score and 579 had data for calculating the aLIU score (excluding those from University of Colorado/The Children's

Hospital of Denver). Causes of PALF in the group with calculated LIU scores were indeterminate (49.9%), acetaminophen toxicity (12.8%), metabolic disease (8.7%), infection (6.7%), autoimmune liver disease (5.2%), and other (16.7%). At 21 days after enrollment, 50.3% of participants were alive without LT, 36.2% underwent LT, and 13.4% died without LT. Demographics, laboratory values, and clinical outcomes were compared between the 461 participants with LIU scores calculated and the 248 who were not included due to missing data (Table I). The significant differences between the 2 groups were that the excluded group, compared with those included in the analysis of the LIU score, tended to be younger at admission to study, to be more likely to have transferred to a study hospital from another hospital, and to have had lower ammonia and INR/PT values. A major reason for participants to be excluded from these analyses was missing ammonia values, as shown by the relatively few participants excluded who had ammonia levels recorded. In addition, the 21-day outcomes differed significantly between the 2 groups ($P < .0001$).

Evaluation of LIU Score

Figure 1 shows the proportion of participants surviving with native liver (without LT) stratified by the previously defined quartiles³ for the LIU score using INR ($n = 460$, 1 participant with unknown date of LT was excluded). The cumulative proportions of participants surviving with native liver at 3 weeks for each quartile using the LIU score with INR were 0.70, 0.38, 0.21, and 0.16, respectively ($P < .0001$). The c-index of the ROC curve for the model with LIU score using INR of survival with native liver was 0.81 (95% CI 0.78-0.85). Table II (available at www.jpeds.com) shows the distribution of participants with PALF by quartile and outcome. Figure 2 (available at www.jpeds.com) shows the survival distribution for quartiles of the LIU score using PT rather than INR ($n = 454$, 1 participant with unknown date of LT was excluded). The c-index for the ROC curves for the model with LIU score using PT (death or LT vs LT-free survival at 21 days) was 0.84 (95% CI 0.80-0.87). The P value was .06 comparing the c-index for LIU score calculated by the INR versus PT, for the 402 participants who had values available for both calculations.

Application of the hdLIU

We next sought to determine the predictive value of the LIU score using laboratory values that were all obtained on the same day, rather than the highest values for the 3 components independent of day of collection. The hdLIU score was calculated daily and the highest hdLIU score value was used in a logistic regression model with the combined end point of death or transplantation within 21 days of enrollment. This model ($n = 454$) had a c-index of 0.81 (95% CI 0.77-0.85), which was comparable with the value of 0.82 (95% CI 0.78-0.85) that was obtained using the highest values regardless of the day of collection (peak LIU score).

Evaluation of the aLIU Score

Figure 1, B shows the cumulative proportions of participants alive with native liver stratified by the previously defined⁴ quartiles for the aLIU score using INR ($n = 578$). The c-index for the model predicting transplantation-free survival was 0.76 (95% CI 0.72-0.79). The cumulative proportions of participants surviving without LT for 3 weeks for each quartile using the aLIU score with INR were 0.69, 0.56, 0.44, and 0.24 ($P < .0001$). Figure 2, B shows survival curves for the aLIU score using PT ($n = 546$). The c-index for predicting transplantation-free survival was 0.76 (95% CI 0.72-0.79). Thus, the aLIU score (using PALF enrollment values) did not predict as well as the LIU score for those who would survive without LT versus those who died or underwent LT.

Comparing the LIU Score for Death Alone versus LT Alone versus Survival Without LT

The LIU score was also used for predicting death alone (vs survival without LT) and LT alone (vs survival without LT). The c-index from ROC curves for predicting LT for the peak LIU using INR was 0.84 (95% CI 0.80-0.87). The c-index for predicting death for the peak LIU using INR was 0.76 (95% CI 0.70-0.82). Thus, discriminating those who survived without LT, the LIU score predicted LT better than it did death without LT. Up to 21 days post-enrollment into the study, the cumulative proportions of participants undergoing LT for each quartile using the LIU score with INR were 0.17, 0.41, 0.61, and 0.67, respectively ($P < .0001$) and the cumulative proportions of death without LT were 0.13, 0.21, 0.18, and 0.18, respectively ($P = .11$). Figure 3 (available at www.jpeds.com) shows the ROC curves for the LIU score using INR based on the outcome of death or LT, LT, or death versus transplant-free survival.

To examine possible factors accounting for the reduced predictive strength of the LIU score for death compared with LT, demographics, treatments, and laboratory values were compared between those experiencing the outcomes of death without LT and LT (Table III; available at www.jpeds.com). Participants who died compared with those undergoing LT tended to be younger ($P = .0001$), were less likely to have a diagnosis of indeterminate PALF (42.0% vs 68.2%), were less likely to have received plasmapheresis (6.5% vs 20.1%, $P = .01$), had lower mean admission total bilirubin (median 9.2 vs 17.4 mg/dL, $P < .0001$), had lower peak total bilirubin (median 15.3 vs 19.4 mg/dL, $P < .002$), and had lower peak INR (median 4.2 vs 4.9, $I = 0.02$).

We sought to analyze whether age was influencing the predictive value of the LIU score and compared the c-index of the LIU score for all participants <6 months of age ($n = 84$) versus participants at least 6 months of age ($n = 377$). The c-index for the ROC curve of LIU scores using peak INR of transplant-free survival versus death or LT in participants at least 6 months of age was 0.83 (95% CI 0.79-0.88, $n = 377$), for predicting LT versus transplant-free survival was 0.85 (95% CI 0.81-0.89, $n = 337$), and for predicting death versus transplant-free survival was 0.78 (95% CI 0.70-0.85, $n = 230$). The c-index of LIU scores using peak INR of transplant-free survival versus death or transplantation in participants <6 months of age was 0.71 (95% CI 0.60-0.82, $n = 84$), for predicting LT versus transplant-free survival was 0.76 (95% CI 0.64-0.89, $n = 62$), and for predicting death versus transplant-free survival was 0.67 (95% CI 0.53-0.81, $n = 64$). The only significant difference between these 2 age groups was for transplant-free survival (0.83 vs 0.71, $P = .04$).

Finally, we sought to determine whether the cause of PALF was associated with the predictive value of the LIU score. Using values from only those with indeterminate PALF ($n = 230$), the c-index from a model using peak INR for predicting 21-day survival with native liver versus death or LT was 0.78 (95% CI 0.72-0.85). For the model predicting LT versus survival ($n = 204$), the c-index using peak INR was 0.80 (95% CI 0.74-0.86), and for predicting death versus survival without LT ($n = 116$), the c-index was 0.71 (95% CI 0.60-0.82). For the participants with PALF with a specific defined etiology ($n = 231$), the c-index for the model predicting survival with native liver was 0.82 (95% CI 0.76-0.87). When predicting LT versus transplant-free survival ($n = 195$), the c-index was 0.83 (95% CI 0.77-0.89), and for predicting death versus transplant-free survival ($n = 178$) was 0.79 (95% CI 0.71-0.87) (Table IV; available at www.jpeds.com). The c-index for the model predicting death without LT was not significantly different for participants with a specific etiology of PALF than for those with an indeterminate cause ($P = .23$).

Analyses were then performed to account for the competing risks between death and LT in the indeterminate and specific etiology subgroups. For the participants with an indeterminate diagnosis, the cumulative incidence of LT for each quartile using the LIU score with INR

were 0.14, 0.28, 0.58, and 0.75, respectively ($P < .0001$) 21 days post-enrollment, and the cumulative incidences of death were 0.14, 0.17, 0.12, and 0.12, respectively ($P = .96$) (Figure 4, A and B). For participants with a specific PALF etiology, the cumulative incidences of LT for each quartile using the LIU score with INR were 0.11, 0.25, 0.30, and 0.60, respectively ($P < .0001$), and the cumulative incidences of death were 0.02, 0.16, 0.39, and 0.21, respectively ($P = .002$) (Figure 4, C and D). Separation of LIU quartiles for death and LT was demonstrated by significant P values for specified diagnosis of PALF, although the cumulative incidence of death in the fourth quartile is below that of the third quartile, whereas only LT had significant separation of quartiles for indeterminate diagnosis. The LIU score was not predictive in participants who died with an indeterminate diagnosis of PALF, but this could be due to the small number of participants ($n = 26$).

Discussion

In this study, using a large multicenter multinational cohort of children with PALF, the LIU score was shown to be strongly predictive of transplant-free survival (c-index 0.81). The aLIU score showed moderate predictive strength (c-index 0.76) at time of enrollment in the PALF Study, similar to the previous study.⁴ The weaker predictive ability of the aLIU score is likely a reflection of the variable time interval between onset of symptoms and admission and the wide spectrum of severity of PALF at clinical presentation to PALF study institutions. The strength of the current study over past evaluations of the LIU score^{3,4} was access to data collected from the largest existing cohort participants with PALF from multiple centers within 3 countries, supporting the broad applicability of these findings. The components of the LIU score are common laboratory tests (total bilirubin, ammonia, PT/INR) that are used qualitatively by clinicians to assess the severity of PALF and the consideration for LT. The LIU scoring system adds a quantitative predictive model based on these tests that may contribute objective criteria useful for the decision-making process regarding listing and timing for LT.

Calculating the LIU score in this study using the hdLIU had comparable predictive value for outcomes as did the overall peak LIU score used in prior analyses.^{3,4} This suggests that the LIU score, based on values obtained on any given day, has the potential to be used in a similar manner to the Model of End-Stage Liver Disease (MELD)/Pediatric End-stage Liver Disease (PELD) scoring systems in chronic liver disease (ie, as a dynamic measure that reflects the risk for transplant-free survival at any given point in time during hospitalization for PALF). Attempts have been made to use the MELD/PELD scoring system to assess prognosis in ALF with varying results.¹⁴ Both MELD and PELD were derived to predict the probability of mortality in patients who had chronic liver disease, which is an entity distinct from PALF. Although components of the MELD were thought to be a predictive factor in ALF (INR and bilirubin), the MELD score itself had poor specificity.¹⁴ Rhee et al¹⁵ eloquently showed that separate prioritizing systems for LT are needed in children with chronic liver disease and with PALF.

Another potential use for the LIU score is for stratifying patients with PALF in future clinical trials for enrollment or data analysis. In addition, the LIU score could be used to ensure that only patients with a poor prognosis would be included preferentially to test new treatments or interventions, reducing the sample size (and cost) needed for the clinical trial. Finally, the possible use of the LIU score as a surrogate marker for treatment outcomes in PALF will require validation.

The LIU score appeared to predict the likelihood of receiving a liver transplant better than the risk of death. Possible explanations for this include that participants who die of PALF may do so rapidly from a complication, thus providing insufficient time for the serum

bilirubin or INR to reach levels observed in participants who survived long enough to undergo LT. In addition, therapeutic manipulations, such as administration of fresh frozen plasma, activated factor VII, or plasmapheresis, could artificially lower the LIU score and impact its predictive value. Ammonia levels can be manipulated with medications or have erroneous values based on how the sample is handled and the source of blood. Despite the fact that these interventions were used in some of the participants with PALF, the LIU score still was predictive of need for LT. It also appears that the LIU score is less predictive of overall outcome (death or LT vs transplant-free survival) in participants <6 months of age compared with those ≥6 months of age. It is not possible to determine what factor related to young age may be playing a role in this discrepancy (eg, cause of PALF or other host factors), given that there were only 84 patients in this age range to evaluate.

Another possible interpretation of the discrepancy in LIU score predictive strength between death and LT is that these 2 populations are not equivalent (ie, a subset of participants with PALF who received a transplant would have survived PALF [and not died] without LT). Of the 547 participants who had aLIU scores calculated, only 51% (283 participants) were ever listed for LT (data not shown). Of those listed, 10% (30 participants) were removed from the transplant list either due to irreversible disease (20% of those removed, 6 participants) or because they were recovering (80%, 24 participants). Thus, only 40% of all participants with PALF were on the transplant waiting list at 1 week and, of those still listed, 67% were transplanted (165 participants), 7% died (19 participants), and 25% survived without LT (62 participants). Although 25% may appear to be a high survival rate in participants listed for LT, it is important to recognize that these 62 participants represented only 11% of the total participants who presented with PALF, suggesting that liver transplant centers are selecting appropriate candidates for LT. However, because the availability of donor organs generally dictates when a patient with PALF is transplanted once the patient is listed, it is likely that some patients with PALF who underwent transplantation would have survived without transplantation, although the quality of their central nervous system recovery without transplantation may have been compromised. In addition, organ availability varies based on a center's experience with different types of LT (ie, living related, deceased cardiac death donor, reduced segments), which would impact outcome.

One concern about the derivation and the validation of the LIU score is that clinicians may have used components of the LIU score to prioritize for LT; thus, the LIU score has an inherent incorporation bias. However, it is important to note that the LIU score itself (and its weighted calculations) was not used in treatment decisions. The LIU score analysis for this study was done retrospectively, so enrolling centers were not aware that this study was to be undertaken and the lack of awareness of the LIU score applicability should have minimized incorporation bias. Another concern is that the excluded group may have been somewhat less ill, based on lower INR and ammonia, better outcome, and less indeterminate cause of PALF; however, LIU scores could not be calculated on this group because of missing data to assess how predictive the LIU score may have been in this group. Another concern about this evaluation of the LIU score might relate to the different time points used in this study for assessing outcome compared with previous studies of the LIU score. The original LIU score derivation was at a single-center site, which had the benefit of longer follow-up at 16 weeks as opposed to 3 weeks in the PALF database. However, the use of a large cohort from a multicenter study of rare pediatric liver disease is invaluable, so the slight differences in methodologies were justified. When the validation cohort ($n = 53$) was reexamined,⁴ all LTs ($n = 7$) occurred before 21 days and 4 of 8 deaths occurred between 3 weeks and 16 weeks. This discrepancy may explain why the LIU score was better at predicting LT than death. However, only 1 death occurred between 3 and 4 weeks, so we would expect the ROC curves that were calculated at 4 weeks in the original study to be comparable with ROC curves calculated at 3 weeks in this study.

In conclusion, using a large, multicenter multinational database study of children with PALF, the LIU score (using highest laboratory values during a 7-day period) was predictive of survival without LT in PALF. The hdLIU score (using laboratory values obtained on the same day) was also predictive of survival. The aLIU score was moderately but less predictive of survival and may not prove to be useful. The LIU score and the hdLIU score may be helpful tools to predict clinical outcomes in PALF, allow for prioritization of patients with PALF who would benefit from LT listing, and should be explored for use in stratifying patients in clinical trials and as a surrogate biomarker for clinical outcomes.

Acknowledgments

PALF Study Group is supported by National Institutes of Health (NIH; U01 DK072146) and University of Colorado Denver Colorado Clinical and Translational Sciences Institute (UL1 TR000154). B.L. is supported by NIH (1 T32 DK067009-01).

Appendix

Additional current and former principal, co-investigators, and research coordinators of the PALF Study Group include: University of Pittsburgh, Pennsylvania: Benjamin L. Shneider, MD; Cincinnati Children's Hospital Medical Center, Ohio: John Bucuvalas, MD, and Mike Leonis MD, PhD; Ann & Robert H. Lurie Children's Hospital of Chicago, Illinois: Estella Alonso, MD; University of Texas Southwestern: Norberto Rodriguez-Baez, MD; Seattle Children's Hospital, Washington: Karen Murray, MD, and Simon Horslen, MD; Children's Hospital Colorado, Aurora: Michael R. Narkewicz, MD; St Louis Children's Hospital, Missouri: David Rudnick, MD, PhD, and Ross W Shepherd, MD; University of California, San Francisco: Philip Rosenthal, MD; Hospital for Sick Children, Toronto, Canada: Vicky Ng, MD; Riley Hospital for Children, Indianapolis, Indiana: Girish Subbarao, MD; Emory University: Rene Romero, MD; Children's Hospital of Philadelphia, Pennsylvania: Elizabeth Rand, MD, and Kathy Loomis, MD; Kings College, London, United Kingdom: Anil Dhawan, MD; Birmingham Children's Hospital, United Kingdom: Dominic Dell Olio, MD, and Deirdre A. Kelly, MD; Texas Children's Hospital: Saul Karpen, MD, PhD; Mt Sinai Medical Center: Nanda Kerkar, MD; University of Michigan: M. James Lopez, MD, PhD; Children's Hospital Medical Center, Boston, Massachusetts: Scott Elisofon, MD, and Maureen Jonas MD; Johns Hopkins University: Kathleen Schwarz, MD; and Columbia University: Steven Lobritto, MD.

Glossary

ALF	Acute liver failure
aLIU	Admission Liver Injury Units
hdLIU	Highest daily Liver Injury Units
INR	International normalized ratio
LIU	Liver Injury Units
LT	Liver transplantation
MELD	Model of End-stage Liver Disease
PALF	Pediatric acute liver failure
PELD	Pediatric End-stage Liver Disease
PT	Prothrombin time

ROC Receiver operating characteristic

References

1. Squires RH. Acute liver failure in children. *Semin Liver Dis.* 2008; 28:153–66. [PubMed: 18452115]
2. Squires RH, Schneider BL, Bucuvalas J, Alonso E, Sokol RJ, Narkewicz MR, et al. Acute liver failure in children: the first 348 patients in the Pediatric Acute Liver Failure Study Group. *J Pediatr.* 2006; 148:652–8. [PubMed: 16737880]
3. Liu E, MacKenzie T, Dobyns EL, Parikh CR, Karrer FM, Narkewicz MR, et al. Characterization of acute liver failure and development of a continuous risk of death staging system in children. *J Hepatol.* 2006; 44:134–41. [PubMed: 16169116]
4. Lu BR, Gralla J, Liu E, Dobyns EL, Narkewicz MR, Sokol RJ. Evaluation of a scoring system for assessing prognosis in pediatric acute liver failure. *Clin Gastrohepatol.* 2008; 6:1140–5.
5. Lee WS, McKiernan P, Kelly DA. Etiology, outcome and prognostic indicators of childhood fulminant hepatic failure in the United Kingdom. *J Pediatr Gastroenterol Nutr.* 2005; 40:575–81. [PubMed: 15861019]
6. Rivera-Penera T, Moreno J, Skaff C, McDiarmid S, Vargas J, Ament ME. Delayed encephalopathy in fulminant hepatic failure in the pediatric population and the role of liver transplantation. *J Pediatr Gastroenterol Nutr.* 1997; 24:128–34. [PubMed: 9106097]
7. Dhawan A, Cheeseman P, Mieli-Vergani G. Approaches to acute liver failure in children. *Pediatr Transplant.* 2004; 8:584–8. [PubMed: 15598330]
8. Ciocca M, Ramonet M, Cuarterolo M, Lopez S, Cernadas C, Alvarez F. Prognostic factors in pediatric acute liver failure. *Arch Dis Chld.* 2008; 93:48–51.
9. Nicolette L, Billmire D, Faulkenstein K, Pierson A, Vinocur C, Weintraub W, et al. Transplantation for acute hepatic failure in children. *J Pediatr Surg.* 1998; 33:998–1003. [PubMed: 9694084]
10. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology.* 1982; 143:29–36. [PubMed: 7063747]
11. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc.* 1958; 53:457–81.
12. Breslow NE. Covariance analysis of censored survival data. *Biometrics.* 1974; 30:89–99. [PubMed: 4813387]
13. Gray RJ. A class of K-sample tests for comparing the cumulative incidence of a competing risk. *Ann Stat.* 1988; 16:1141–54.
14. Polson J. Assessment of prognosis in acute liver failure. *Semin Liver Dis.* 2008; 28:218–25. [PubMed: 18452121]
15. Rhee C, Narsinh K, Venick RS, Molina RA, Nga V, Engelhardt R, et al. Predictors of clinical outcome in children undergoing orthotopic liver transplantation for acute and chronic liver disease. *Liver Transpl.* 2006; 12:1347–56. [PubMed: 16741901]

Figure 1.

A, Distribution of the LIU score by previously derived quartiles using INR for participants from the PALF Study Group. Survival without LT with native liver is stratified by quartiles of the LIU score. $LIU = 3.507 \times \text{peak total bilirubin (mg/dL)} + 45.51 \times \text{peak INR (seconds)} + 0.254 \times \text{peak ammonia } (\mu\text{mol/L})$. **B**, Distribution of the aLIU score by previously derived quartiles using INR for participants from the PALF Study Group. Survival without LT with native liver stratified by quartiles of the aLIU score. $aLIU = 8.4 \times \text{admission bilirubin (mg/dL)} + 50.0 \times \text{admission peak INR}$.

Figure 2.

A, Distribution of the LIU score by previously derived quartiles using PT on participants from the PALF Study Group. Survival without LT with native liver is stratified by quartiles of the LIU score. $LIU = 3.584 \times \text{peak total bilirubin (mg/dL)} + 1.809 \times \text{peak PT (seconds)} + 0.307 \times \text{peak ammonia } (\mu\text{mol/L})$. **B**, Distribution of the aLIU score by previously derived quartiles using PT for participants from the PALF Study Group. Survival without LT with native liver stratified by quartiles of the aLIU score. $aLIU = 6.9 \times \text{admission bilirubin (mg/dL)} + 4.0 \times \text{admission peak PT (seconds)}$.

Figure 3.

A, ROC curve of LIU score using INR for predicting 21-day outcome of death or LT versus LT-free survival (n = 461). **B**, ROC curve of LIU score using INR for predicting 21-day outcome of LT versus LT-free survival (n = 399). **C**, ROC curve of LIU score using INR for predicting 21-day outcome of death versus LT-free survival (n = 294).

Figure 4.

A, Likelihood of LT divided by quartiles of the LIU score using INR in participants with PALF with an indeterminate diagnosis. **B,** Likelihood of death without LT divided by quartiles of the LIU score using INR in participants with PALF with an indeterminate diagnosis. **C,** Likelihood of transplantation divided by quartiles of the LIU score using INR in participants with PALF with a specified diagnosis. **D,** Likelihood of death without LT divided by quartiles of the LIU score using INR in participants with PALF with a specified diagnosis.

Table I

Demographics, clinical data, and laboratory values for the patients included in the LIU score analysis (with peak INR) and for patients excluded from analysis because of missing laboratory values

	Included in calculating LIU (n = 461), No. (%)	Excluded in calculating LIU (n = 248), No. (%)	P
Sex			
Male	229 (49.7)	124 (50.0)	.93 [†]
Female	232 (50.3)	124 (50.0)	
Median age (25%, 75%)	5.5 (1.1, 13.5)	3.8 (0.3, 13.3)	.02 [*]
0-<6 mo	84 (18.2)	72 (29.0)	.004 [†]
6 mo-<3 y	93 (20.2)	46 (18.6)	
3 y	284 (61.6)	130 (52.4)	
Patients from outside hospital	346 (75.2)	211 (85.1)	.002 [†]
Final diagnosis			
Indeterminate	230 (49.9)	96 (38.7)	.08 [†]
Acetaminophen overdose	59 (12.8)	35 (14.1)	
Metabolic	40 (8.7)	30 (12.1)	
Autoimmune	24 (5.2)	21 (8.4)	
Infection	31 (6.7)	19 (7.7)	
Other	77 (16.7)	47 (19.0)	
Outcome on day 21			
Alive without LT	232 (50.3)	166 (66.9)	<.0001 [†]
LT	167 (36.2)	55 (22.2)	
Death without LT	62 (13.4)	27 (10.9)	

Laboratory value	n, median (25%, 75%)	n, median (25%, 75%)	
Total bilirubin level, mg/dL			
Admission	441, 10.5 (3.2, 18.5)	219, 9.6 (2.8, 16.5)	.16 [*]
Highest value	461, 13.3 (4.1, 21.4)	229, 11.8 (3.1, 20.2)	.08 [*]
PT, s			
Admission	391, 28.0 (23.0, 39.0)	192, 25.3 (20.8, 33.7)	.002 [*]
Highest value	402, 33.3 (25.1, 45.5)	196, 28.9 (21.4, 39.0)	.0003 [*]
INR			
Admission	449, 2.8 (2.1, 4.3)	168, 2.7 (2.0, 3.9)	.04 [*]
Highest value	461, 3.6 (2.5, 5.4)	171, 2.9 (2.1, 4.5)	.0004 [*]
Ammonia level, μ mol/L			
Admission	412, 71.5 (44.0, 116.0)	56, 55.0 (37.5, 92.0)	.04 [*]
Highest value	461, 90.0 (57.0, 135.0)	67, 57.0 (41.0, 97.0)	.0002 [*]

* Wilcoxon rank sum test.

[†] 2 Test.

Table II

Patients with PALF distribution based on peak LIU score quartile using INR and outcome of death, LT, or survival without LT (n = 460)

Peak LIU score quartile using INR	Death (n = 62; 13.4%)	LT (n = 167; 36.2%)	Survival without LT (n = 232; 50.3%)
1. 0-209	13 (7.2 % of 1st quartile)	21 (11.7% of 1st quartile)	146 (81.1% of 1st quartile)
2. 210-296	21 (19.1% of 2nd quartile)	40 (36.4% of 2nd quartile)	49 (44.5% of 2nd quartile)
3. 297-369	12 (17.1% of 3rd quartile)	41 (58.6% of 3rd quartile)	17 (24.3% of 3rd quartile)
4. 370	16 (15.8% of 4th quartile)	65 (64.4% of 4th quartile)	20 (19.8% of 4th quartile)

Table III

Demographics, clinical data, and laboratory values for patients with PALF (who had LIU scores using INR calculated) who either died without LT or underwent LT

	Death without LT (n = 62), No. (%)	LT (n = 167), No. (%)	P
Sex			.66 [*]
Male	34 (54.8)	97 (58.1)	
Female	28 (45.2)	70 (41.9)	
Median age (25%, 75%)	1.9 (0.1, 10.0)	5.8 (1.7, 12.2)	.0001 [*]
0-<6 mo	22 (35.5)	20 (12.0)	
6 mo-3 y	13 (21.0)	34 (20.4)	
3 y	27 (43.5)	113 (67.6)	
Outside hospital referrals	51 (82.0)	126 (75.7)	.27 [*]
Final diagnosis			<.0001 [‡]
Indeterminate	26 (42.0)	114 (68.2)	
Acetaminophen overdose	1 (1.6)	5 (3.0)	
Metabolic	7 (11.3)	16 (9.6)	
Autoimmune	1 (1.6)	10 (6.0)	
Infection	9 (14.5)	8 (4.8)	
Other	18 (29.0)	14 (8.4)	
Received fresh frozen plasma	56 (91.8)	131 (81.4)	.06 [*]
Received plasmapheresis	4 (6.5)	33 (20.1)	.01 [*]

Laboratory values	N, median (25%, 75%)	N, median (25%, 75%)	P [‡]
Total bilirubin, mg/dL			
Admission	60, 9.2 (3.9, 15.7)	158, 17.4 (11.9, 22.9)	<.0001
Highest value	62, 15.3 (6.5, 22.8)	167, 19.4 (14.7, 25.6)	.002
PT, s			
Admission	50, 31.4 (25.2, 40.1)	140, 31.7 (24.2, 44.3)	.76
Highest value	50, 35.6 (28.6, 46.0)	144, 40.0 (31.9, 57.3)	.046
INR			
Admission	61, 3.5 (2.4, 4.7)	164, 3.4 (2.5, 5.3)	.56
Highest value	62, 4.2 (2.9, 5.5)	167, 4.9 (3.4, 7.7)	.02
Serum ammonia, μ mol/L			
Admission	54, 116.5 (65.0, 157.0)	148, 94.0 (52.5, 135.0)	.13
Highest value	62, 130.0 (86.0, 197.7)	167, 119.0 (79.0, 163.0)	.27

* ² Test.

[‡] Wilcoxon rank sum test.

[‡] Exact ² test.

Table IV

Area under ROC curve (c-index) of LIU score using INR for predicting 21-day outcome for all patients with PALF, those defined as “indeterminate” or those with a specific PALF cause

21-d Outcome	c-Index of LIU score using INR		
	All PALF causes (95% CI)	Indeterminate PALF cause (95% CI)	Specific PALF cause (95% CI)
Death or LT vs transplant-free survival	0.81 (0.76-0.85) n = 461	0.78 (0.72-0.85) [*] n = 230	0.82 (0.76-0.87) [*] n = 231
LT vs transplant-free survival	0.84 (0.80-0.87) n = 399	0.80 (0.74-0.86) [†] n = 204	0.83 (0.77-0.89) [†] n = 195
Death vs transplant-free survival	0.76 (0.70-0.82) n = 294	0.71 (0.60-0.82) [‡] n = 116	0.79 (0.71-0.87) [‡] n = 178

^{*} $P = .44$.

[†] $P = .48$.

[‡] $P = .23$.