Gastro Hep Advances 2022;1:538–545

ORIGINAL RESEARCH—BASIC

Accounting for Liver Transplant in Acute Liver Failure Research

Sherry I. Livingston and Valerie Durkalski-Mauldin

Department of Public Health Sciences, Medical University of South Carolina, Charleston, South Carolina

BACKGROUND AND AIMS: Acute liver failure (ALF) is a rare but serious disease with challenging clinical decisions, including the possibility of liver transplantation. Although there is interest in predicting who will need a transplant, that outcome is difficult to define as the decision to transplant includes many extraneous factors. The majority of research in this setting focuses on identifying factors that can provide guidance on a patient's likelihood of survival without a liver transplant. The question that arises is whether death and transplant should be combined as a poor outcome or should alternative approaches be used to account for transplant in this setting. Furthermore, does the approach to incorporating transplant information impact the accuracy of predicting survival. We aim to compare alternative analytic methods for the ALF setting to provide guidance to the clinical research community on how to handle transplant when the outcome of interest is survival without a transplant. **METHODS:** Five analysis approaches are compared based on model performance using existing registry data from 2100 ALF patients: logistic regression with transplant as part of the outcome, logistic regression with transplant as a covariate, inverse probability weighting, survival analysis, and multiple imputation. **RESULTS:** The various models exhibit comparable model fit with each providing advantages and challenges in implementation. **CONCLUSION:** There are alternative modeling approaches in the ALF setting, leaving researchers with multiple valid options for how to include transplant when examining factors that may influence transplant-free survival.

Keywords: Prognostic Modeling; Transplant-Free Survival; Liver Transplant Modeling; Spontaneous Survival

Acute liver failure (ALF) is a rare condition in which loss of liver cell function occurs rapidly over a few days or weeks in those without established liver disease. Characteristic features include coagulopathy (international normalized ratio [INR] $\geq$1.5) and hepatic encephalopathy (HE). Those patients who develop ALF with a prolonged coagulopathy and HE carry an overall mortality of 30% even in the age of liver transplantation.[1] While approximately 65% of acetaminophen-induced ALF patients will spontaneously recover without liver transplantation, several other etiologies of ALF carry a much lower probability of transplant-free survival (TFS). The variation in the outcome by the cause of ALF is important to note since etiology tends to guide treatment. Clinical decision-making in the setting of adult ALF is challenging for several reasons, including the possibility of the patient receiving a liver transplant. Beyond transplant, treatments for ALF are few, with only N-acetylcysteine being proven to be effective in patients with ALF due to acetaminophen overdose and other ALF etiologies.[2] Because new therapies are not on the near horizon for this condition, researchers in this field often conduct prognostic modeling to identify clinical factors that are associated with the 'good outcome', meaning surviving without a liver transplant.

Prognostic modeling is used in clinical research to determine the probability of an event of clinical interest occurring given a set of potential risk factors. In a clinical setting, it is common for the event of interest to be binary (eg, is the disease present, will the patient survive). Being an acute illness, the clinical outcome of interest in the ALF population is often TFS. If a liver transplant occurs, the patient is likely to survive (probability of survival after transplant is approximately 90%)[3]; however, the quality of life tends to be lower than the average US population and similar to patients who survive without a transplant.[4] For this reason, when clinicians are interested in identifying risk factors associated with survival, the outcome is often dichotomized as TFS vs death or transplant. There is particular interest in predicting who will need a transplant and thus listing for transplant, but that outcome is difficult to define as the decision to transplant includes many extraneous factors. Several prognostic indices have been developed and validated for use in the ALF population including King's criteria that provide a measure of the poor outcome, the Model for End-Stage Liver Disease score that provides severity of disease, and more recently, the ALF prognostic index that provides the probability of TFS.[5–7] Although these are good bedside resources for the treating

Abbreviations used in this paper: **ALF, acute liver failure; ALFSG, Acute Liver Failure Study Group; AUROC, area under the receiver operator characteristic; HE, hepatic encephalopathy; INR, international normalized ratio; MI, multiple imputation; ROC, receiver operating characteristic; TFS, transplant-free survival.**

Most current article

2772-5723
https://doi.org/10.1016/j.gastha.2022.02.026

physician, the research community continues to explore additional factors that may impact TFS including clinical labs or specific biomarkers. Several manuscripts have been published that conduct a retrospective analysis from existing registry data to examine the relationship between potential prognostic variables and TFS.[8,9] The approach to examining this relationship raises an important question as to whether it is appropriate for transplant to be considered a 'poor outcome' along with death. An alternative option is to define the outcome as alive or dead and account for transplant as a covariate in the model. Statistically there are several methods available to accommodate either approach. What is needed is clarification on the pros and cons of defining the outcome in these various ways and a better understanding of the potential impact on results. The current study compares available analytic methods for the assessment of clinical risk factors associated with the outcome, highlighting the benefits and limitations of choosing one method over the other using real data from an ALF registry.

Setting

The Acute Liver Failure Study Group (ALFSG) is a multicenter network funded since 1998 by the National Institute of Diabetes and Digestive and Kidney Diseases to study the etiology and outcomes of ALF and severe acute liver injury patients. The group maintains a registry of ALF patients who are hospitalized at participating institutions representing 30 transplant centers across the United States and 1 in Canada (ClinicalTrials.gov: NCT00518440). The ALFSG definition of ALF is coagulopathy (INR $\geq$1.5) and any degree of HE within 26 weeks of the first symptoms without preexisting liver disease. A primary clinical outcome of research interest to the ALFSG and clinical community is TFS at 21 days from registry enrollment (21-day TFS).

Method

The registry data collected from January 1998 through December 2018 are used in 5 analysis approaches: (1) logistic regression using an outcome defined as 21-day TFS vs death or transplant, (2) logistic regression with an outcome defined as 21-day survival vs death and transplant status (yes or no) as an independent variable in the model, (3) inverse probability weighting with an outcome defined as 21-day survival vs death and transplant status incorporated into the model using a prognostic score, (4) survival analysis with an outcome defined as 21-day survival vs death, with transplant included in the model as a time-varying covariate, and (5) logistic regression with an outcome defined as 21-day survival vs death, and patients who receive a transplant are considered as having a missing outcome, which is then imputed using multiple imputation (MI).

Our goal is not to create a new prognostic model, but to look at the impact of alternative ways to handle transplant when assessing potential prognostic factors related to survival without a transplant. To that end, the independent variables in each model are consistent across the various approaches and include the following previously identified prognostic variables: low-risk or high-risk etiology (defined as acetaminophen overdose, pregnancy, ischemia, or hepatitis A for low risk and all other etiologies as high risk); use of vasopressors; coma grade (1–2 or 3–4); log bilirubin; and log INR.[7,10] The different models are compared based on the concordance statistic (c-statistic) which measures goodness of fit for binary outcomes and is equivalent to the area under the receiver operator characteristic (AUROC) curve. In addition, predictive accuracy was assessed using the Brier score, and calibration for agreement between predictive and observed probabilities was examined using slopes of calibration plots.

Analysis Approaches

Logistic Regression (Models 1 and 2). Logistic regression is a common statistical modeling approach used to examine the relationship between a single binary outcome and one or more binary, continuous, and/or ordinal independent variables. In the prognostic setting, this model is often used to determine how a set of patient characteristics predict the outcome of interest.[11] Model 1 was created using a logistic regression model with a binary outcome of 21-day TFS vs transplant or death. The purpose of this model is to examine factors associated with TFS and predict a patient's expected probability of surviving at 21 days from enrollment without having to receive a transplant. Model 2 uses logistic regression with a binary outcome of survival vs death, and the variable transplant is used as a binary covariate in the model. Logistic regression is a robust statistical model used for binary outcomes; however, a limitation is that a missing outcome or covariate variables result in removing the record from the analysis set.

Inverse Probability Weighting (Model 3). Inverse probability weighting with propensity scores is used in observational studies to account for potential differences in baseline characteristics (independent variables) of the study population. A propensity score is the predicted probability of an event based on the baseline characteristics, and in theory, patients with the same propensity score will have similar distributions of observed baseline covariates.[12] The calculated propensity score is used to create the inverse probability weight, which is equal to the reciprocal of the predicted probability of an event. In our setting, a prognostic score is calculated to represent the predicted probability of a patient receiving a transplant. Predictors for transplant that were included in the propensity score model were admission coma grade, etiology, vasopressor use, bilirubin level, INR, race, age, and use of continuous veno-venous hemofiltration. The inverse of the prognostic score creates a "pseudo-population," where baseline covariates are independent of transplant status. The use of stabilized

weights can reduce the effect of very high weights that result from patients with extreme (high/low) propensity scores. The goal of the propensity score is to create a balance between the groups who did and did not receive a transplant. It is therefore important to verify that patients in these 2 groups with similar propensity scores are balanced on the covariates that were used in the calculation of the propensity score.[13] If balance is not achieved, then the method is not valid. If the balance assumption holds, a logistic regression model can be used to predict the probability of 21-day survival, using the inverse probabilities as weights and including transplant in the model as an independent variable.

Survival Analysis (Model 4). Survival models are used to model time-to-event data. Hazard ratios are estimated to provide information on the probability that a patient will experience the event of interest at a given time point assuming the patient has not experienced the event at that particular point in time. A key aspect of survival models is censoring, which is how the model handles incomplete data (unknown time of events whether due to the event not occurring or the date of occurrence being unknown). In the setting of ALF, receipt of a liver transplant alters the probability of the event of interest (eg, survival); therefore, ignoring transplant or treating transplant as noninformative censoring can provide misleading results. For example, if we estimate the probability of 21-day survival for an ALF patient, this probability will change if the patient receives a transplant. However, we can use a survival model and consider including transplant as a time-dependent covariate in a Cox proportional hazards model.[14,15]

Multiple Imputation (Model 5). MI is a method for dealing with missing data.[16] For patients who receive a transplant, it is unknown whether they would have survived without that transplant, and so their TFS outcome can be treated as missing. In this case, we are assuming that the outcome is "missing at random," meaning that the missing outcome can be predicted by observed variables. Baseline covariates are used to impute the outcome of each transplanted patient, assuming they did not receive a transplant. This imputation is done multiple times, and the results are pooled to obtain one predicted outcome for whether that patient would have survived without a transplant. A logistic model is then created using the actual 21-day survival outcome for nontransplanted patients and the imputed outcome for transplanted patients. Transplant is not a covariate in the model.

These methods can be implemented in various statistical packages including R and SAS software. The output and data analysis for this study were generated using SAS software, Version 9.4 (SAS Institute, Cary, NC).

Results

The data set contained 2249 ALF registry participants, of whom 2100 patients had all 5 predictor variables nonmissing. The baseline characteristics and outcomes of the analysis population are listed in Table 1. A total of 473 patients (22.5%) received a liver transplant.

Models 1 and 2 are logistic models as previously described. Model 3 utilizes inverse probability weighting. In this model, propensity scores were calculated to predict the probability that a patient will receive a transplant based on their baseline covariates. The propensity score was then used to calculate stabilized weights for each patient. For patients who received a transplant, the weight equals 0.2252/(propensity score), and for patients who did not receive a transplant, weight equals (1 – 0.2252)/(1 – propensity score), where 0.2252 is the proportion of patients who received a transplant. A weighted logistic model was then built using 21-day survival as the outcome that is weighted by the calculated stabilized weight, and transplant is included as a covariate in the model. Model 4 is a Cox proportional hazards survival model, with time to death as the outcome and all subjects censored at 21 days if they were still alive. Transplant was also included as a time-varying covariate in this model, and the probability of surviving to 21 days was calculated. Model 5 is the missing data model which used 25 imputations to impute a hypothetical outcome for transplanted patients if they had not received a transplant, and then a logistic model was built using this imputed 21-day survival outcome.

For comparison purposes, all models were built using the same 5 predictors (defined in the methods section). Figure 1 illustrates the odds ratios (for models 1, 2, 3, and 5) and hazard ratios for the survival model (model 4) for each covariate in the multivariable model. There is minimal variability between models in the estimate of the odds ratio for each prognostic variable. We see this same trend in the hazard ratios which differ from the odds ratio because the survival model is modeling the instantaneous risk rather than the overall probability of survival when accounting for transplant.

In addition to estimates of the relationship between the outcome and covariate, we can also glean from each model a predicted probability (between 0 and 1) of the outcome. For clinical decision-making, a threshold value for the predicted probability of survival would generally be implemented. A receiver operating characteristic (ROC) curve illustrates performance of a model across all such thresholds and is therefore useful for model comparison. The ROC curve plots the model's false positive rate (1-specificity) on the x-axis against the sensitivity on the y-axis, and each point on the graph is generated by a different threshold for the decision.[17] The AUROC is a useful summary variable that ranges from 0 to 1, with larger values indicating better prediction accuracy. For the survival model (model 4), the AUROC is calculated using the specific time point of 21 days. Because model 1 defines the outcome differently than the other models (defining a poor outcome as death or transplant) and model 5 uses an imputed outcome for transplanted patients, we first examine model performance in the population that did not receive a transplant. Figure 2 illustrates the AUROC and ROC curves and shows that the 5 models

Table 1. Demographic, Baseline, and Outcome Variables for Subjects With Complete Baseline Covariates

Variables	Total n = 2100 n (%) or median (IQR)	Transplant n = 473 n (%) or median (IQR)	No transplant n = 1627 n (%) or median (IQR)
Age	40.0 (29.0–52.0)	39.0 (27.0–50.0)	40.0 (29.0–53.0)
Sex			
Female	1455 (69.3%)	329 (69.6%)	1126 (69.2%)
Male	644 (30.7)	144 (30.4%)	500 (30.8%)
Race			
White	1586 (75.6%)	324 (68.5%)	1262 (77.6%)
Black	306 (14.6%)	78 (16.5%)	228 (14.0%)
Asian	97 (4.6%)	33 (7.0%)	64 (3.9%)
Other	110 (5.2%)	38 (8.0%)	72 (4.4%)
Etiology			
Acetaminophen	963 (45.9%)	79 (16.7%)	884 (54.3%)
ALF in pregnancy	18 (0.9%)	1 (0.2%)	17 (1.0%)
Hepatitis A	37 (1.8%)	11 (17.6%)	26 (1.6%)
Shock/ischemia	123 (5.9%)	3 (0.6%)	120 (7.4%)
Autoimmune hepatitis	149 (7.1%)	83 (17.6%)	66 (4.1%)
Drug induced liver injury	229 (10.9%)	79 (16.7%)	150 (9.2%)
Indeterminate	259 (12.3%)	108 (22.8%)	151 (9.3%)
Other	322 (15.3%)	109 (23.0%)	213 (13.1%)
Coma grade[a]			
I or II	1122 (53.4%)	267 (56.5%)	855 (52.6%)
III or IV	978 (46.6%)	206 (43.6%)	772 (47.5%)
Vasopressor use[a]			
Yes	410 (19.5%)	67 (14.2%)	343 (21.1%)
No	1690 (80.5%)	406 (85.8%)	1284 (78.9%)
Bilirubin, mg/dL[a]	7.1 (3.7–19.8)	20.8 (9.8–28.0)	5.7 (3.2–13.1)
INR[a]	2.7 (2.0–4.1)	3.2 (2.3–4.7)	2.6 (1.9–3.9)
Model for End-Stage Liver Disease score	31.4 (24.3–38.1)	33.3 (27.4–39.6)	30.5 (27.4–39.6)
Outcome (21-d)			
Alive	1268 (60.4%)	377 (79.7%)	891 (54.8%)
Dead	615 (29.3%)	37 (7.8%)	578 (35.5%)
Unknown	217 (10.3%)	59 (12.5%)	158 (9.7%)

IQR, interquartile range.
[a]Coma grade, vasopressor use, bilirubin level, and INR all taken at hospital admission.

have quite similar model performance in the absence of transplant. Note that model 5 is similar to model 2 since no transplant cases are included. Figure 3 also illustrates the similarity between models when applied to the entire population. In this case, model 5 is using the mean AUROC of the 25 imputed data sets. We looked at model performance for the population listed for transplant and found similar results for models 2–5. The AUROC was reduced in model 1 due to the increased probability of receiving a transplant in the waitlist population as compared to the overall population and the fact that transplant is part of the outcome (see Tables A1 and A2).

In addition to model discrimination, Table 2 lists the performance aspects of positive and negative predictive values and predictive error and agreement in the overall population as well as only those that did not receive a transplant. We see similar performance aspects across the models with a slight inflation of the negative predictive value in model 1, which is driven by the inclusion of transplant as a poor outcome.

Discussion

The choice of analysis strategies is relatively vast in the prognostic modeling setting. Regression modeling is one approach to identify prognostic factors, and we examined the impact of different regression modeling approaches in the ALF setting. Our particular focus was how to handle transplant and if the various methods produce different results based on transplant as an outcome or as a covariate in the model. Our results indicate that the different models have similar performance. This is insightful as it provides researchers various analysis approaches and mitigates the concern of inflating model performance when transplant is part of the defined outcome. We did see a slight inflation in the negative predictive value when including transplant in the outcome definition, but this is not necessarily a sign of better model performance than the other models. The inflation is driven by the increase in the number of poor outcomes when transplant is included in the definition. The inclusion of transplant in the outcome definition acts as a

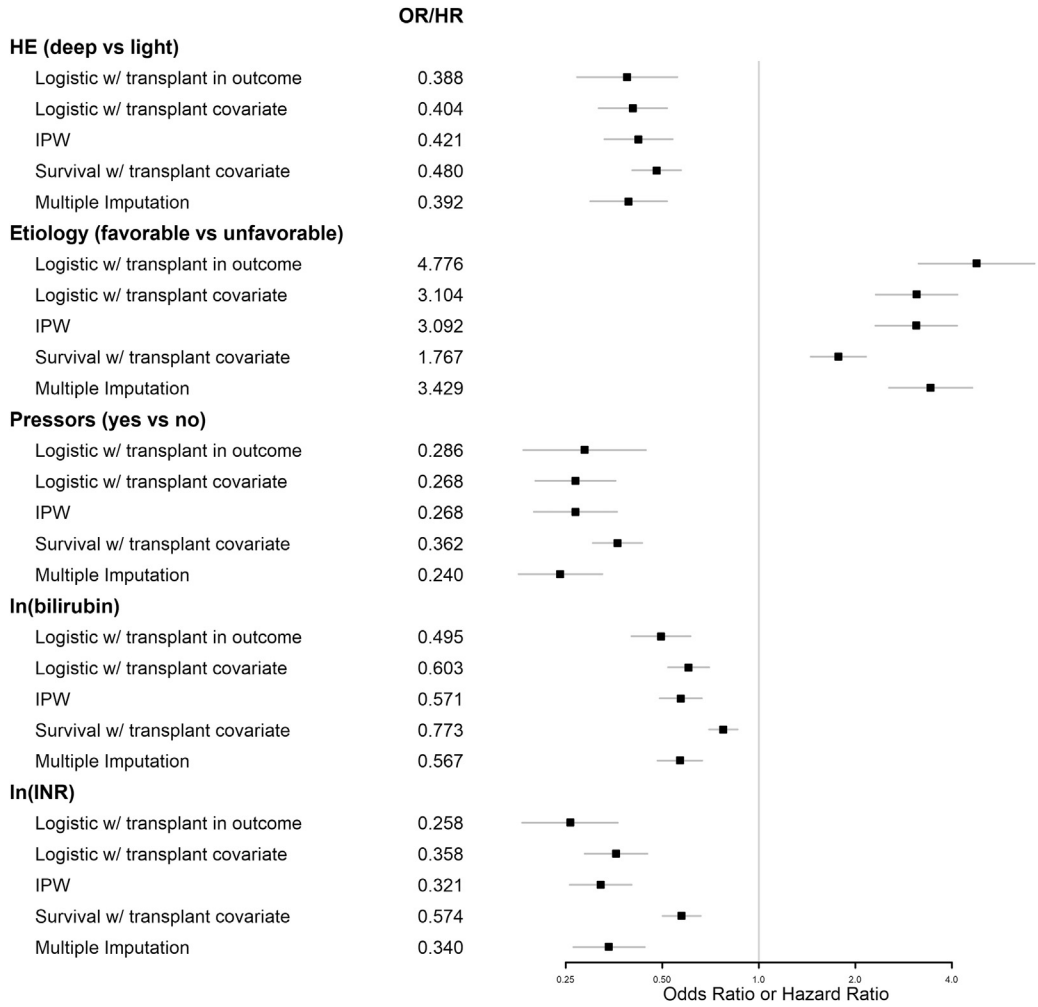

Figure 1. The forest plot of odds ratios (or hazard ratio) for survival given by the 5 constructed models.

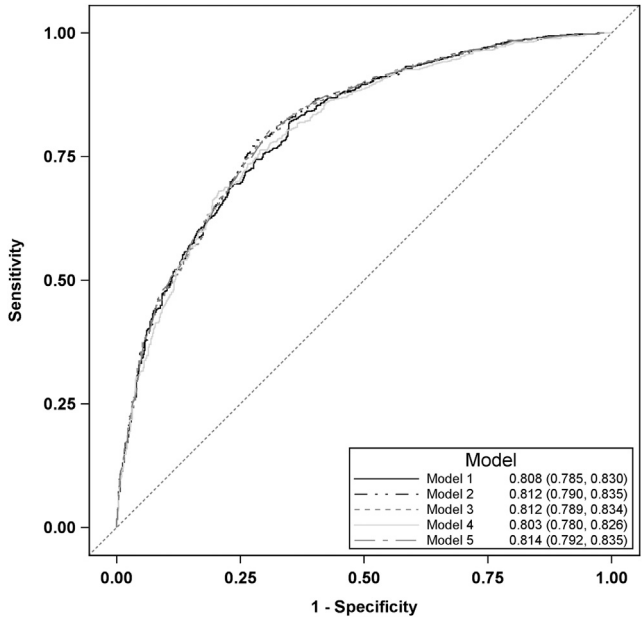

Figure 2. ROC curves for models applied to patients who did not receive a transplant.

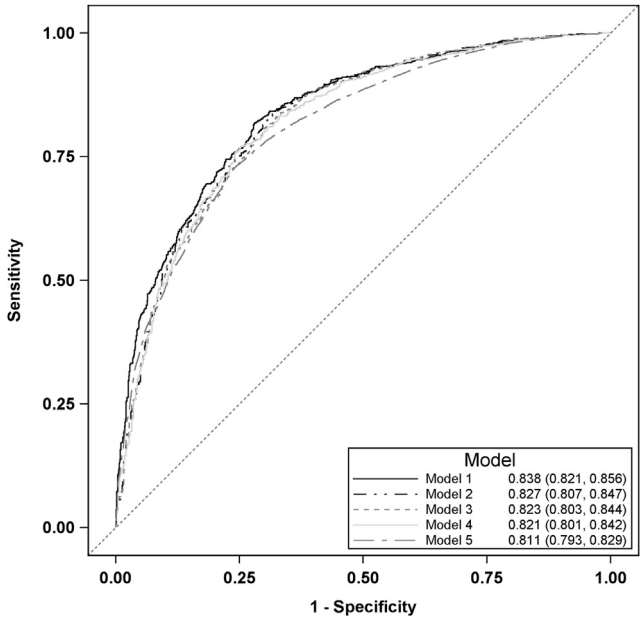

Figure 3. ROC curve for models applied to all patients.

Table 2. Model Performance: Predictive Error and Agreement, Overall and in the Nontransplant Population

Models[a]	AUROC (95% CI)		Brier score[b]		Calibration slope (*P*-value for test of intercept = 0 and slope = 1)[c]		Positive predictive value using an 80% threshold		Negative predictive value using an 80% threshold	
	All patients	Patients without transplant	All patients	Patients without transplant	All patients	Patients without transplant	All patients	Patients without transplant	All patients	Patients without transplant
Model 1 logistic	0.838 (0.821, 0.856)	0.808 (0.785, 0.820)	0.163	0.180	0.995 (0.55)	0.859 (<.001)	0.885	0.909	0.646	0.496
Model 2 logistic	0.827 (0.807, 0.847)	0.812 (0.790, 0.835)	0.152	0.170	1.000 (1.00)	1.081 (0.43)	0.903	0.897	0.502	0.516
Model 3 inverse probability	0.823 (0.803, 0.844)	0.812 (0.789, 0.834)	0.154	0.170	1.026 (0.28)	1.043 (0.75)	0.909	0.897	0.490	0.520
Model 4 survival	0.821 (0.801, 0.842)	0.803 (0.780, 0.826)	0.159	0.181	1.152 (<.001)	1.221 (<.001)	0.892	0.877	0.509	0.515
Model 5 multiple imputation	0.811 (0.793, 0.829)	0.814 (0.792, 0.835)	0.175	0.168	1.240 (<.001)	1.002 (0.97)	0.893	0.890	0.584	0.524

CI, confidence interval.
[a]The outcome for model 1 is defined as alive without transplant (vs death or transplant) at 21 d. The outcome is defined as alive at 21 d and transplant is a covariate in models 2, 3, and 4. For model 5, each performance measure was calculated for each of the 25 imputations and the mean of those 25 values reported.
[b]The Brier score is a measure of calibration: the mean squared difference between the predicted probability and the actual outcome. A lower Brier score indicates better calibration.
[c]A slope close to 1 and intercept close to 0 on a calibration plot indicate good overall calibration across the range of individuals or subgroups.

protective effect and provides a conservative approach to estimating the probability of survival without a transplant.

It is important to note that some of the examined models (models 2, 3, and 4) include transplant as a covariate, so if a clinician was using one of these models as a bedside tool to predict patient outcome, they would substitute 1 or 0 for the transplant covariate to predict survival of the patient with or without transplant, respectively. This could allow for the added insight as to how much a transplant improves a patient's probability of survival, but it does require understanding of the model and its interpretation. The other 2 models (models 1 and 5) do not contain a transplant covariate and only yield a prediction of survival without transplant.

The observed statistical models share similar findings, but there are some advantages to certain models worth highlighting (see Table 3). As noted previously, combining transplant and death in the outcome assumes that patients who receive a transplant would have died otherwise, treating both transplant and death as a poor outcome. Some have questioned the validity of this assumption; however, our results show that this model is comparable to the other models when the goal is to predict survival without a transplant. Model 1 is a simple logistic regression method, and it also allows for inclusion of patients who had a transplant even if their alive/dead outcome at 21 days was unknown. The alternative logistic model (model 2) with transplant as a covariate is also straightforward to implement and can provide the probability of surviving without a transplant as well as the probability of surviving with a transplant. Inverse probability weighting (model 3) is similar to the logistic model, but it requires an additional model-building step, including variable selection, to calculate propensity scores, and it is necessary to check balance for each covariate in the final model. For some data sets with significant imbalance between covariates in the transplant and nontransplant groups, this type of model might offer an advantage. The predictions of this model also depend on the accuracy of the propensity scores. In our data set, propensity scores were roughly 81% accurate in their prediction of transplant when using a 0.5 cutoff for the prediction. The survival model (model 4) requires time-to-event information that may not always be easy to obtain, but it does allow for inclusion (as right censored) of patients whose outcome is unknown. Because transplant is included as a time-varying covariate in the survival model, the resulting estimate for the baseline hazard should be used with caution. Survival models can offer more information than logistic models in regard to probability of survival at different time points, but they are more complicated to build, and in the acute liver setting where we are interested in the outcome at a specific time point (21 days from enrollment), a logistic model may be more appropriate, given the relatively short time period. Additionally, in the acute setting, time to transplant is a relatively short period, making the time-varying covariate less impactful. The survival modeling approach may be more appropriate in a chronic setting where time to transplant

Table 3. Attributes of Models 1–5

Model	Definition	Pros	Cons
Model 1—logistic regression	21-d TFS—composite death/ transplant	• Easy to implement • Easy to interpret • Transplant patients can be included in modeling even if 21-d survival is unknown	• Only models transplant-free survival.
Model 2—logistic regression	21-d survival—logistic model with transplant as a covariate	• Easy to implement. • Able to predict the outcome with or without transplant.	• None
Model 3—inverse probability weighting	21-d survival—logistic model with transplant as a covariate, using inverse probability weighting	• Provides balanced groups without a random sample. • The final model is easy to interpret. • Able to predict the outcome with or without transplant.	• Must choose variables to model the propensity score. • Slightly more difficult to implement. • Must check balance between groups.
Model 4—survival model	Time to death, with all subjects censored at 21 d—transplant included as a time-varying covariate	• Right censoring instead of missingness for subjects with unknown outcomes. • The time-to-event outcome might be of greater interest than a binary outcome at a single time point.	• Requires obtaining exact time of event. • Addition of a time-varying covariate can be difficult to implement.
Model 5—multiple imputation	21-d TFS—using the imputed outcome for transplanted patients for the outcome if they had not received a transplant.	• The final model is easy to interpret. • Good model fit. • Additional variables can be used to impute the missing outcome.	• More complex to implement than other models. • Use caution with high percentages of transplants.

and time to the outcome could be much longer. Also in the survival model setting, a strategy for handling events that compete with the primary outcome event is known as 'competing risk analysis',[18,19] and this strategy might be appropriate for some research questions involving transplant. For example, transplant is a competing risk to 'dying before transplant' in a study on waitlist mortality. However, in a setting such as ours where the interest is in predicting TFS, transplant and death are both events of interest and treating them as a composite outcome or including transplant as a time-varying covariate is more appropriate. Although the models were similar in performance, the MI model (model 5) had the lowest AUROC for the models we examined when applied to all patients. MI can be a complex process to implement and requires creation of multiple data sets to provide a summary measure of the predicted probability. Due to the high percent (22.5%) of transplant cases in the ALF setting, one should be cautious about using this method, as higher levels of missingness can lower the quality of the prediction.

This work was performed on data collected from a North American ALF patient population and focused on the 21-day outcome since it was the defined outcome of the registry protocol. Other settings may choose different time points or have different decision guidelines for transplant. This could alter the proportion of outcome events, but not necessarily the performance of the explored models. In addition, the performances of the 5 analysis methods are not limited to ALF. Other liver conditions that may require liver transplant, such as chronic liver diseases, could also find value in our results. The key consideration when choosing one of these methods is the clinical question. If the question is how the patient might improve without a transplant, then these models are relevant in that setting. Overall, we have shown that there is flexibility on handling transplant in this setting and the choice can be driven by preference in analysis methods. The simplest model for predicting TFS combines death and transplant as a single composite outcome, and this study shows that this model gives results comparable to more complex methods. However, if there is additional

interest in predicting survival with or without transplant, then transplant can also be included as a covariate. More complex models, with possible advantages shown in Table 3, might be chosen based on the nature of the data or the goal of the study. The important point is that the researchers clearly describe their approach including the outcome and handling of transplant.

Supplementary Materials

Material associated with this article can be found in the online version at https://doi.org/10.1016/j.gastha.2022.02.026.

References

1. Stravitz R, Lee W. Acute liver failure. Lancet 2019; 394:869–881.
2. Lee WM, Hynan LS, Rossaro L, et al. Intravenous N-acetylcysteine improves transplant-free survival in early stage non-acetaminophen acute liver failure [published correction appears in Gastroenterology. 2013 Sep;145(3):695. Dosage error in article text]. Gastroenterology 2009;137:856–864.e1.
3. Reuben A, Tillman H, Fontana RJ, et al. Outcomes in adults with acute liver failure between 1998 and 2013: an observational cohort study. Ann Intern Med 2016; 164:724–732.
4. Rangnekar AS, Ellerbe C, Durkalski V, et al. Quality of life is significantly impaired in long-term survivors of acute liver failure and particularly in acetaminophen-overdose patients. Liver Transpl 2013;19:991–1000.
5. O'Grady JG, Alexander GJ, Hayllar KM, et al. Early indicators of prognosis in fulminant hepatic failure. Gastroenterology 1989;97:439–445.
6. Zaman MB, Hoti E, Quasim A, et al. MELD score as a prognostic model for listing acute liver failure patients for liver transplantation. Transplant Proc 2006;38:2091–2098.
7. Koch DG, Tillman H, Durkalski V, et al. Development of a model to predict transplant-free survival of patients with acute liver failure. Clin Gastroenterol Hepatol 2016; 14:1199–1206.
8. Patidar KR, Davis BC, Slaven JE, et al. Admission factor V predicts transplant-free survival in acute liver failure. Dig Dis Sci 2021;66:619–627.
9. Cardoso FS, Gottfried M, Tujios S, et al. Continuous renal replacement therapy is associated with reduced serum ammonia levels and mortality in acute liver failure. Hepatology 2018;67:711–720.
10. Rutherford A, King LY, Hynan LS, et al. Development of an accurate index for predicting outcomes of patients with acute liver failure. Gastroenterology 2012; 143:1237–1243.
11. Pavlou M, Ambler G, Seaman SR, et al. How to develop a more accurate risk prediction model when there are few events. BMJ 2015;351:h3868.
12. Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. Stat Med 2015;34:3661–3679.
13. Crowson CS, Schenck LA, Green AB, et al. The basics of propensity scoring and marginal structural models (technical report #84). Rochester: Mayo Clinic, 2013:1–37.
14. Bradburn MJ, Clark TG, Love SB, et al. Survival analysis part II: multivariate data analysis – an introduction to concepts and methods. Br J Cancer 2003; 89:431–436.
15. Clark TG, Bradburn MJ, Love SB, et al. Survival analysis part IV: further concepts and methods in survival analysis. Br J Cancer 2003;89:781–786.
16. Groenwold RH, White IR, Donders AR, et al. Missing covariate data in clinical research: when and when not to use the missing-indicator method for analysis. CMAJ 2012;184:1265–1269.
17. Zhou X, McClish D, Obuchowski N. Statistical methods in diagnostic medicine. 2nd ed. Hoboken: Wiley, 2011:24–32.
18. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc 1999;94:496–509.
19. Gooley TA, Leisenring W, Crowley J, Storer BE. Estimation of failure probabilities in the presence of competing risks: new representations of old estimators. Stat Med 1999;18:695–706.

Received July 30, 2021. Accepted February 23, 2022.

Correspondence:
Address correspondence to: Sherry I. Livingston, MA, Department of Public Health Sciences, Medical University of South Carolina, 135 Cannon Street, Suite 303 MSC 835, Charleston, South Carolina 29425. e-mail: livingss@musc.edu.

Authors' Contributions:
Both named authors contributed to the conception and design of work, data interpretation, and final approval of the article version to be published. Sherry I. Livingston conducted the data analyses and wrote the original draft, while Dr Valerie Durkalski-Mauldin provided critical revision of the article.

Conflicts of Interest:
The authors disclose no conflicts.

Funding:
This study was supported in part by a grant from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) to the Acute Liver Failure Study Group (U01 DK058369).

Ethical Statement:
The corresponding author, on behalf of all authors, jointly and severally, certifies that their institution has approved the protocol for any investigation involving humans or animals and that all experimentation was conducted in conformity with ethical and humane principles of research.

Data Transparency Statement:
The NIH-funded ALFSG registry data will be made publicly available in the NIH data repository and can be requested by individuals from the NIH archived clinical data sets available online. The authors can share SAS analysis code upon request.