Received: 18 December 2018 | Revised: 21 May 2019 | Accepted: 27 June 2019

DOI: 10.1111/petr.13554

ORIGINAL ARTICLE

WILEY

Predicting ideal outcome after pediatric liver transplantation: An exploratory study using machine learning analyses to leverage Studies of Pediatric Liver Transplantation Data

Sharad Indur Wadhwani[1] | Evelyn K. Hsu[2] | Michele L. Shaffer[3] | Ravinder Anand[4] | Vicky Lee Ng[5] | John C. Bucuvalas[6]

[1]Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio

[2]University of Washington School of Medicine, Seattle Children's Hospital, Seattle, Washington

[3]University of Washington, Seattle, Washington

[4]EMMES Corporation, Rockville, Maryland

[5]Transplant and Regenerative Medicine Center, Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada

[6]Icahn School of Medicine at Mount Sinai, Kravis Children's Hospital, New York, New York

Correspondence
Evelyn K. Hsu, University of Washington School of Medicine, Seattle Children's Hospital, Seattle, WA.
Email: evelyn.hsu@seattlechildrens.org

Funding information
National Institute of Diabetes and Digestive and Kidney Diseases, Grant/Award Number: T32DK007727-24 (PI: Lee Denson, support for S.I.W.); Ashley's Angels Fund

Abstract

Machine learning analyses allow for the consideration of numerous variables in order to accommodate complex relationships that would not otherwise be apparent in traditional statistical methods to better classify patient risk. The SPLIT registry data were analyzed to determine whether baseline demographic factors and clinical/biochemical factors in the first-year post–transplant could predict ideal outcome at 3 years (IO-3) after LT. Participants who received their first, isolated LT between 2002 and 2006 and had follow-up data 3 years post–LT were included. IO-3 was defined as alive at 3 years, normal ALT (<50) or GGT (<50), normal GFR, no non-liver transplants, no cytopenias, and no PTLD. Heat map analysis and RFA were used to characterize the impact of baseline and 1-year factors on IO-3. 887/1482 SPLIT participants met inclusion criteria; 334 had IO-3. Demographic, biochemical, and clinical variables did not elucidate a visual signal on heat map analysis. RFA identified non-white race (vs white race), increased length of operation, vascular and biliary complications within 30 days, and duct-to-duct biliary anastomosis to be negatively associated with IO-3. UNOS regions 2 and 5 were also identified as important factors. RFA had an accuracy rate of 0.71 (95% CI: 0.68-0.74), PPV = 0.83, and NPV = 0.70. RFA identified participant variables that predicted IO-3. These findings may allow for better risk stratification and personalization of care following pediatric liver transplantation.

KEYWORDS
ideal outcome, machine learning, pediatric liver transplant

1 | INTRODUCTION

Pediatric recipients of liver transplantation have the potential to live full and productive lives. Transplant data agencies primarily monitor patient and graft survival,[1] and 1-year patient survival and 5-year patient survival are upward of 90% and 85%, respectively.[1] However, especially in pediatric patients, sustained allograft health without comorbidities or sequelae from long-term immunosuppression remains the ultimate goal.[2] Attaining over seven decades of comorbidity-free survival in pediatric liver transplant recipients will require directed research priorities.[3] Based on this premise, Ng, et

Abbreviations: ALT, alanine aminotransferase; GFR, glomerular filtration rate; GGT, gamma-glutamyltransferase; IO, ideal outcome; IO-3, ideal outcome at 3 years; IRB, institutional review board; LT, liver transplantation; NPV, negative predictive value; PELD, pediatric end-stage liver disease; PPV, Positive predictive value; PTLD, Post–transplant lymphoproliferative disease; RFA, random forests analysis; SPLIT, Studies in Pediatric Liver Transplantation; SRTR, Scientific Registry of Transplant Recipients; UNOS, United Network for Organ Sharing.

al[4] published the first comprehensive description of health status in the long-term follow-up of pediatric liver transplant recipients using the SPLIT database. This publication proposed the "ideal" long-term outcome as a composite concept, broadly defined as normal graft function and avoidance of immune and non-immune complications of immunosuppression. The authors found that only 32% of children met these criteria 10 years after transplant. However, this estimate is likely optimistic because it does not account for silent, immune-mediated allograft injury.[3]

After the first year of transplant, patients move into a chronic management phase[5] in which the goals shift from survival to sustained health without complications of therapy. Being able to stratify those at 1 year who are likely to have long-term complications could inform research and intervention efforts and allow for more direct personalization of care. The critical first step is to determine whether it is possible to predict those who are likely to attain the ideal outcome and conversely, those who are unlikely to achieve the ideal outcome. This approach, in line with the idea of precision medicine,[6] could allow for risk stratification and the more appropriate allocation of resources to those at highest risk for morbidity. A dynamic examination of many variables could identify targeted areas of focus for future research priorities, yet using traditional modeling techniques to study a multitude of variables simultaneously can lead to overfitting of the data.

In the quest to leverage the richness of registry data to define critical questions for research and targets for improvement, advanced analytics and intelligent techniques have emerged.[7,8] The application of these methodologies to transplant databases is emerging yet limited.[7,9,10] The SPLIT registry enrolls and prospectively follows patients under 18 years of age who were listed and received a liver transplant at over 40 institutions in the United States and Canada—making it the largest, international database of pediatric liver transplant recipients. The rich longitudinal data in the SPLIT registry offer the unique opportunity to utilize advanced analytics to define distinct phenotypes of survivors of pediatric LT. Previous work has identified variables that predict 6-month patient and graft survival.[11] However, using a data-driven approach to evaluate a multitude of variables to identify patients at risk for long-term morbidity could aid clinicians in minimizing their bias and in uncovering complex interactions among myriad variables.[12]

RFA[13-16] is a machine learning classifier that can uncover complex relationships among predictor variables to classify an observation (ie, participant) to an outcome while minimizing bias. Our objective was to use RFA to examine factors available at 1 year (the start of the chronic management phase of care) that predict IO of pediatric liver transplant recipients at 3 years. Our rationale is that machine learning could be a useful tool to identify those at risk for long-term complications following liver transplantation and move the field closer toward personalized care algorithms. Multiple predictor variables were used to identify novel predictors that may not have emerged with traditional statistical techniques. We hypothesized that novel predictors of IO would be identified using machine learning techniques.

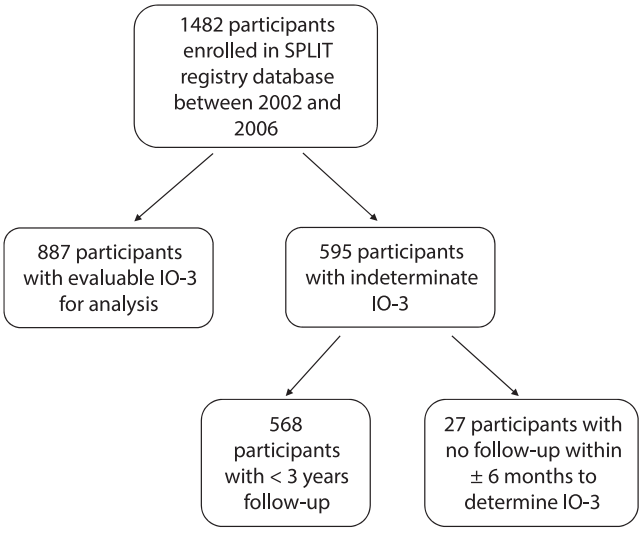

FIGURE 1 Participant inclusion/exclusion diagram

2 | PARTICIPANTS AND METHODS

2.1 | Participants

This project was reviewed and approved by the Seattle Children's Hospital Institutional Review Board (15444 NHS).

All centers obtained local IRB approval and informed consent prior to participant data collection and submission to SPLIT. De-identified data, including clinical, laboratory, operative, medical treatment, complications, and outcomes, were submitted to the SPLIT data coordinating center starting at the time of LT. The specific data collected are described elsewhere.[17]

Participants who received a liver transplant between February 2002 and 2006 were eligible for inclusion (n = 1482). This time period was chosen because it encompassed the period when PELD was implemented and when the registry was supported by the National Institutes of Health (2004-2009) and allowed for robust 3-year follow-up data. Specifically, centers were being compensated for participation; therefore, data were more in-depth and of higher quality. Eligible participants included subjects who had not received combined organ transplants, had not had a second transplant prior to the study period, and were within the study period. Of those, participants were excluded if they did not have complete data for the components of the IO at 3 years following transplant. Of 1482 eligible participants, 887 met inclusion criteria for the study. Figure 1 depicts participant inclusion/exclusion.

2.2 | Study design

This was a prospective cohort study using registry data from the SPLIT database. The predictor variables used were available at 1 year post–transplant. Predictor variables included demographic characteristics, allocation characteristics (ie, UNOS region, waitlist priority, and donor type), pretransplant health characteristics, peri-operative

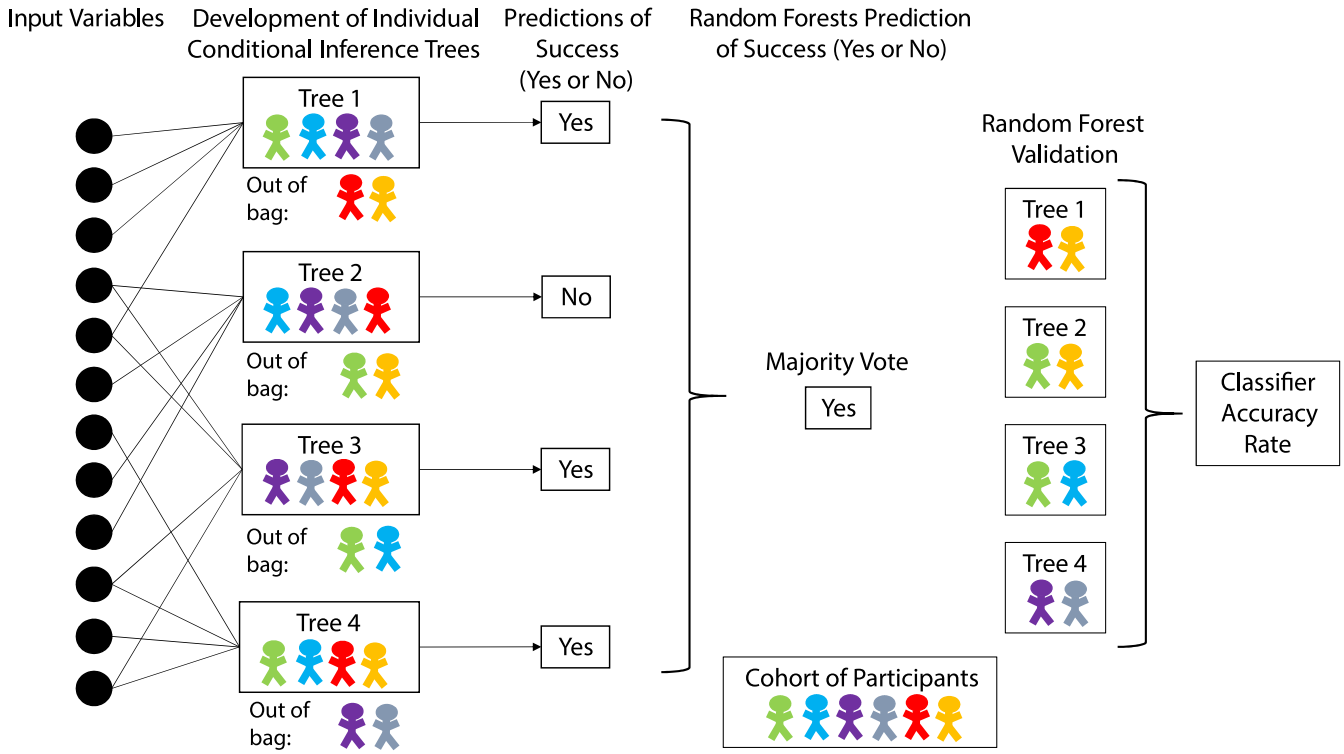

FIGURE 2 Example schematic of random forests analysis. In this example schematic, the classifier is predicting success (yes or no). A subset of participant variables from a subset of participants is used in each conditional inference tree to generate a prediction. The subsets of variables and participants can differ for each tree. Participants in the subset used to build a tree are in bag, and those outside of the subset are out-of-bag. Rather than splitting the data set once for training and then validation, as if often done with other methods, random forests analysis incorporates training and testing within individual trees by always holding some participants out (ie, out-of-bag). The number of variables used to build a tree is tuned because allowing use of all variables can limit generalizability due to overfitting. Each of the individual trees may have high bias for overfitting the data, yet in random forests analysis, the average of each of the individual trees is used to generate the classifier prediction. The classifier then uses out-of-bag error measurement to determine the accuracy rate by validating the classifier on participants (about a third of the total data) that were not used in building each tree

characteristics, and post–operative characteristics. All categorical variables were recoded to binary indicator variables with meaningful reference categories with the exception of recipient and donor blood type match which was recoded to a trinary variable (identical, compatible, and mismatch). For 6-month and 12-month complications, visits could have occurred within a window of ± 3 months. A total of 76 variables were included in the descriptive analysis, and 69 predictor variables were included in the random forests analysis.

The primary outcome was an IO at 3 years post–transplant (IO-3). IO-3 was modified from the original IO as defined by Ng, et al[4] due to data availability. For this study, the IO-3 composite was defined as: (a) alive with first allograft, (b) with "normal" liver tests (ALT < 50 IU/L and GGT < 50 IU/L), (c) with no PTLD, (d) with no non-liver transplants, (e) with no cytopenias, and (f) with normal GFR ascertained by Schwartz formula.[18] Data were obtained from follow-up visits occurring within a window of ± 6 months of the 3-year anniversary from LT surgery. Participants were classified as having (IO-3) or not having (non-IO-3) IO at 3 years post–transplant.

Two transplant hepatologists (VN, EH) reviewed and categorized open-text fields for diagnosis categories and complication categories. Discrepancies were adjudicated to arrive at a final classification.

2.3 | Heat map analysis

We first sorted participants by attaining IO-3 or not attaining IO-3. Predictor variables were added in a stepwise fashion and were categorized into demographic, allocation-related, pretransplant health status–related, peri-operative, and post–operative to generate a descriptive heat map stratified by IO-3 to determine whether there were discernible patterns of variables (phenotypes) associated with IO-3.

2.4 | Random forests analysis

Random forests,[13-15] using ensembles of conditional inference trees, were used to determine the importance of candidate variables in classifying participants as attaining or not attaining IO-3. Performance was measured via out-of-bag accuracy rate, positive predictive value, and negative predictive value. RFA uses multiple decision trees to generate a prediction (Figure 2). Decision trees have low bias; however, they tend to overfit the data provided, making them relatively unstable. In RFA, each decision tree utilizes a subset of the data and a subset of the variables to generate a prediction. The ensemble of

decision trees reduces the noise that is present within each tree and identifies complex interactions among the predictor variables, which further strengthens predictive ability. The accuracy was calculated using the out-of-bag method[15] by testing the classifier on about a third of the data. Importantly, these data were not used in generating specific decision trees. As a result, the accuracy rate reflects the predictive ability based on data the classifier has not encountered.

2.5 | Statistical analysis

Descriptive statistics were prepared for all variables including quartiles, means, standard deviations, and ranges for quantitative variables and frequencies and percentages for categorical variables to characterize the sample and assess for completeness. Participants were excluded if there was insufficient information to determine ideal outcome at 3 years. We compared demographic characteristics of the included and excluded participants to determine whether and how these participants differed using chi-square tests for categorical variables, and *t* test or Wilcoxon rank sum tests for quantitative variables. Similar analyses were conducted to determine whether and how participants who achieved IO-3 differed from those who did not. Participants with missing predictor information were maintained in the study as all analytic methods accommodate missing predictor data. Specifically, the random forests used handle missing data via surrogate variables. If a predictor variable is selected for the next split in a tree, observations that have a missing value in this variable are processed further down the tree using a surrogate variable that is not missing. The surrogate variable is selected such that it is the best predictor for the split in the originally chosen variable. Analyses were conducted using SAS version 9.4 (SAS Institute Inc.) and R version 3.0.3 (The R Foundation for Statistical Computing).

3 | RESULTS

3.1 | Participants

Table 1 displays demographic characteristics of the included and excluded participants, and the included participants are further stratified by achievement or failure to achieve IO-3. Comparative analyses were done to compare included/excluded participants and those with/without IO-3.

Included participants were more likely to have chronic liver disease/cholestasis as an indication for transplant, be below school age, be white race, and have received a partial organ transplant. Additionally, there was geographic variation in the participants who were included vs excluded with a higher proportion of participants from regions 2, 3, 4, 5, 6, 9, and 11 excluded. Of note, diagnosis, participant education level, and whether recipient received a whole organ have the same relative proportion of participants across included/excluded participant groups.

Of the participants included in the analysis, 553 of 887 (62%) failed to achieve IO-3. Participants with IO-3 were more likely to be from 2 parent households, white race, from UNOS regions 4, 5, 7, 10, and 12, and have a lower calculated PELD score at listing.

3.2 | Subgroup analysis of non–IO-3 participants

Participants who did not achieve IO were further analyzed to better understand the IO components that were not met at 3 years (Figure 3). The majority of participants who did not attain IO-3 had only one abnormal IO-3 component. The most likely components to be abnormal were elevated ALT or GGT and decreased GFR. Risk factors for each component are described elsewhere.[4]

3.3 | Heat map analysis

Figure 4 depicts the heat map and predictor variables included in the analysis. The 76 resulting variables are listed in Figure 4. No obvious visual signal was evident to any of the authors. In order to better uncover any complex relationships between the predictors and the outcome measure, machine learning techniques were used to attempt to predict IO-3.

3.4 | Random forests analysis (RFA)

RFA was used to develop a classifier for IO-3. The classifier had a predictive accuracy of 0.71 (95% CI: 0.68-0.74). The PPV was 0.83 (95% CI: 0.76-0.89), and the NPV was 0.70 (95%CI: 0.68-0.71). The naïve prediction classifier is 0.62 (ie, the prevalence of not attaining IO-3 is 0.62; so if the classifier predicted everyone to not have IO-3, the classifier would be correct in 0.62 of the instances).

Figure 5 depicts the relative variable importance in the classifier. Variable importance is a ranking of variables in their relative importance in the model for predicting IO-3. They do not provide any indication on the magnitude of the effect. In order of highest to lowest importance, variables predicting achievement of IO-3 include:

1. White race: Participants designated as white as opposed to non-white were predictive of achieving IO-3.
2. Length of operation (in hours): Shorter duration of transplant surgery was predictive of achieving IO-3.
3. UNOS region 2: Being from region 2 was predictive of not achieving IO-3.
4. UNOS region 5: Being from region 5 was predictive of achieving IO-3.
5. Vascular complications within 30 days of transplant: Absence of vascular complications within 30 days of transplant was predictive of achieving IO-3.
6. Pretransplant supplemental feedings: Absence of supplemental feedings pretransplant was predictive of achieving IO-3.
7. Biliary complications within 30 days of transplant: The absence of biliary complications within 30 days of transplant was predictive of achieving IO-3.
8. Biliary anastomosis: Roux limb was predictive of achieving IO-3.

TABLE 1 Demographic data by group (%, mean [SD], or median (IQR))

Variable	Included n = 887%	Excluded n = 595%	*P-Value*	No IO n = 553%	IO n = 334%	*P-Value*
Diagnosis			*P* < .01			*P* = .76
Cholestasis/chronic liver disease	64.3	58.2		62.9	66.5	
ALF	14.8	21.3		15.7	13.2	
Inborn error metabolism	9.1	10.8		9.0	9.3	
Dx tumor	9.2	6.4		9.4	9.0	
Dx other	2.6	3.4		2.9	2.1	
Missing	0.0	0.0		0.0	0.0	
Patient education level			*P* < .01			*P* = .23
Above grade level	0.0	0.7		0.0	0.0	
At grade level	21.6	27.2		22.6	20.1	
Below grade level	4.1	5.2		4.0	4.2	
Homeschooling	2.7	2.7		3.4	1.5	
Not at school age	70.3	64.2		68.4	73.7	
Missing	1.2	6.4		1.6	0.6	
Caretaker marriage status			*P* = .32			*P* < .05
Married/intact household	76.3	72.1		73.6	80.8	
Single-parent/non-intact household	22.3	27.9		25.1	17.7	
Missing	1.4	3.9		1.3%	1.5	
Race			*P* < .001			*P* < .01
White	63.5	60.3		61.3	67.1	
Non-white	15.2	25.7		18.4	9.9	
No race designated	21.3	13.9		20.3	23.1	
Missing	0.0	0.0		0.0	0.0	
Insurance			*P* = .36			*P* = .67
Public insurance	39.9	43.9		40.9	38.3	
Non-public insurance	51.5	56.1		51.4	51.8	
Missing	8.6	5.2		7.8	9.9	
UNOS region			*P* < .001			*P* < .001
Region 1	2.3	1.3		2.5	1.8	
Region 2	11.3	19.5		13.7	7.2	
Region 3	9.5	13.1		11.6	6.0	
Region 4	8.1	11.1		6.5	10.8	
Region 5	12.0	16.0		9.8	15.6	
Region 6	0.3	1.3		0.5	0.0	
Region 7	12.5	8.4		11.8	13.8	
Region 8	14.1	7.6		14.8	12.9	
Region 9	3.5	3.7		4.2	2.4	
Region 10	14.9	6.4		13.6	17.1	
Region 11	2.1	2.2		2.4	1.8	
Region 12	9.5	9.4		8.7	10.8	
Missing	0.0	0.0		0.0	0.0	
Wait time (d)	54 (140.5)	65 (166.0)	*P* = .10	52.0 (132.8)	57.5 (145.8)	*P* = .36

(Continues)

TABLE 1 (Continued)

Variable	Included n = 887%	Excluded n = 595%	*P-Value*	No IO n = 553%	IO n = 334%	*P-Value*
Status 1a/1b						
Yes	29.8	26.9	*P* = .28	31.6	26.6	*P* = .16
No	61.4	74.3		60.6	62.9	
Missing	12.5	4.2		11.4	14.4	
PELD	12.7 [14.5]	14.1 [14.4]	*P* = .11	13.6 [14.8]	11.3 [14.0]	*P* = .02
Donor age (y)	11 (18)	13 (22)	*P* = .21	12 (19)	10 (18)	*P* = .23
Organ type			*P* = .43			*P* = .92
Cadaveric	87.9	85.9		88.2	87.4	
Living-related donor	11.3	12.4		11.0	11.7	
Living-unrelated donor	0.8	1.7		0.7	0.9	
Missing	0.0	0.3		0.0	0.0	
Whole organ			*P* < .05			*P* = .22
Yes	53.3	58.3		55.3	50.0	
No	44.4	41.7		43.0	46.7	
Missing	2.3	4.7		1.6	3.3	

Note: Missing data were not included in statistical comparisons across groups.
Abbreviations: ALF, acute liver failure; dx, diagnosis; IO, ideal outcome; IQR, interquartile range; PELD, pediatric end-stage liver disease; SD, standard deviation.

4 | DISCUSSION

As the field of pediatric liver transplantation has evolved, we seek to personalize treatment strategies to optimize outcome and value. As patients move through the pathway of selection, waitlist management, peri-transplant, and post–transplant, we asked if it was possible to predict who at 1 year will have the IO-3 at 3 years. As a proof of concept and to move the field closer to personalized medicine, we utilized intelligent techniques to leverage available data in the SPLIT database. RFA allows for a completely different evaluation of registry data that is not traditionally employed. While no obvious signal was evident on the heat map, RFA allowed us to identify several variables for predicting IO-3 at 3 years. The RFA classifier had an accuracy of 0.71 which exceeds that of the naïve prediction classifier. Furthermore, accuracy was assessed using the out-of-bag method, which supports development of a robust classifier. This methodology, over traditional statistical approaches, allows for the flexible use of multiple variables to aid in *classification* as opposed to uncovering the specific relationship between predictor and outcome variables, while simultaneously minimizing the risk of overfitting the data.

Notably, non-white race was the variable of most importance and predictive of *not* attaining the IO-3. This was the variable of most importance despite including a variable for insurance (public or private,

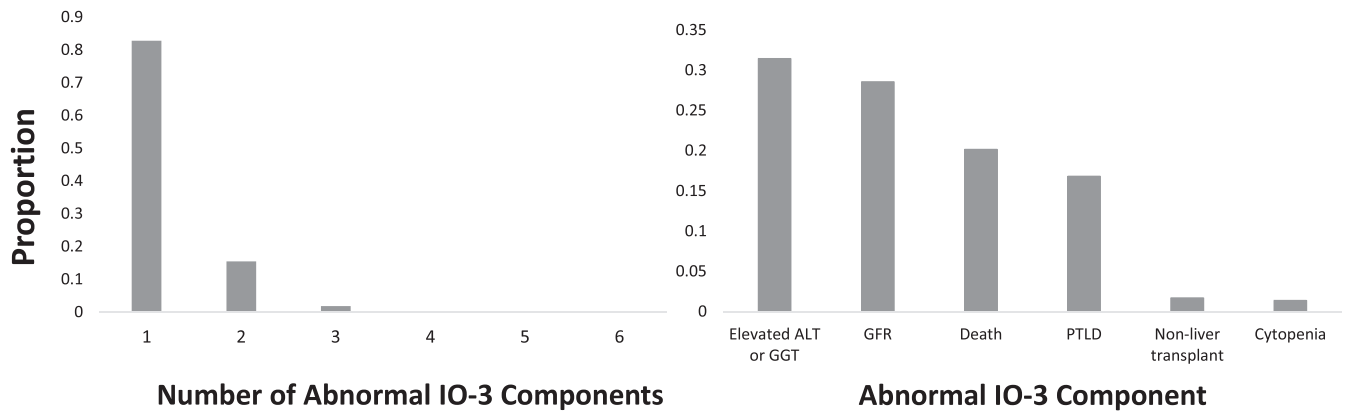

FIGURE 3 Reason(s) for failure to achieve IO-3 profile by A) number of abnormal IO-3 component variables for participants not achieving IO-3 and B) frequency of component variables for participants not achieving IO-3. ALT, alanine aminotransferase; GGT, gamma-glutamyltransferase; PTLD, post–transplant lymphoproliferative disease

a marker of socioeconomic status) in the RFA. The reasons for this are likely multifactorial and may reflect systemic bias (both societal and health system) experienced by minorities across the phases of care. Previous studies in pediatric LT have suggested a disparity in access to care, presentation to care, and waitlist priority for patients of non-white race.[19-22] Black adults who received kidney transplants were found to have increased prevalence of cardiac disease pretransplant, which could suggest that comorbidities may contribute to this racial gap.[23] Furthermore, this finding could reflect differential socioeconomic backgrounds and immunosuppressant adherence rates.[24] Our findings support the case for continued attention in racial disparity and equity in access to care for children within our system.

Pretransplant supplemental feedings and length of operation may reflect the severity of illness prior to transplant. Length of operation has not been previously implicated in affecting ideal outcomes. Increasing length of operation may reflect the stability, and health of patient prior to transplantation, graft type, may be associated with previous operations, difficulty in explant hepatectomy, transplant center, or likelihood of open abdomen and delayed closure. It is likely a surrogate marker for complexity of patient, but perhaps prior planning could decrease the importance of this variable. Its relationship to overall outcomes is intriguing and introduces the idea of establishing a time threshold or benchmark in order to improve overall outcomes.

Participant's UNOS region was also found to be predictive of IO-3. UNOS divides the United States into 11 different regions for purposes of organ allocation. Notably, participants from region two (Delaware, Maryland, New Jersey, Pennsylvania, West Virginia and Washington DC) were less likely to have the IO and participants from region five (Arizona, California, Nevada, Utah, and New Mexico) were more likely to have IO. Once again, the reasons for this are likely multifactorial and may reflect organ availability, institutional differences, and patient-level characteristics. However, SPLIT does not have equal representation across UNOS regions, and as highlighted in Table 1, there were differential inclusion/exclusion rates across UNOS regions. Therefore, definitive conclusions cannot be drawn due to selection bias.

The majority of participants without IO-3 had abnormal liver enzymes. Our estimates were based on an ALT and GGT cutoff

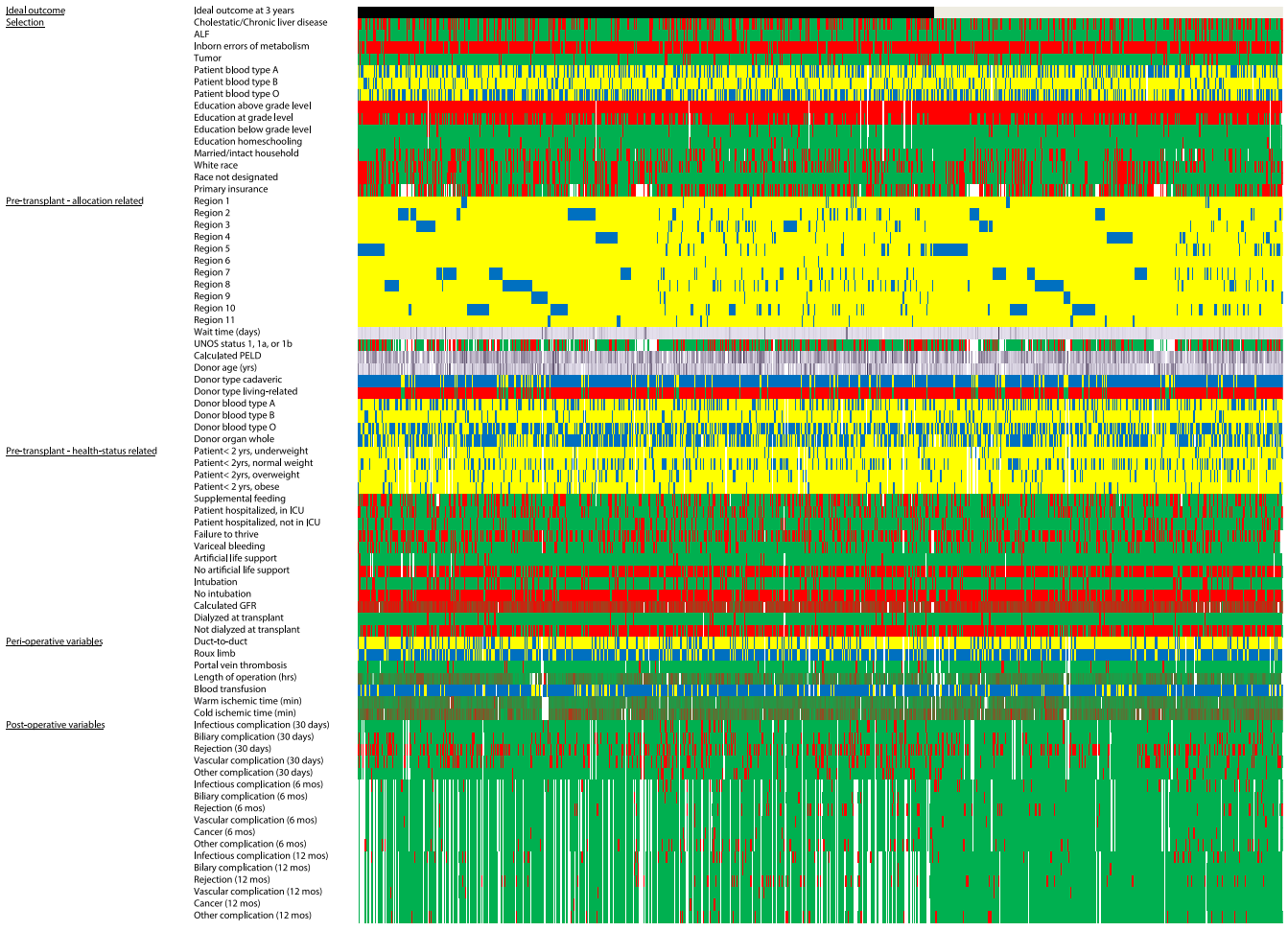

FIGURE 4 Descriptive heat map of predictor variables available at 1 y and ideal outcome Legend: Ideal outcome at 3 y: Black indicates participants who did not meet definition of IO-3; light gray indicates participants who did meet definition of IO-3. Green/red: green—favorable; red—unfavorable; Blue/yellow: blue—yes; yellow—no; Purple—continuous variable; higher value is depicted with deeper purple; ALF, acute liver failure; UNOS, United Network for Organ Sharing; ICU, intensive care unit; GFR, glomerular filtration rate; hrs, hours; min, minutes; mos, months

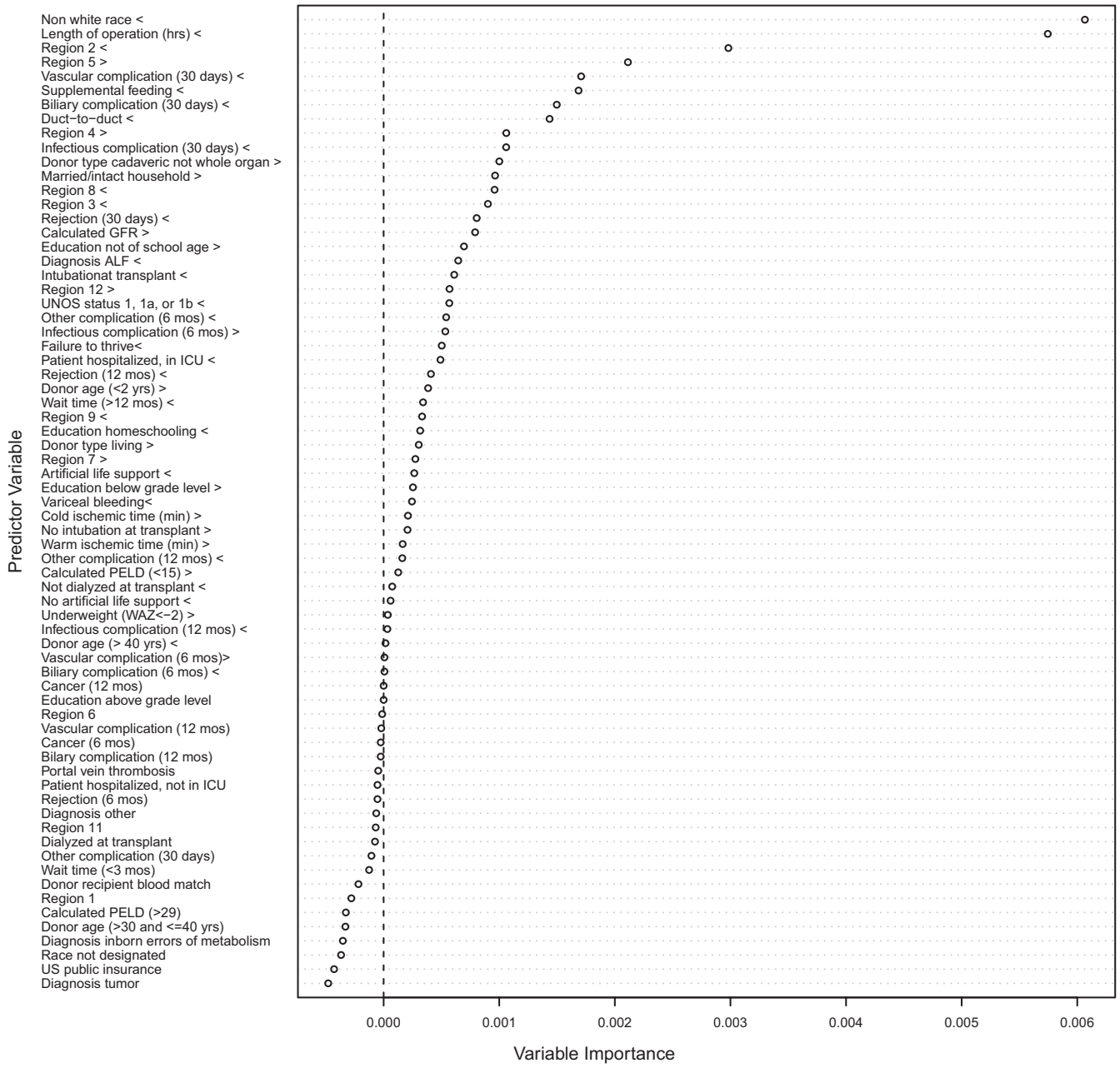

FIGURE 5 Random forests analysis ranking of variable importance Legend: <signifies that the variable predicts no ideal outcome; >signifies that the variable predicts ideal outcome. Hrs, hours; GFR, glomerular filtration rate; ALF, acute liver failure; UNOS, United Network for Organ Sharing; ICU, intensive care unit; min, minutes; yrs, years; mos, months; PELD, Pediatric end-stage liver disease; US, United States

of 50, which likely underestimates the incidence of patients with ongoing inflammation.[25] This is further supported by data on patients ineligible for participation in a multicenter immunosuppression withdrawal trial due to silent immune-mediated liver injury despite appearing clinically stable.[3,26] Kidney injury is the second most frequent complication and likely reflects, at least in part, pre-existing kidney disease, episodes of acute kidney injury and non-immune complications of immunosuppressive medications. These findings further support the need for research strategies that personalize immunosuppression and optimize allograft health without complications.

This study has several limitations. Notably, there were differences in the demographic characteristics of participants who had IO-3 data available and those who did not. This may bias findings from the analyses. Furthermore, there were baseline differences in those who had the IO-3 and those who did not. Specifically, participants without IO-3 were more likely to come from households without intact marriages, be of non-white race, and have higher calculated PELD score. There are also limitations to RFA. RFA is able to determine the relative importance of variables; however, it is unable to ascertain the direct relationships. This is both a strength and limitation of RFA. In one regard, RFA is able to ascertain complex relationships between

variables to strengthen predictive ability. Conversely, because of this, it is difficult to determine the specific individual relationships among predictors and outcomes. Finally, limitations common to registry studies such as representative sample, missing data, and quality of the data apply to this study. Specifically, the SPLIT registry is not a mandatory reporting database of all LT recipients (like UNOS) so it may not represent the entire population of transplanted US children.

Long-term morbidity is significantly increased for LT recipients who survive the first year compared with age-matched controls in the general population. Pediatric recipients compared to adult recipients face increased risk of morbidity given their potential for longer life expectancy and therefore increased likelihood for longer cumulative exposure to immunosuppression. The challenge remains to ensure optimal allograft health and functional outcomes, while striving to minimize the complications of immunosuppression. Interestingly, in our cohort, only 38% attained IO-3, while Ng, et al[4] found that 32% of participants attained IO at 10 years post–transplant. This suggests that patients are at highest risk of morbidity in the few years immediately following transplant. Being able to identify subgroups of pediatric LT recipients who require additional care could unlock targeted interventions for those at highest risk and prevent morbidity such as retransplantation, given the significant cost estimates of retransplantations being upward of $300 000. Conversely, if we can predict from variables at 1 year who is likely to have long-term success, it may allow resources to be targeted toward those at higher risk. However, future work is needed to identify what predictive variables are modifiable and whether that affects the long-term outcomes of pediatric LT survivors. This is aligned with the national push for precision medicine and a newer concept—precision public health.[27,28] Precision public health "can be simply viewed as providing the right intervention to the right population at the right time".[27] This study sought to use machine learning algorithms to better predict who is at risk for not attaining the IO-3. The authors hope this will catalyze future research that ultimately leads to greater personalization of care for pediatric transplant recipients.

ACKNOWLEDGMENTS

The research was supported by NIH T32DK007727-24 (PI: Lee Denson; support for SIW) and supported in part by the Ashley's Angels Fund. The authors would like to thank the SPLIT centers in the registry from where these data came and Cole Brokamp, PhD, for reviewing aspects of the manuscript.

CONFLICT OF INTEREST

The authors of this manuscript have no conflicts of interest to disclose as described by Pediatric Transplantation.

AUTHOR CONTRIBUTIONS

EKH, VLN, and JCB: Designed study; EKH, RA, and VLN: Performed data acquisition; SIW, EKH, MLS, RA, VLN, and JCB: Analyzed and interpreted the data; SIW and EKH: Drafted the manuscript; SIW, EKH, MLS, RA, VLN, and JCB: Revised the manuscript.

ORCID

Sharad Indur Wadhwani https://orcid.org/0000-0001-5565-6670
Vicky Lee Ng https://orcid.org/0000-0002-9998-5692

REFERENCES

1. Kim WR, Lake JR, Smith JM, et al. OPTN/SRTR 2016 annual data report: liver. *Am J Transplant*. 2018;18(Suppl 1):172-253.
2. Porter ME. What is value in health care? *N Engl J Med*. 2010;363(26):2477-2481.
3. Feng S, Bucuvalas J. Tolerance after liver transplantation: where are we? *Liver Transpl*. 2017;23(12):1601-1614.
4. Ng VL, Alonso EM, Bucuvalas JC, et al. Health status of children alive 10 years after pediatric liver transplantation performed in the US and Canada: report of the studies of pediatric liver transplantation experience. *J Pediatr*. 2012;160(5):820.e3-826.e3.
5. Wagner EH. Chronic disease management: what will it take to improve care for chronic illness? *Eff Clin Pract*. 1998;1(1):2-4.
6. Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med*. 2015;372(9):793-795.
7. Nilsson J, Ohlsson M, Hoglund P, Ekmehag B, Koul B, Andersson B. The international heart transplant survival algorithm (IHTSA): a new model to improve organ sharing and survival. *PLoS ONE*. 2015;10(3):e0118644.
8. Nilsson J, Ohlsson M, Thulin L, Hoglund P, Nashef SA, Brandt J. Risk factor identification and mortality prediction in cardiac surgery using artificial neural networks. *J Thorac Cardiovasc Surg*. 2006;132(1):12-19.
9. Sousa FS, Hummel AD, Maciel RF, et al. Application of the intelligent techniques in transplantation databases: a review of articles published in 2009 and 2010. *Transplant Proc*. 2011;43(4):1340-1342.
10. Srinivas TR, Taber DJ, Su Z, et al. Big data, predictive analytics, and quality improvement in kidney transplantation: a proof of concept. *Am J Transplant*. 2017;17(3):671-681.
11. McDiarmid SV, Anand R, Martz K, Millis MJ, Mazariegos G. A multivariate analysis of pre-, peri-, and post–transplant factors affecting outcome after pediatric liver transplantation. *Ann Surg*. 2011;254(1):145-154.
12. Chawla NV, Davis DA. Bringing big data to personalized healthcare: a patient-centered framework. *J Gen Intern Med*. 2013;28(Suppl 3):660-665.
13. Strobl C, Malley J, Tutz G. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychol Methods*. 2009;14(4):323-348.
14. Strobl C, Boulesteix AL, Kneib T, Augustin T, Zeileis A. Conditional variable importance for random forests. *BMC Bioinformatics*. 2008;9:307.
15. Breiman L. Random forests. *Mach Learn*. 2001;45(1):5-32.
16. Donges N. The Random Forest Algorithm – Towards Data Science. 2018; https://towardsdatascience.com/the-random-forest-algorithm-d457d499ffcd. Accessed June 19, 2018.
17. McDiarmid SV, Anand R, Lindblad AS, SPLIT Research Group. Studies of pediatric liver transplantation: 2002 update. An overview of demographics, indications, timing, and immunosuppressive practices in pediatric liver transplantation in the United States and Canada. *Pediatr Transplant*. 2004;8(3):284-294.

18. Schwartz GJ, Haycock GB, Edelmann CM Jr, Spitzer A. A simple estimate of glomerular filtration rate in children derived from body length and plasma creatinine. *Pediatrics*. 1976;58(2):259-263.
19. Braun HJ, Perito ER, Dodge JL, Rhee S, Roberts JP. Nonstandard exception requests impact outcomes for pediatric liver transplant candidates. *Am J Transplant*. 2016;16(11):3181-3191.
20. Hsu EK, Bucuvalas J. The trouble with exceptional exceptions. *Am J Transplant*. 2016;16(11):3073-3074.
21. Hsu EK, Shaffer M, Bradford M, Mayer-Hamblett N, Horslen S. Heterogeneity and disparities in the use of exception scores in pediatric liver allocation. *Am J Transplant*. 2015;15(2):436-444.
22. Thammana RV, Knechtle SJ, Romero R, Heffron TG, Daniels CT, Patzer RE. Racial and socioeconomic disparities in pediatric and young adult liver transplant outcomes. *Liver Transpl*. 2014;20(1):100-115.
23. Taber DJ, Douglass K, Srinivas T, et al. Significant racial differences in the key factors associated with early graft loss in kidney transplant recipients. *Am J Nephrol*. 2014;40(1):19-28.
24. Shemesh E, Bucuvalas JC, Anand R, et al. The medication level variability index (MLVI) predicts poor liver transplant outcomes: a prospective multi-site study. *Am J Transplant*. 2017;17(10):2668-2678.
25. Schwimmer JB, Dunn W, Norman GJ, et al. SAFETY study: alanine aminotransferase cutoff values are set too high for reliable detection of pediatric chronic liver disease. *Gastroenterology*. 2010;138(4):1357-1364.e1352.
26. Feng S, Bucuvalas JC, Demetris AJ, et al. Evidence of chronic allograft injury in liver biopsies from long-term pediatric recipients of liver transplants. *Gastroenterology*. 2018.
27. Khoury MJ, Iademarco MF, Riley WT. Precision public health for the era of precision medicine. *Am J Prev Med*. 2016;50(3):398-401.
28. Bayer R, Galea S. Public health in the precision-medicine era. *N Engl J Med*. 2015;373(6):499-501.

How to cite this article: Wadhwani SI, Hsu EK, Shaffer ML, Anand R, Ng VL, Bucuvalas JC. Predicting ideal outcome after pediatric liver transplantation: An exploratory study using machine learning analyses to leverage Studies of Pediatric Liver Transplantation Data. *Pediatr Transplant*. 2019;23:e13554. https://doi.org/10.1111/petr.13554