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

Authors
Wadhwani, Sharad Indur
Hsu, Evelyn K
Shaffer, Michele L
et al.

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

Sharad Indur Wadhwani, M.D. M.P.H.,
Cincinnati Children's Hospital Medical Center, Cincinnati, OH

Evelyn K. Hsu, M.D.,
University of Washington School of Medicine, Seattle Children's Hospital, Seattle, WA

Michele L. Shaffer, Ph.D.,
University of Washington, Seattle, WA

Ravinder Anand, Ph.D.,
EMMES Corporation, Rockville, MD

Vicky Lee Ng, M.D.,
Hospital for Sick Children, Transplant and Regenerative Medicine Center, University of Toronto, Toronto, Canada

John C. Bucuvalas, M.D.
Icahn School of Medicine at Mount Sinai, Kravis Children's Hospital New York, NY

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 Studies of Pediatric Liver Transplantation (SPLIT) registry data was 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)

Corresponding author evelyn.hsu@seattlechildrens.org.
Author Contributions:

Author	Study Design	Data acquisition	Data Analysis & Interpretation	Drafting Manuscript	Revising manuscript
SIW			X	X	X
EKH	X	X	X	X	X
MLS			X		X
RA		X	X		X
VLN	X	X	X		X
JCB	X		X		X

after liver transplantation (LT). Participants who received their first, isolated LT between 2002–2006 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 glomerular filtration rate (GFR), no non-liver transplants, no cytopenias and no post-transplant lymphoproliferative disease (PTLD). Heat map and random forests analyses (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, 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

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 one- and five-year patient survival are upwards 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 7 decades of comorbidity-free survival in pediatric liver transplant recipients will require directed research priorities.[3] Based on this premise, Ng, et al.[4] published the first comprehensive description of health status in the long-term follow-up of pediatric liver transplant recipients using the Studies in Pediatric Liver Transplantation (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 one 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 if it is possible to predict those 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 are 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 offers 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]

Random forests analysis (RFA)[13–16] is a machine learning classifier that can uncover complex relationships amongst predictor variables to classify an observation (i.e. participant) to an outcome while minimizing bias. Our objective was to use RFA to examine factors available at one year (the start of the chronic management phase of care) that predict "ideal outcome" (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 towards 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.

2. Participants & 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 institutional review board (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 pediatric end-stage liver disease (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 "ideal outcome" 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 (i.e. United Network for Organ Sharing (UNOS) region, waitlist priority and donor type), pre-transplant health characteristics, peri-operative 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, 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 Ideal Outcome (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: 1) alive with first allograft, with 2) "normal" liver tests (ALT < 50 IU/L and GGT < 50 IU/L), 3) no post-transplant lymphoproliferative disorder (PTLD) 4) no non-liver transplants 5) no cytopenias **and** 6) normal glomerular filtration rate (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, pre-transplant health status related, peri-operative and post-operative to generate a descriptive heat map stratified by IO-3 to determine if 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 over fit 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 reduce the noise that is present within each tree and identifies complex interactions amongst 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., Cary, NC, USA) and R Version 3.0.3 (The R Foundation for Statistical Computing, Vienna, Austria).

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 versus 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, 553/887 (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 alanine aminotransferase (ALT) or gamma-glutamyl transferase (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 positive predictive value (PPV) was 0.83 (95%CI: 0.76–0.89) and the negative predictive value (NPV) was 0.70 (95%CI: 0.68–0.71). The naïve prediction classifier is 0.62 (i.e. 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 was 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. Pre-transplant supplemental feedings: Absence of supplemental feedings pre-transplant 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.

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, wait list 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; a marker of socio-economic status) in the RFA. The reasons for this are likely multi-factorial 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 pre-transplant which could suggest that comorbidities may contribute to this racial gap.[23] Furthermore, this finding could reflect differential socio-economic 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.

Pre-transplant 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 2 (Delaware, Maryland, New Jersey, Pennsylvania, West Virginia and Washington DC) were less likely to have the IO and participants from region 5 (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 cut-off of 50, which likely under-estimates the incidence of patients with ongoing inflammation.[25] This is further supported by data on patients ineligible for participation in a multi-center 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. 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 U.S. 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 thereby 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 re-transplantation, given the significant cost estimates of re-transplantations being upwards 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 towards 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 lead to greater personalization of care for pediatric transplant recipients.

Acknowledgments:

Abbreviations:

ALT	alanine aminotransferase
GGT	gamma-glutamyl transferase
IO	Ideal Outcome
IO-3	Ideal Outcome at 3 years
IRB	Institutional Review Board
LT	Liver transplantation
PELD	Pediatric end-stage liver disease
SPLIT	Studies in Pediatric Liver Transplantation
SRTR	Scientific Registry of Transplant Recipients
UNOS	United Network for Organ Sharing
PTLD	Post-transplant lyphoproliferative disease
RFA	Random Forests Analysis

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. [PubMed: 29292603]

2. Porter ME. What is value in health care? N Engl J Med. 2010;363(26):2477–2481. [PubMed: 21142528]

3. Feng S, Bucuvalas J. Tolerance after liver transplantation: Where are we? Liver Transpl. 2017;23(12):1601–1614. [PubMed: 28834221]

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):820–826 e823. [PubMed: 22192813]

5. Wagner EH. Chronic disease management: what will it take to improve care for chronic illness? Effective clinical practice : ECP. 1998;1(1):2–4. [PubMed: 10345255]

6. Collins FS, Varmus H. A new initiative on precision medicine. N Engl J Med. 2015;372(9):793–795. [PubMed: 25635347]

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. The Journal of thoracic and cardiovascular surgery. 2006;132(1):12–19. [PubMed: 16798296]

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. [PubMed: 21620124]

10. Srinivas TR, Taber DJ, Su Z, et al. Big Data, Predictive Analytics, and Quality Improvement in Kidney Transplantation: A Proof of Concept. American Journal of Transplantation. 2017;17(3):671–681. [PubMed: 27804279]

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. Annals of surgery. 2011;254(1):145–154. [PubMed: 21606838]

12. Chawla NV, Davis DA. Bringing Big Data to Personalized Healthcare: A Patient-Centered Framework. Journal of general internal medicine. 2013;28(Suppl 3):660–665. [PubMed: 23225256]

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. [PubMed: 19968396]

14. Strobl C, Boulesteix AL, Kneib T, Augustin T, Zeileis A. Conditional variable importance for random forests. BMC Bioinformatics. Vol 92008:307.

15. Breiman L. Random forests. Machine Learning. 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 6/19/2018, 2018.

17. McDiarmid SV, Anand R, Lindblad AS, Group SR. 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. [PubMed: 15176967]

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. [PubMed: 951142]

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. [PubMed: 27214757]

20. Hsu EK, Bucuvalas J. The Trouble With Exceptional Exceptions. Am J Transplant. 2016;16(11):3073–3074. [PubMed: 27305346]

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. [PubMed: 25612496]

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. [PubMed: 24136785]

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. [PubMed: 24969370]

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. [PubMed: 28321975]

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. [PubMed: 20064512]

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. [PubMed: 26547538]

28. Bayer R, Galea S. Public Health in the Precision-Medicine Era. New England Journal of Medicine. 2015;373(6):499–501.
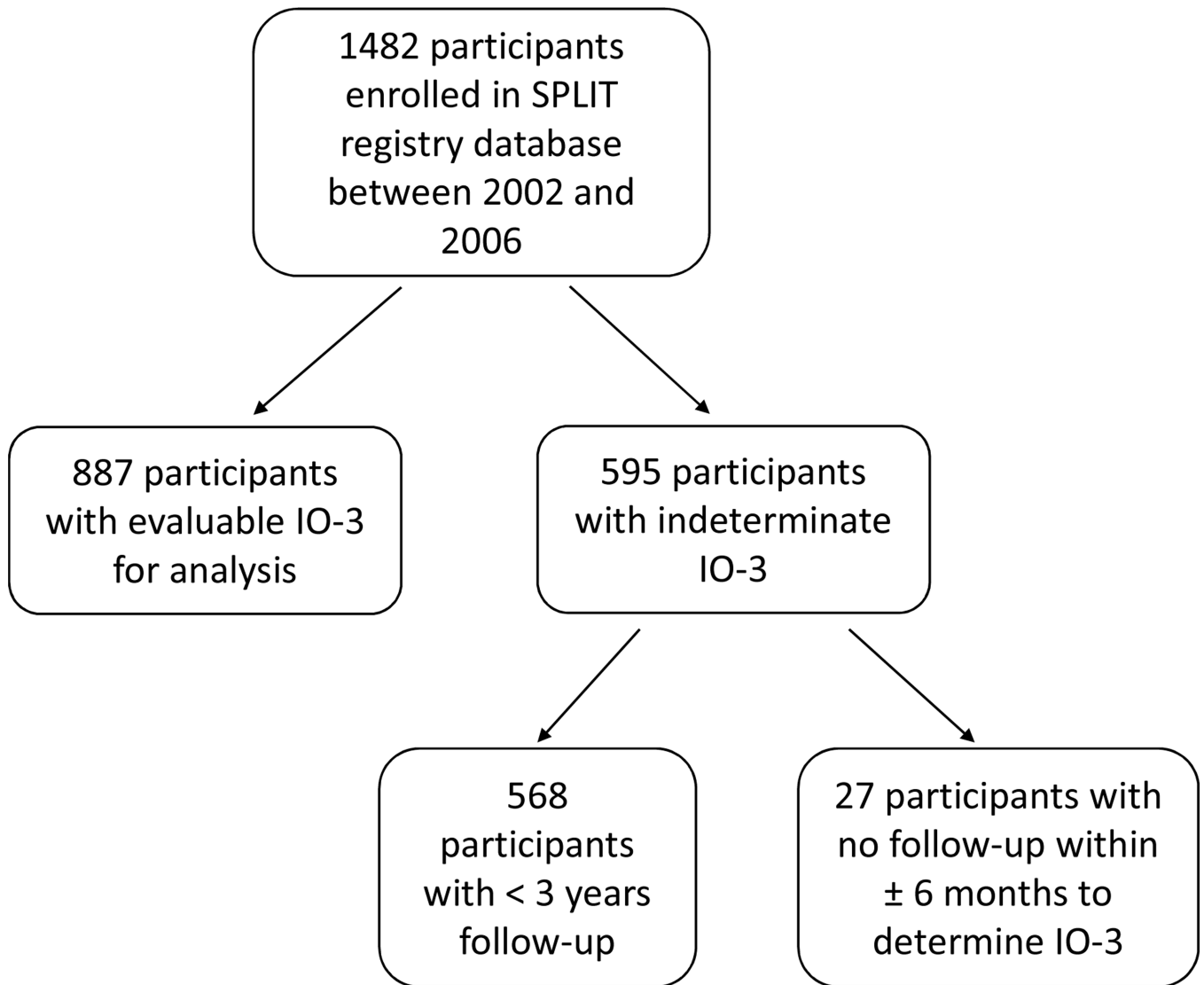
Figure 1.
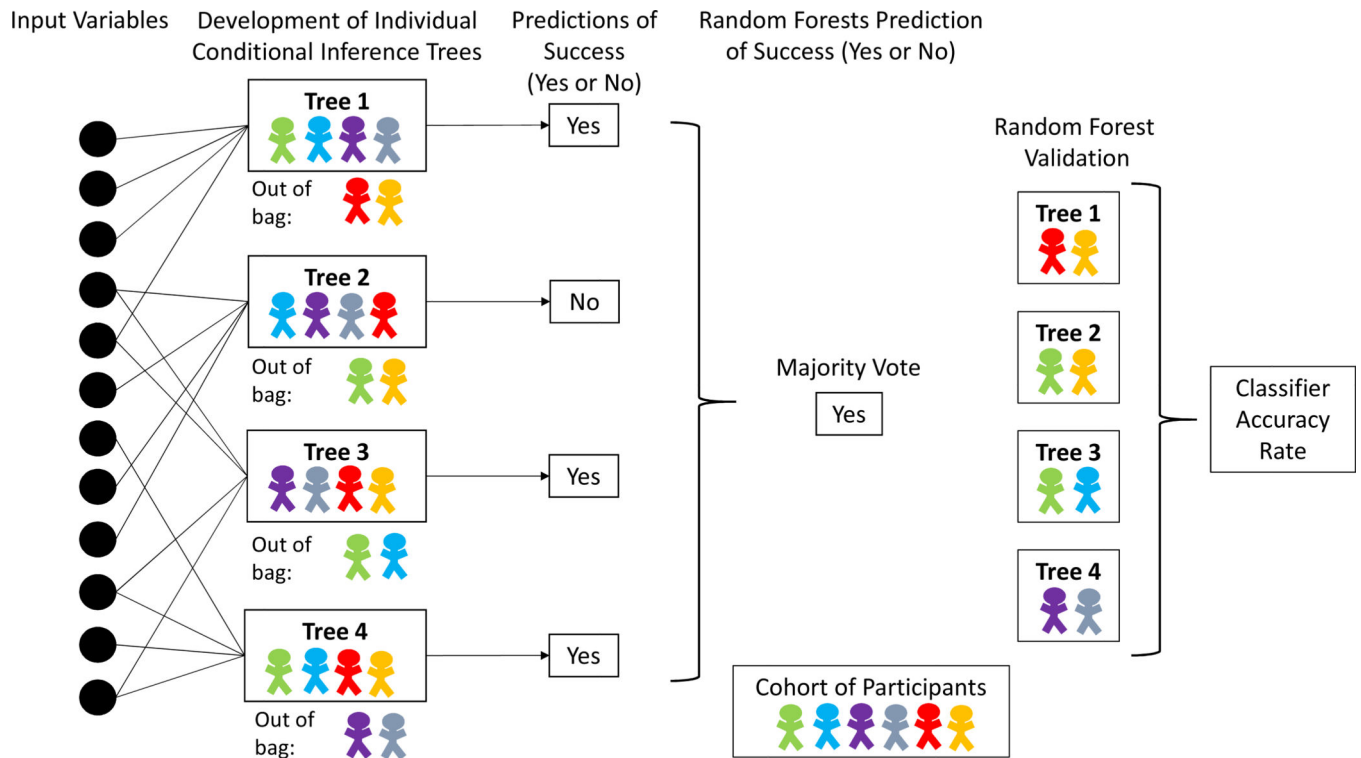Participant inclusion/exclusion diagram

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 incorporates training and testing within individual trees by always holding some participants out (i.e. 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.

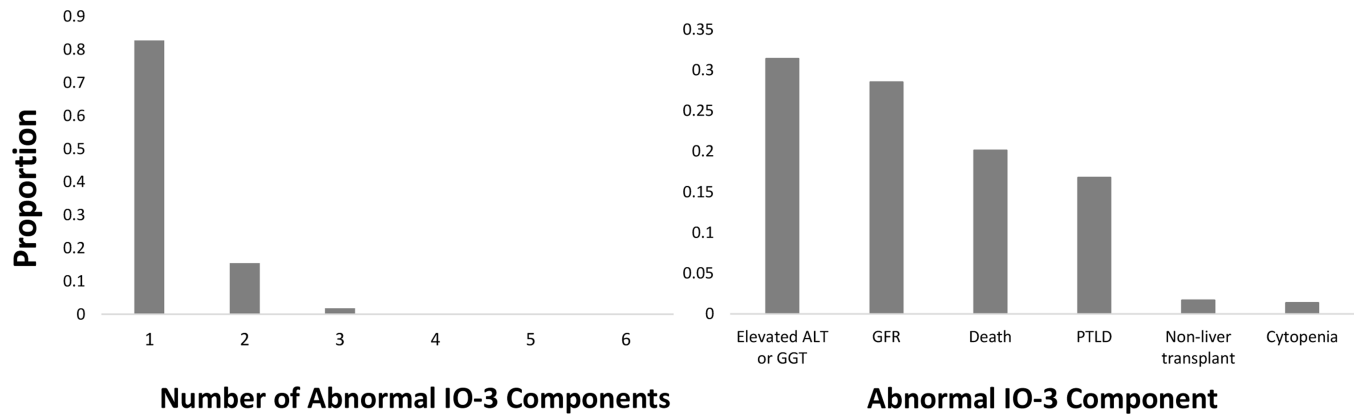**Number of Abnormal IO-3 Components**

Abnormal IO-3 Component

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-glutamyl transferase; PTLD – post-transplant lymphoproliferative disease

Figure 4.

Descriptive heat map of predictor variables available at 1-year and ideal outcome Legend:

Ideal outcome at 3 years: 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.
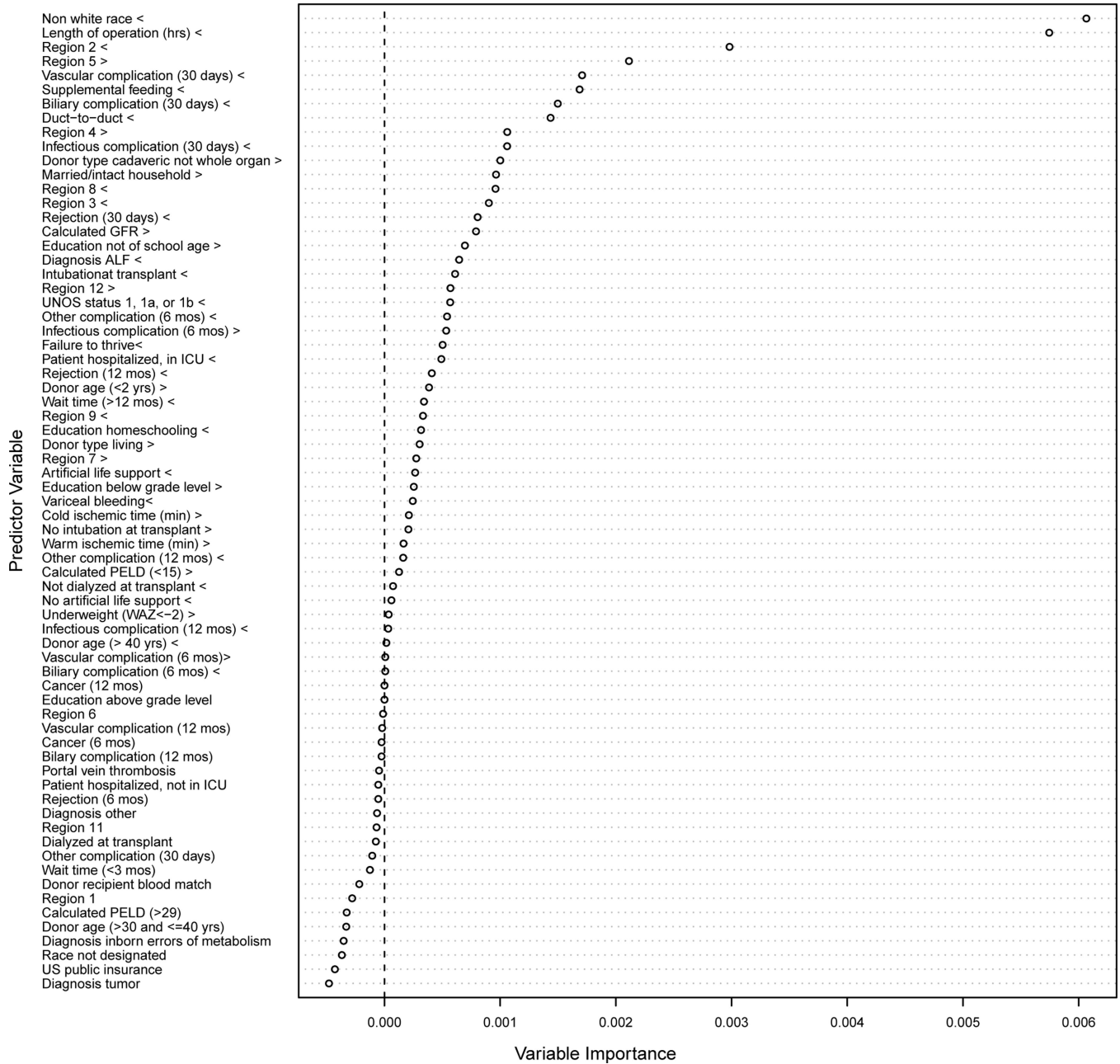Random forests analysis ranking of variable importance Legend: < signifies that the variable predicts no ideal outcome; > signifies that the variable predicts ideal outcome hrs – hours; GFR – glomerular filtration rate; ALF – acute liver failure; UNOS – United Network for Organ Sharing; ICU – intensive care unit; min – minutes; yrs – years; mos – months; PELD – Pediatric end-stage liver disease; US – United States

Author	Study Design	Data acquisition	Data Analysis & Interpretation	Drafting Manuscript	Revising manuscript
SIW			X	X	X
EKH	X	X	X	X	X
MLS			X		X
RA		X	X		X
VLN	X	X	X		X
JCB	X		X		X

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 < 0.01*			*p = 0.76*
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 < 0.01*			*p = 0.23*
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 = 0.32*			*p < 0.05*
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 < 0.001*			*p < 0.01*
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 = 0.36*			*p = 0.67*
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 < 0.001*			*p < 0.001*
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%	

Variable	Included n = 887	Excluded n = 595	p-value	No IO n = 553	IO n = 334	p-value
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 (days)	54 (140.5)	65 (166.0)	p = 0.10	52.0 (132.8)	57.5 (145.8)	p = 0.36
Status 1a/1b						
Yes	29.8%	26.9%	p = 0.28	31.6%	26.6%	p = 0.16
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=0.11	13.6 [14.8]	11.3 [14.0]	p= 0.02
Donor age (years)	11 (18)	13 (22)	p=0.21	12 (19)	10 (18)	p=0.23
Organ Type			p = 0.43			p = 0.92
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 < 0.05			p = 0.22
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%	

SD: standard deviation; IQR: interquartile range; IO: ideal outcome; ALF: acute liver failure; dx: diagnosis; PELD: pediatric end-stage liver disease

Missing data were not included in statistical comparisons across groups