Advancing Pediatric Heart Transplantation

Developing Predictive Models for Donor Selection

ABSTRACT

The ORCHID dataset contains granular information grouped into referral, event, and death data tables, including demographics, cause of death, lab values (chemistry, hematology, ABGs), serologic and culture data, fluid and hemodynamic measurements, and time-stamped process milestones such as family approach, authorization, and procurement. Preprocessing will include evaluation of missing data and one-hot encoding of categorical variables. Exploratory data analysis will explore distributions, correlations, and potential anomalies using statistical methods such as Chi-square testing and correlation matrices.

Modeling will employ supervised learning techniques, with algorithms such as XGBoost, support vector machines, random forests, and logistic regression. Evaluation of the machine learning modules will be include metrics such as: accuracy, F1 score, precision, recall, and ROC/AUC. Cross-validation (k=5) and calibration plots will assess robustness and generalizability.

The project is structured over five weeks: initial data filtering and preprocessing (Week 1–2), model development and validation (Weeks 3–4), and final analysis and write-up (Week 5). Ultimately, this work aims to generate actionable, data-driven insights to support clinical decision-making in donor selection and improve pediatric transplant outcomes.

1. Introduction / Background

Pediatric heart transplantation is a life-saving intervention for children with end-stage heart failure, yet the process of donor selection remains largely subjective and lacks standardized, data-driven tools. Given the scarcity of pediatric donor hearts and the critical nature of timely decision-making, improving donor selection is essential for optimizing transplant outcomes and reducing organ discard rates. While clinical judgment plays an important role, the increasing availability of large, structured donor datasets presents an opportunity to enhance selection protocols through machine learning. This project leverages the ORCHID dataset, a comprehensive, multi-center registry of over 133,000 deceased donor referrals and detailed clinical variables, to develop predictive classification models aimed at improving donor selection for pediatric heart transplantation.

2. Related Work

Pediatric heart transplantation remains a critical treatment for children with end-stage heart failure, but the donor selection process is often limited by subjective judgment and lacks standardized, data-driven methodologies. Clinical decision-making frequently hinges on factors such as donor age, ischemia time, infection status, and size matching, yet these variables are weighed inconsistently across transplant centers.[1-3] Moreover, ethical considerations, such as the role of donor obesity, intellectual disability, and evolving debates around xenotransplantation, complicate allocation frameworks.[4-7] As a result, potentially suitable organs may be discarded while other marginal grafts are accepted under pressure, highlighting the need for improved risk stratification and evidence-based donor evaluation.

Efforts to refine donor selection have increasingly turned to outcome predictors such as pre-transplant mechanical support, allosensitization, and diagnosis-specific risk, which are known to impact survival and recovery.[8-10] Registry data also reveal substantial disparities in donor utilization across regions and age groups, with global trends showing rising donor age and variable waitlist mortality.[1] These issues underscore the necessity of tools that can integrate diverse clinical variables to support donor acceptance decisions. Recent clinical innovations, including xenotransplantation trials and algorithmic approaches to high-risk donor assessment, point toward a future in which machine learning may offer scalable and objective solutions.[9] In this context, our project applies predictive modeling techniques to the ORCHID dataset to enhance precision in pediatric donor selection, with the goal of improving both utilization and post-transplant outcomes.

3. Proposed Work (Updated)

3.1 Pre-Processing

3.12 Missing data

Columns that had more than 50% missing data were removed. The data was split into training and test sets, and the training dataset was imputed using KNN (K=5). Only the training set was imputed to prevent data leakage.

3.13 Categorical Variables

Categorical variables were converted to 0 and 1 using one-hot encoding.

3.14 Time Variables

Time variables were parsed accordingly: approached to transplant, approached to authorization, etc

3.2 Exploratory Data Analysis

The dataset was split into transplanted vs non-transplanted. Counts and percentages were obtained for the categorial variables. SD, IQR were obtained for the continuous variables.

Across the heart-donor referrals in this cohort, several binary referral-and-testing flags differ markedly between recipients who ultimately underwent transplantation (group 1) versus those who did not (group 0). Notably, brain death status shows a dramatic shift: almost all non-transplanted donors were not brain-dead (99.4 % vs. 0.6 %), whereas over half of transplanted donors were (51.6 % vs. 48.4 %), $p < 1 \times 10^{-107}$. Referral type also matters: tissue and eye referrals were both more common among transplanted donors (tissue 34.9 % vs. 22.6 %, $p = 3 \times 10^{-4}$; eye 35.4 % vs. 22.7 %, $p = 8 \times 10^{-5}$), suggesting that centers with multi-tissue interests may more aggressively pursue hearts. Several microbiology flags (e.g. positive Gram-stains of the right and left bronchi, sputum) were also significantly more frequent in transplanted cases (for example, R Bronch Gm St: 73 % vs. 50 %, $p = 6 \times 10^{-5}$), which may reflect more intensive peri-donation evaluation rather than a barrier to procurement.

Age of the donor differed only slightly (mean 7.7 years non-transplanted vs. 8.7 years transplanted, $p \approx 0.005$), while Height and Weight were modestly higher in transplanted donors (Height 48.4 in vs. 50.95 in, $p \approx 0.03$; Weight 34.8 kg vs. 41.5 kg, $p \approx 1 \times 10^{-4}$). Timing intervals from brain death or asystole to referral, approach, authorization, and procurement showed no consistent directional differences (all $p > 0.05$), indicating that logistical delays were similar regardless of ultimate transplant status.

Overall, the strongest discriminators of transplantation were clinical donor characteristics (especially brain-death status) and the engagement of multiple referral streams, rather than differences in donor size or process timing.

3.3 Modeling

Will split the data into a standard train/test 80/20 split. Will start with a standard logistic regression upon which all the other models will be compared. Will experiment with XGboost, support vector machines, Random Forest. Will tune the machine learning models as able such as class weights, SMOTE, sensitivities, number of trees, depth of trees, learning rates, kernal changes, gamma changes, and regularization. May consider RandomizedSearchCV for GridSearchCV to help search for optima parameters.

4. Evaluation

Will evaluate the models using confusion matrix, ROC and PR plots, accuracy, F1 score, precision, recall, calibration, and learning curves.

Will evaluate generalization via k-fold method (k=5).

Will also perform a pre and post power analysis to ensure the study was adequately powered.

5. Discussion

The proposed work should be completed within 5 weeks. May add a week during the machine learning modeling phase in case changes need to be made (longest duration 6 weeks). Some potential changes could include the size of the database (could be computationally expensive). May have to decrease the number of models if the dataset is too large.

6. Conclusion

In summary, this project aims to identify the specific donor characteristics that predict suitability for transplantation. Multiple machine learning models will be developed and optimized to classify transplant versus non-transplant outcomes. Future directions may include incorporating additional metrics to refine donor assessment, expanding criteria for extended donor status, leveraging real-time sensor data to improve donor evaluation, and integrating recipient data for a more comprehensive predictive framework.

ACKNOWLEDGMENTS

None

REFERENCES

1. Elizer S, Mantell BS. Risk factors affecting mortality in pediatric heart transplantation: A comprehensive review of pre- and post-transplant contributors. JHLT Open. 2025;9:100309.
2. Conway J, Ballweg JA, Fenton M, Kindel S, Chrisant M, Weintraub RG, et al. Review of the impact of donor characteristics on pediatric heart transplant outcomes. Pediatr Transplant. 2020;24(3):e13680.
3. Donne M, De Pauw M, Vandekerckhove K, Bove T, Panzer J. Ethical and practical dilemmas in cardiac transplantation in infants: a literature review. Eur J Pediatr. 2021;180(8):2359-65.
4. Wilkens SJ, Gossett JG, Patel A. Ethical Considerations in Pediatric Heart Transplantation. In: Mavroudis C, Cook JT, Mavroudis CD, editors. Bioethical Controversies in Pediatric Cardiology and Cardiac Surgery. Cham: Springer International Publishing; 2020. p. 217-35.
5. Berkman E, Wightman A, Friedland-Little JM, Albers EL, Diekema D, Lewis-Newby M. An ethical analysis of obesity as a determinant of pediatric heart transplant candidacy. Pediatr Transplant. 2021;25(3):e13913.

6. Hurst DJ, Padilla L, Merlocco A, Rodger D, Bobier C, Gray WH, et al. Pediatric Cardiac Xenotransplantation: Recommendations for the Ethical Design of Clinical Trials. Transplantation. 2024;108(10):e292-e300.
7. Singh TP, Hsich E, Cherikh WS, Chambers DC, Harhay MO, Hayes D, Jr., et al. The International Thoracic Organ Transplant Registry of the International Society for Heart and Lung Transplantation: 23rd pediatric heart transplantation report-2020; focus on deceased donor characteristics. J Heart Lung Transplant. 2020;39(10):1028-37.
8. Singh N, Raees MA, Zafar F. Donor considerations in pediatric heart transplantation. Transl Pediatr. 2019;8(4):284-9.
9. Bearl DW. Ethical issues in access, listing and regulation of pediatric heart transplantation. Transl Pediatr. 2019;8(4):278-83.
10. Canter CE, Shaddy RE, Bernstein D, Hsu DT, Chrisant MR, Kirklin JK, et al. Indications for heart transplantation in pediatric heart disease: a scientific statement from the American Heart Association Council on Cardiovascular Disease in the Young; the Councils on Clinical Cardiology, Cardiovascular Nursing, and Cardiovascular Surgery and Anesthesia; and the Quality of Care and Outcomes Research Interdisciplinary Working Group. Circulation. 2007;115(5):658-76.