Analyzing Transplant Success through Statistical and Data Science Modeling: Uncovering Determinants of Patient Survival

Charlotte Bony, Ashley Colvin, Emma Choo, and Krisha Mansukhani

Introduction

Organ transplantation is a critical and complex medical intervention that saves thousands of lives annually, yet significant disparities in outcomes remain prevalent. Variables such as donor type, organ type, recipient demographics, and socioeconomic status are known to influence success rates, but the interplay between these factors is not fully understood. For example, recipients from ethnic minorities and those relying on Medicaid exhibit consistently lower survival rates than privately insured counterparts, even when clinical factors are adjusted (McElroy et al., 2022). Furthermore, living donor transplants have shown remarkably higher success rates compared to deceased donor transplants, particularly for kidneys and livers, with survival rates increasing by up to 35% as reported in the OPTN (Organ Procurement and Transplantation Network) and SRTR (Scientific Registry of Transplant Recipients) 2020 Annual Data Report.

Despite the growing awareness of these obstacles, there remains a need for a more comprehensive way to assess outcomes across diverse populations. While other studies may have focused on other limited variables or single-center datasets which may not reflect the nationwide population, our project aims to bridge the knowledge gap by analyzing a national dataset that has extensive information on organ donors and recipients from transplant centers across the country, including demographic details, donor types, insurance categories, waitlist information and post transplant survival metrics. With the need for organ transplantation increasing annually, the need to better understand key predictors becomes ever more crucial (Merola et al., 2016). Through the use of statistical and machine learning models, our goal is to identify actionable predictors of one year post transplant survival and in doing so, offer insights to optimize patient care and address disparities with improved policy or clinical decision making.

Methods

General Data Collection and Overview

Data were accessed and downloaded through the United States Health Resources & Services Administration (HRSA). These data are records from Organ Transplant Centers (OTC) and hospitals across the USA that include a national and state overview of the number of donated organs, the number of transplanted organs, donor and recipient demographics, OTC state/county service areas, insurance types, waitlist demographics, waitlist removals, and transplant outcomes. The snapshot of records we obtained are cumulative from January 1st, 2019 to December 16, 2024. A majority of the data types are numeric count data since the records per transplanted organ or OTC are aggregated as the number of patients meeting a variable's criteria. The remainder of the data types are percentages representing survival outcomes or categorical data representing the organ donated, OTC name, state name, etc.

Data Cleaning, Outlier Removal, and Preprocessing

The HRSA data were provided in a multi-sheet excel spreadsheet format, so sheet separation was necessary. Each sheet was extracted individually from the master spreadsheet document using Pandas and a copy was created of each. Each sheet copy was converted into a Pandas DataFrame and any records marked as "Suppressed" were converted to NAs and subsequently dropped. Leading whitespaces and periods, which were frequent throughout every dataframe, were also cleaned. Any spreadsheet containing percentages were converted into floats rather than rounded integers to retain the fractional decimal data that may have been important in model creation. Lastly, before outlier removal, variable names were truncated and/or renamed for ease of access.

Variables of interest were identified and subsetted from their parent dataframes to make outlier removal easier. Outliers on the targeted response variables were then identified using the Interquartile Range (IQR) method before being dropped. Outliers were only processed on response variables because retaining the spread and variance of input/dependent count data may be crucial in model creation. The specific response variables, 'Percent of Patients Alive at 30 Days' and 'Percent of Patients Alive at 1 Year' were not only chosen as responses because of the nature of transplantation procedures, but also because they were the only data expressed as percentages rather than counts. After the outliers were removed, multiple dataframes were inner-joined together on common variables of OTC-state, organ type transplanted, and donor type alive/deceased. Two main frames resulted from this preprocessing; the first being 'gender_outcomes' that includes OTC-state, organ type, donor type, the count of patients in every age group over the age of 18, and recipient sex with response variable 'Percent of Patients Alive at 1 Year'. The second dataframe, 'insurance_subset' included OTC-state, organ type, donor type, counts of recipients in each of the four most populated ethnicities, recipient insurance type, and the response variable 'Percent of Patients Alive at 1 Year'.

To prepare the categorical variables for modeling, we opted to use encoding to create dummy variables using Pandas where a variable's categories were represented as integers. Most modeling packages in python require this step because, unlike R, a "factor" variable type that retains the name of a category as a string is not supported.

Exploratory Data Analysis

For the exploratory data analysis (EDA), we investigated patterns and relationships within transplant outcome data across demographic and clinical variables. The analysis was conducted in Python with libraries including pandas, seaborn, matplotlib, scipy, and statsmodels. For statistical assessment, we used the Shapiro-Wilk test and Q-Q plots to evaluate normality. Since most transplant-related variables were not normally distributed, we opted for non-parametric tests, specifically the Kruskal-Wallis H test for comparisons across more than two groups (e.g., organ type, age, insurance, and ethnicity) and the Mann-Whitney U test for pairwise group comparisons (e.g., male vs. female, living vs. deceased donors).

Visualizations were key in identifying trends. We used seaborn and matplotlib to generate boxplots for outcome distribution comparisons, pie charts to summarize categorical proportions (e.g., gender distribution), and line charts for analyzing transplant distribution by age group. To explore the strength of relationships between variables, we constructed scatter plots with regression lines and reported R^2 values to quantify model fit. These were helpful in evaluating the predictive relationship between graft function and survival outcomes at both 30 days and 1 year. To ensure comparability across transplant centers of varying sizes, demographic counts (e.g., race, insurance) were normalized by the total number of patients per center, allowing for proportional comparisons. Additionally, we computed Spearman correlation matrices between normalized demographic proportions (e.g., ethnicity or insurance) and transplant outcome metrics, visualized using heatmaps.

Model Selection

The most difficult part of the process was finding modeling types that were suitable for the types of data available for the analyses. It was imperative that the response variables were properly represented in the model as the original percentages they were recorded as, so research was performed on the types of models that can accurately represent percentage or proportional data. There exists research on a model that accurately represents percentages expressed as proportions out of 1.00 called a beta regression model as presented by Ferrari, S., & Cribari-Neto (2004). The model described in their 2004 paper is a regression model where the response variables are bound between [0,1) and the data may or may not be normally distributed (Ferrari, S., & Cribari-Neto 2004). Beta regression would therefore be the first model

type selected. We opted to use one beta-distributed variable, ‘Percent of Patients Alive After 30 Days’ (post transplant), so a transformation was necessary (see the Model Creation subsection for more details).

To be able to cross examine multiple models to determine which most accurately represent the trends in the HRSA data, we opted to use a more high-level approach with our second model. Rather than comparing a common linear regression model with the beta regression model, a Keras model using the Tensorflow/Keras packages was chosen instead. We chose to use a Keras model to accommodate potential non-linear relationships in the data and also gain experience in a widely used framework in modern machine learning workflows.

Model Creation

The Beta Regression Model

To finalize the preparations of the data for the beta regression model, we needed to first convert the percentages in the response variables to proportions out of 1.00 and ensure that the counts of patients per variable type are also converted to proportions out of the reported patient totals. Even though input variables are not required to be expressed as proportions like the response variables are, we did this to ensure that we captured the scale of the variable in relation to the totals. The input variables were then scaled using StandardScaler from sklearn to transform the data to a mean of 0 and standard deviation of 1.

To ensure that the variables selected for the model were meaningful and to reduce the dimensionality of the data, we calculated the Variance Inflation Factor using statsmodels for each variable. The most important benefit of performing a VIF analysis was to reduce multicollinearity between variables in the dataset. As an example, there are only two options in our dataset for organ donor type: alive and deceased donor. Since there are only two, logic dictates that if a donor was not alive, they must be deceased. These variables are, by design, 100% collinear and one must be eliminated to measure if other variables are collinear with it instead of just itself. We decided to eliminate alive donors and male transplants because these were binary and having both provided redundant information - a non-alive donor implies a deceased donor and a non-male recipient implies a female recipient. By default, it is recommended to drop VIF values that are above 5.00 and necessary to drop values above 10.00. Unfortunately, the variables all had VIF values above 5 and many above 10, so we elected to keep all values and try another test. We then piped the variables into a lasso regression model to further improve the variable selection. The variables that had the highest parameter importance were selected and kept from the lasso regression model. After the lasso regression, one last pass through a VIF analysis was performed and all VIF values were found to be less than 5, indicating a successful collinearity reduction and variable selection process. A constant was also added as it was necessary from the statsmodels and sklearn documentation.

Therefore, the variables selected to be in the final dataframe for the beta regression model were as follows:

Input Variable Name	**Description**
‘const’	A constant as required by regression models coded using the default method from the documentation
'Alive30Days'	Percentage of recipients out of the total that were reported to be alive 30 days post-transplant expressed as a proportion out of 1.00
'Donor_Deceased Donor'	The recipients received an organ from a deceased donor
'Organ_Kidney'	The recipients had a kidney transplanted
'Organ_Liver'	The recipients had a liver transplanted
'Organ_Lung'	The recipients had a lung transplanted
Response Variable Name	**Description**
'Alive1Year'	Percentage of patients that were reported to be alive after 1 year post-transplant (converted to a proportion out of 1.00)

The last step in final pre-model preparation was to apply a transformation to the data that would allow the model to distinguish between zeroes that were meant as categorical dummy variables and zeroes that meant a count that was numerically zero. The transformation was also helpful in fitting the data to work better with beta regression models since not every beta distribution is normal, and because the input variable ‘Percent of Patients Alive After 30 Days’ is a beta-distributed proportion. After extensive research, we decided to use a transformation method from Smithson, M., & Verkuilen, J. (2006):

[1]
$$y_i = \frac{x_i(n-1)+0.5}{n}$$
(Smithson, M., & Verkuilen, J., 2006)

where y is the transformation output for each observation/record, x is the input value, n is the total patient count, and the added 0.5 acts as a ε value to set values strictly between 0 and 1, a

necessary condition for the beta regression model that requires values to be within said specific interval:

[2]
$$y_i = x_i\beta + \epsilon_i$$

After the preparation, the data were split into training and test sets using the train_test_split function with an added reproducible seed parameter from sklearn. The defaults for this function were used, so the training set was set to 25% of the data and the testing set was set to 75% of the data. The data, now properly prepared, were passed through a beta regression model from the statsmodels library (which, ironically, was in its beta phase of production). The results are discussed in the following major section.

The Keras Neural Network Model

The Keras Neural Network model was developed using the same processed testing and training set as the beta regression model. Since Keras is a high-level Tensorflow package, the model creation was much more brief.

For this model, we opted to use three hidden layers on a Sequential model framework from the Keras package. The model included a dense input layer with a shape of 14 input columns and 64 neurons and a Rectified Linear Unit (ReLu) activation function. This was followed by two more duplicate dense layers of 64 neurons each and ReLu activation functions. The output dense layer was one neuron representing the single number output we want as a result and it used a sigmoid activation function. The sigmoid activation was necessary because our response is bounded between 1 and 0.

The model was then compiled with a mean squared error (MSE) loss function, an adam optimizer, and set to display MSE as our response metric to determine model performance. An early stop and model save function was included so that, in the many epochs of training, the model will stop when it sees the best fit and save the result at that point for further analysis.

The completed model framework was fitted using the training sets for inputs and outputs, set to train on 100 epochs, and set to carve out a small bit of validation data from the testing input and output sets so that the entire workflow of the model's processing can be completed in one step. Predictions were then made from the fitted and trained model and the model MSE was measured.

In evaluating both models' performance, the MSE was prioritized over R^2. While MSE provides a direct measure of prediction accuracy by quantifying the average difference between predicted and actual values, R^2 describes the strength of a linear relationship between predictions and outcomes. In this context, R^2 value can be misleading because of the nature of this complex and

noisy clinical dataset where relationships may not be strictly linear. Given the center-aggregated nature of the transplant data and the expected heterogeneity, the low MSE values obtained indicate strong predictive performance, even if the R^2 values are not close to 1. Therefore, the MSE was treated as the primary indicator of model success.

Results

Exploratory Data Analysis (EDA)

To evaluate factors influencing transplant success, we conducted a comprehensive analysis of post-transplant graft function and survival outcomes, stratified by organ type, demographic variables, donor type, insurance status, and ethnicity. In all exploratory analyses, survival outcome (Percent of Patients Alive at 1 Year) was treated as the primary dependent variable. Both visualizations and statistical tests were used to assess relationships and group differences, using non-parametric methods due to non-normality of the data (all Shapiro-Wilk p-values < 0.05).

Age Group Distribution

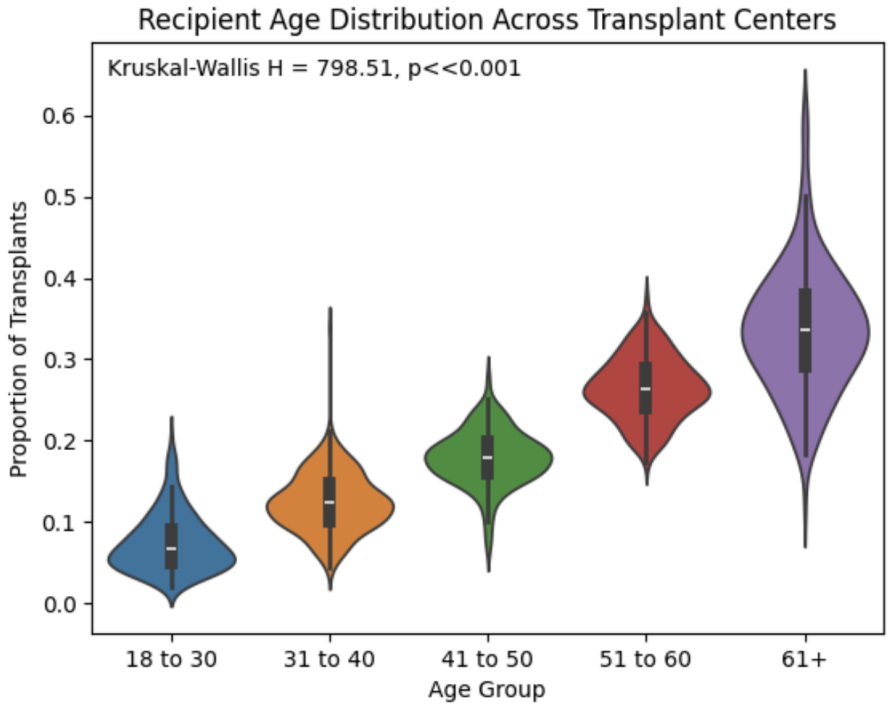

Figure 1. Age played a significant role in transplant patterns. We assessed the distribution of transplant recipients across different age groups (18to30, 31to40, 41to50, 51to60, 61plus) at transplant centers. Violin plots, when scaled by the number of transplant centers, showed a clear upward trend in transplant frequency with increasing age, with the 61+ group receiving the

largest proportion of transplants. This pattern reflects clinical reality, as older individuals are more likely to develop late stage organ disease requiring transplantation. The Kruskal-Wallis test produced a highly significant H value of 798.5076 ($p < 0.0001$), confirming substantial differences across age groups. This plot reveals that older age groups (61+) consistently make up a larger share of transplants across centers. Most centers had a high proportion of transplants in the 61+ age category, while younger recipients (18-30) were far less common. These findings suggest that patient demographics can shape transplant practices and resource allocation.

Figure 2. Regression plots show a positive association between the proportion of younger transplant recipients and 1-year survival. This relationship was most notable in the 18–30 and 31–40 age groups, where higher transplant proportions were modestly correlated with better survival outcomes ($R^2 = 0.19$, slope = 39.85; $R^2 = 0.13$, slope = 28.09). The 41–50 group showed a weaker relationship ($R^2 = 0.05$, slope = 15.26). In contrast, the 51–60 and 61+ age groups demonstrated negative associations with 1-year survival ($R^2 = 0.10$, slope = –16.80; $R^2 = 0.14$, slope = –12.55), indicating that centers performing a greater share of transplants in older patients tended to have lower survival rates. These results suggest that age composition may influence transplant outcomes, with younger recipient populations associated with more favorable post-transplant survival.

Gender-Based Differences

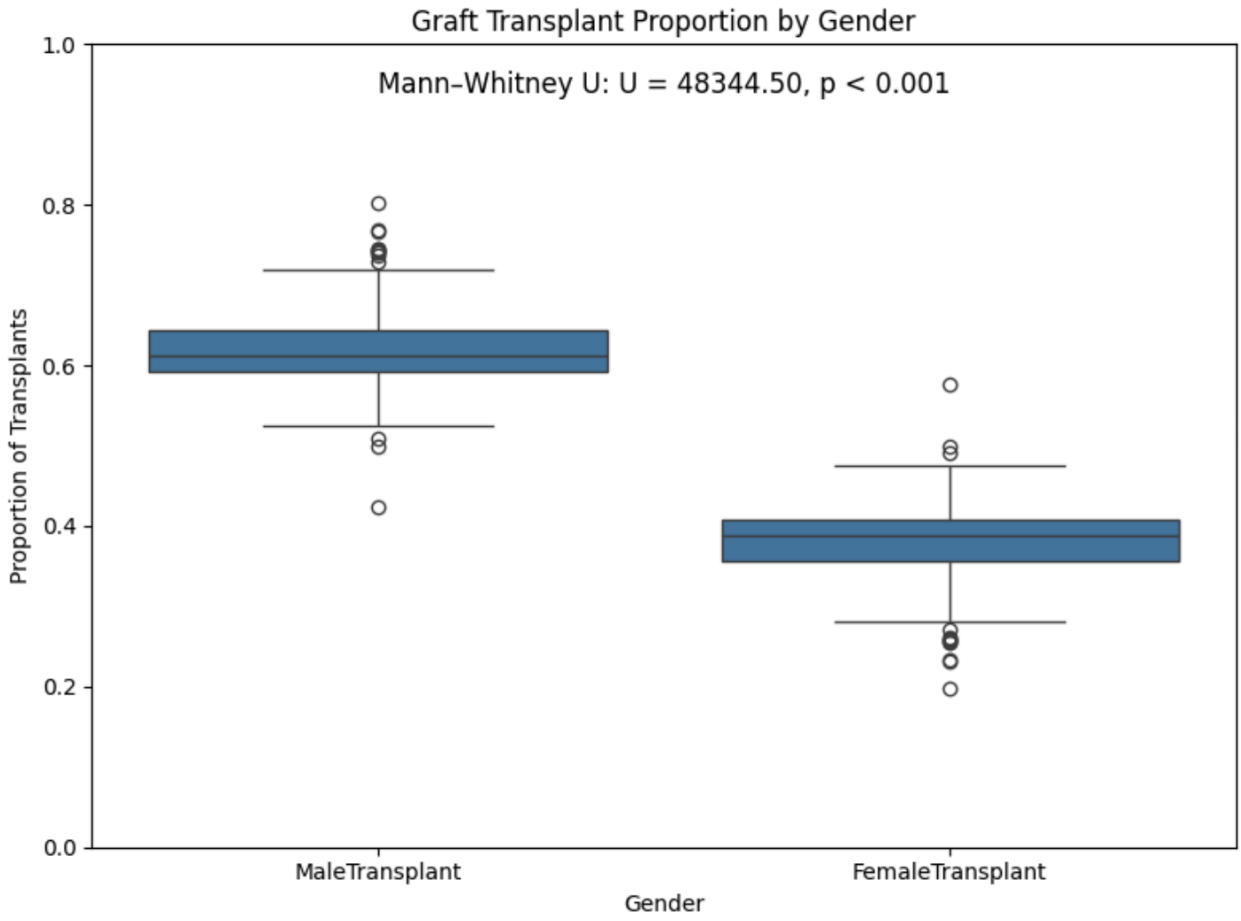

Figure 3. We analyzed the distribution of male and female transplant recipients across centers using the Female/Transplant and Male/Transplant proportions. Analysis of gender differences in transplant outcomes showed that male recipients received a higher proportion of transplants across centers, with a median transplant proportion of 62% compared to 38% for female recipients. A Mann–Whitney U test confirmed that this difference was statistically significant ($p < 0.0001$), indicating that the distributions of transplant proportions differ meaningfully between genders.

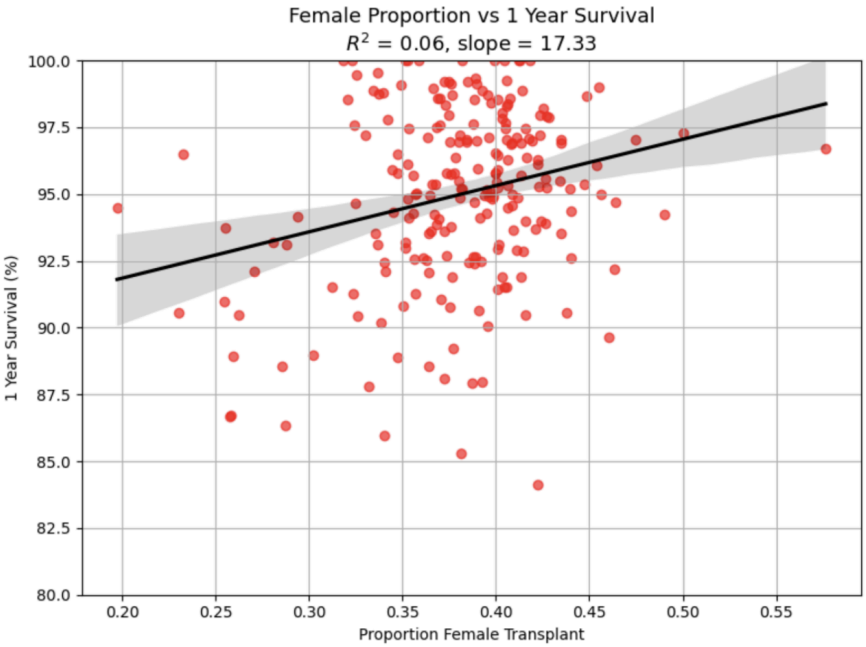

Figure 4. However, regression analysis revealed no strong relationship between gender proportion and 1 year survival for either group (R^2 = 0.06 for both male and female proportions), suggesting that gender distribution alone does not account for variation in outcomes. Male and female transplant proportions are inversely related, so the observed trends reflect the same underlying distribution. While male recipients were more commonly represented, survival outcomes appear comparable between genders at the center level. These findings suggest that although gender disparities exist in transplant distribution, they may not translate into differences in survival at 1 year.

Donor Type and Survival

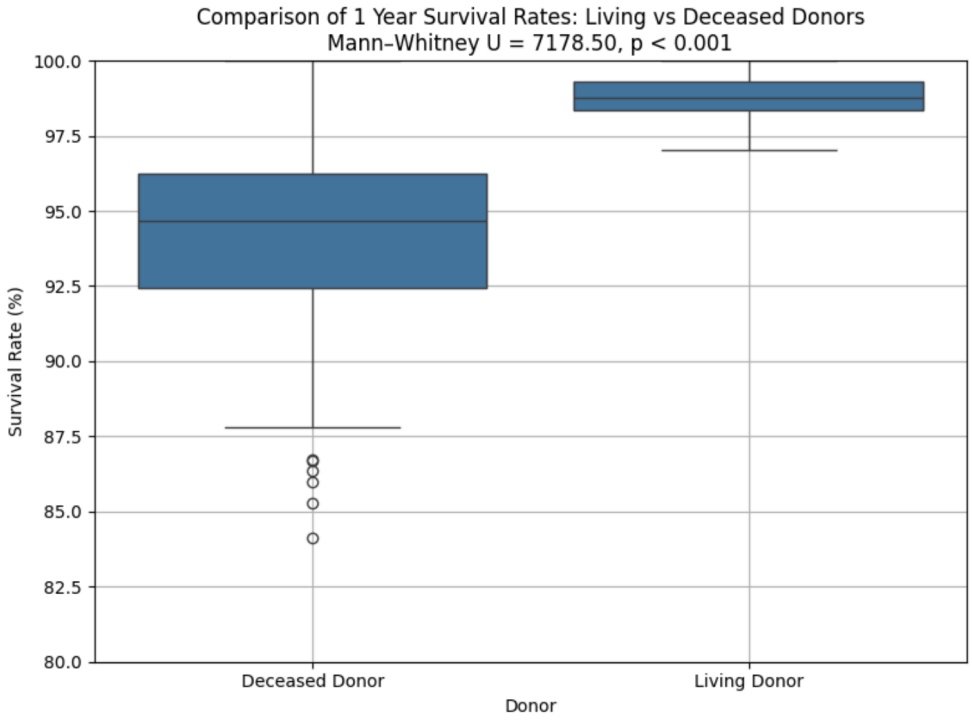

Figure 5. We compared 1 year survival outcomes between living donor and deceased donor transplants using the Donor variable. Mann–Whitney U tests were applied to detect survival differences, and violin plots were used to visualize outcome distributions across donor types. Comparing living versus deceased donors, boxplots revealed that recipients of organs from living donors had significantly higher 1 year survival rates. The Mann-Whitney U test confirmed this finding ($p < 0.0001$). These findings reinforce the positive clinical outcomes associated with living donor transplants, potentially due to improved organ quality.

Insurance Coverage and Outcomes

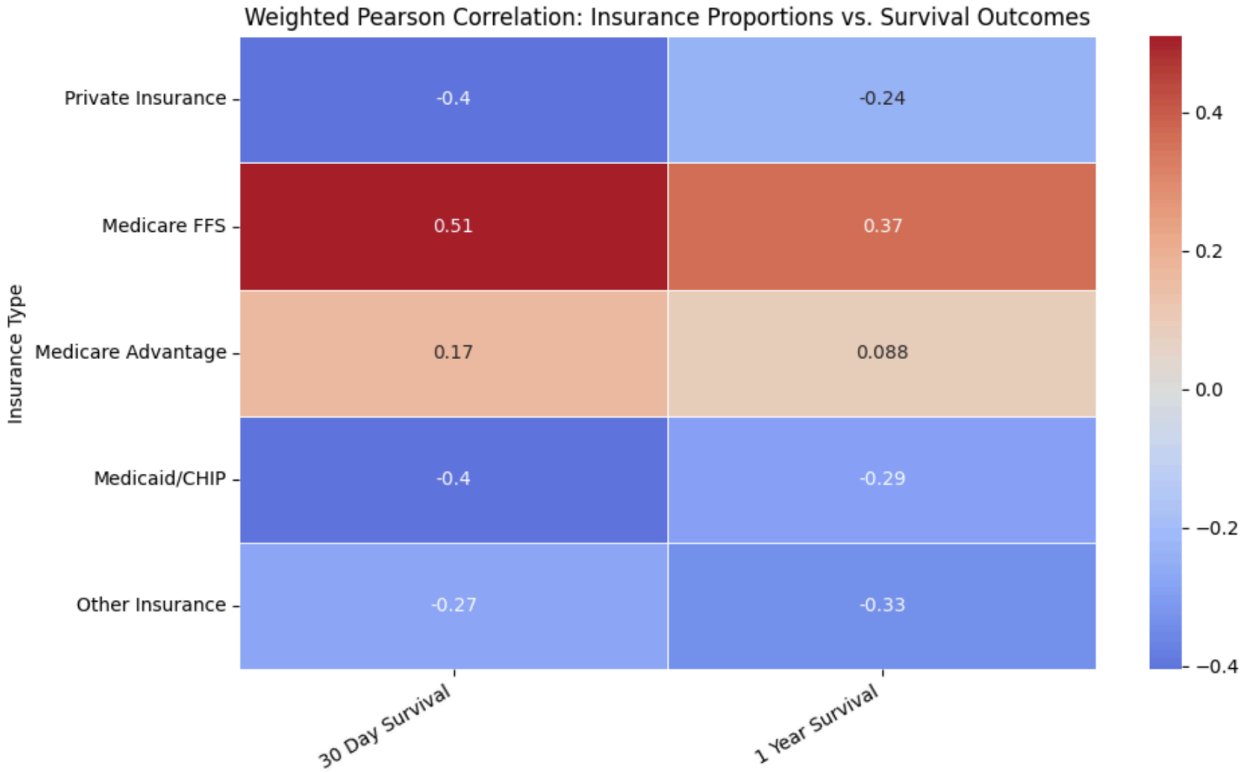

Figure 6. We investigated how insurance type distributions (Medicare/Transplant, Medicaid CHIP/Transplant, Private/Transplant, Other/Transplant) across transplant centers correlated with 1 year survival outcomesThe weighted, center-level correlation heatmap revealed disparities in 1 year . transplant survival outcomes across insurance types. Centers with Medicare FFS recipients showed a positive correlation with 1 year survival (r = 0.37), suggesting improved long-term outcomes. In contrast, Medicaid/CHIP and especially “Other Insurance” proportions were negatively correlated with 1 year survival, with “Other Insurance” reaching r = -0.33. Because correlations were weighted by total transplants per center, these findings reflect consistent patterns across high-volume programs and underscore the potential influence of payer type on transplant success.

Ethnicity and Outcomes

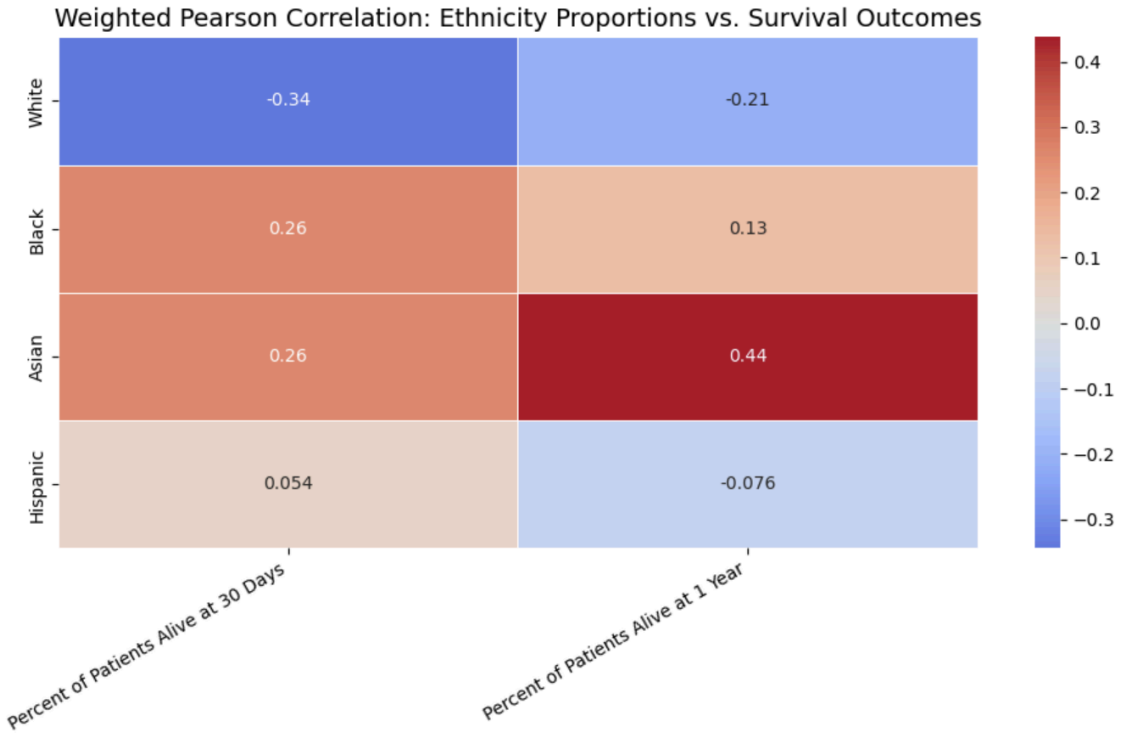

Figure 7. At the center level, weighted Pearson correlation analysis revealed associations between ethnicity composition and transplant survival outcomes. We explored the relationship between ethnic group distributions (White/Transplant, Black/Transplant, Asian/Transplant, Hispanic/Transplant, Other/Transplant) and patient survival outcomes. Centers with a higher proportion of Asian recipients showed the strongest positive correlation with 1 year survival (r = 0.44), suggesting better outcomes in these populations. Black recipient proportions were also positively correlated for 1 year survival (r = 0.13), indicating modestly favorable trends. In contrast, higher proportions of White recipients were associated with lower survival outcomes (r =-0.21 for 1 year survival). Hispanic recipient proportions were only weakly correlated with outcomes, with a slightly negative correlation (r = -0.08) at 1 year. These findings suggest that ethnicity distributions at the center level may be associated with differences in transplant outcomes.

Organ Type and Clinical Outcomes

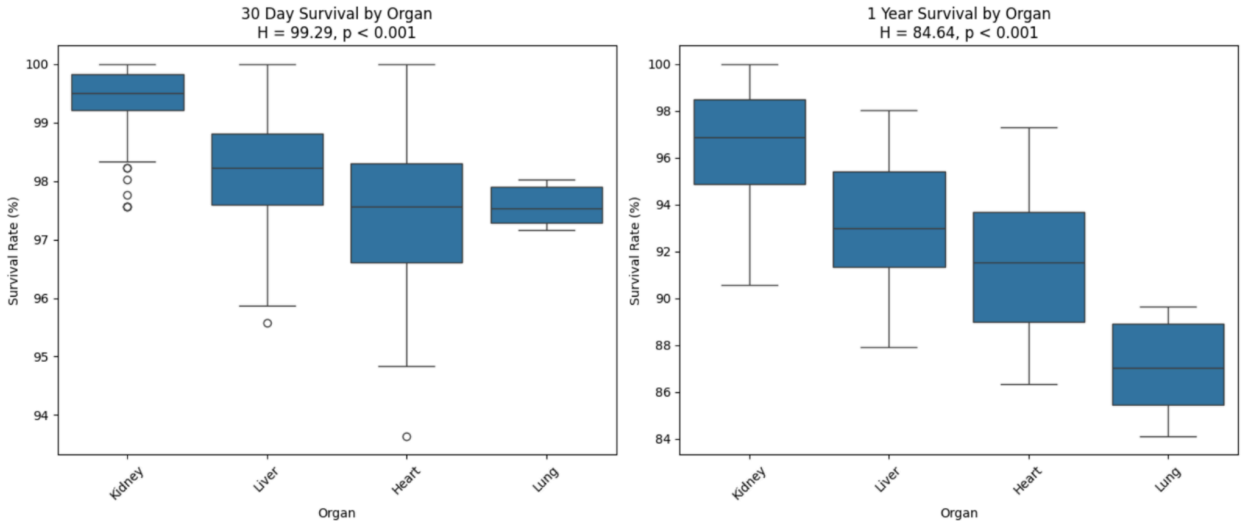

Figure 8. Differences in transplant outcomes varied notably by organ type. We evaluated differences in survival outcomes across organ types (Organ: Kidney, Liver, Lung, Heart. Kidney transplants had the most favorable results, with high survival rates at both 30 days and 1 year and low variability across centers. Liver and heart transplants showed slightly lower survival and broader distributions, reflecting more variable outcomes. Lung transplants had the poorest performance, with 1 year survival medians near 87% and limited spread, indicating consistently lower success rates. These differences were statistically significant (Kruskal-Wallis $H > 84$, $p < 0.0001$), highlighting the strong influence of organ type on post-transplant survival.

Graft Function and Survival Correlations

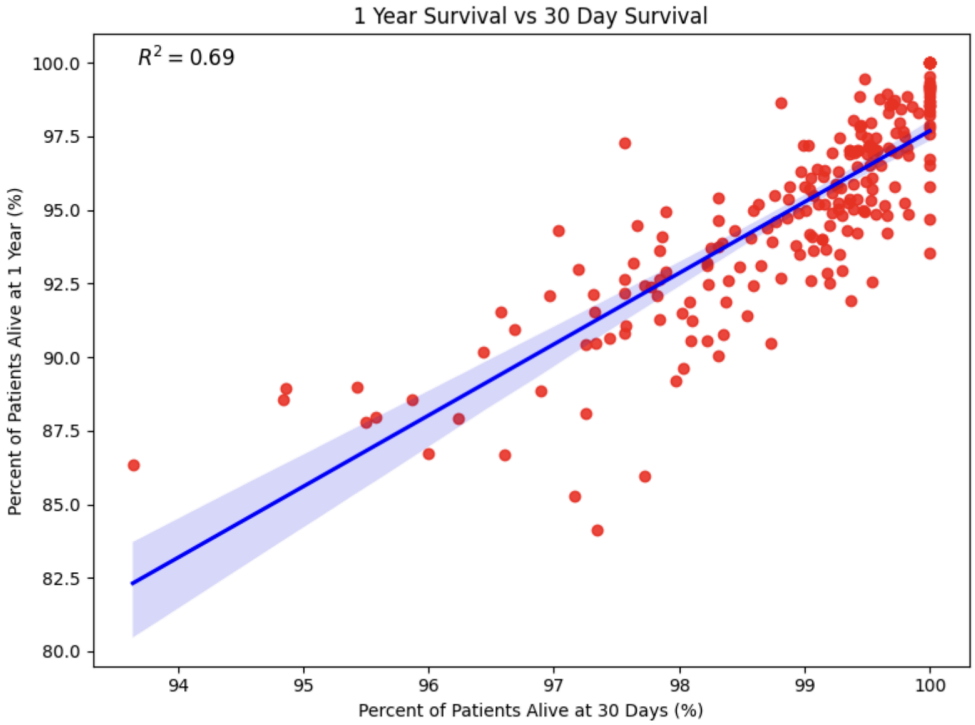

Figure 9. A moderately strong positive relationship was observed between 30 day and 1 year survival (R^2 = 0.69), indicating that early post-transplant recovery is a meaningful predictor of longer-term patient outcomes. This finding suggests that short-term survival metrics may serve as practical early indicators of overall transplant success.

The Beta Regression Model

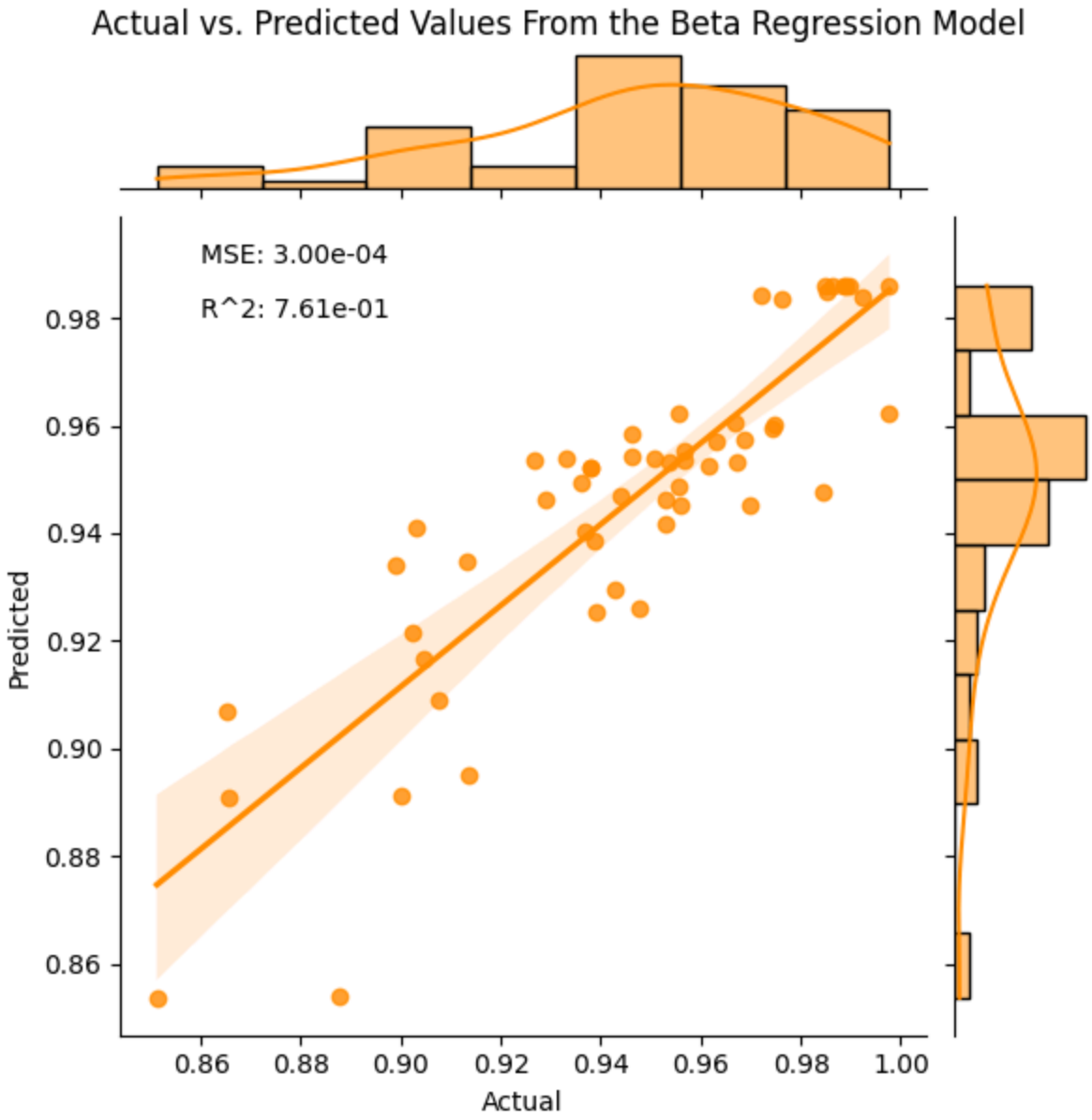

Figure 10. Actual versus predicted response values from the beta regression model. Actual values from the testing set are on the x-axis and the predicted values from that same testing set are on the y-axis. The model performed exceptionally well with an MSE well below 0. The R^2 value, while not indicating a high linear correlation, is presented to show the preference of using MSE during model evaluation in cases of complex input data relationships similar to that found in our dataset; indicating that the variance captured cannot be explained using linear methods.

Beta Regression Model: **Proportion of Patients Alive at 1 Year Post-Transplant Surgery**

N = 165, df residuals = 158, df model = 5, Log-likelihood = 447.98, AIC = -882.0, BIC = -860.2, **MSE = 0.000285**

Variable	coefficient	Standard error	z	P>|z|	[0.025	0.975]
Constant	3.9619	0.154	25.766	p < 0.001	3.661	4.263
Lung Transplanted	-0.6757	0.141	-4.802	**p < 0.001**	-0.951	-0.400
Alive 30 Days Post-Transplant	0.3306	0.034	9.812	**p < 0.001**	0.265	0.397
Deceased Donor	-1.0139	0.122	-8.284	**p < 0.001**	-1.254	-0.774
Kidney Transplanted	-0.0228	0.121	-0.188	0.851	-0.260	0.215
Liver Transplanted	-0.1221	0.103	-1.181	0.238	-0.325	0.081
Precision	4.9061	0.112	43.926	p < 0.001	4.687	5.125

Figure 11. Results of the beta regression model with the MSE of the model listed in the top subtitle. Variables considered to be significant in the model ($p<0.001$) were found to be if the transplanted organ was a lung, if the patient was alive 30 days post-transplant, and if the organ donor was deceased.

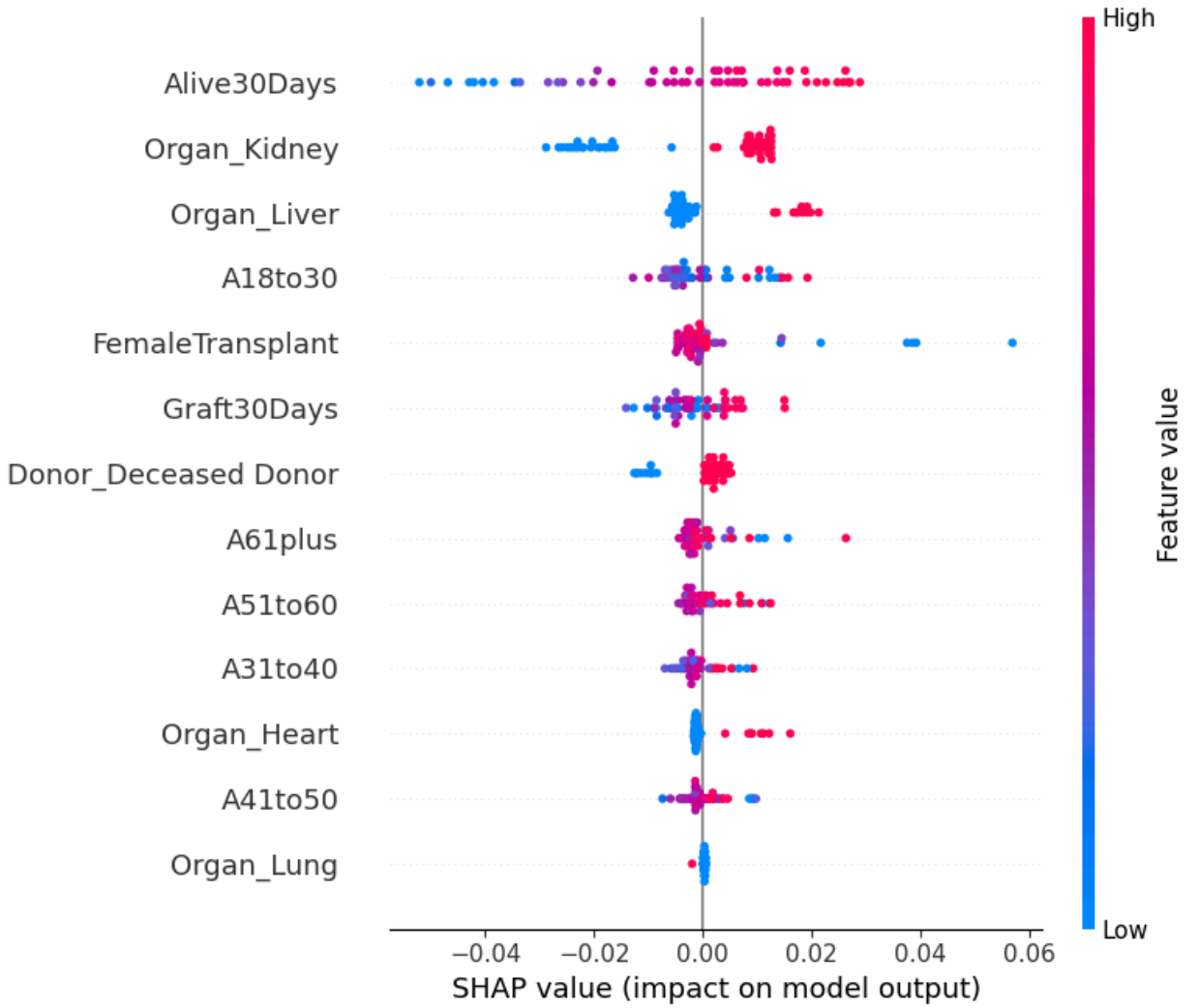

Figure 12. A visualization of the global importance of each parameter in predicting transplant outcomes using SHAP values across a sample of 100 random data points from the Keras Neural Network model. Features displayed in red contribute positively (increase predictions of transplant success), whereas features in blue contribute negatively (decrease predictions of transplant success). The parameters with the greatest global influence include whether patients were alive after 30 days (“Alive30Days”), received a kidney transplant (“Organ_Kidney”), and donor type (“Donor_Living Donor”). Each point represents an individual observation, highlighting the overall impact of each parameter on the model predictions across the entire dataset, with colors representing feature values from high (red) to low (blue).

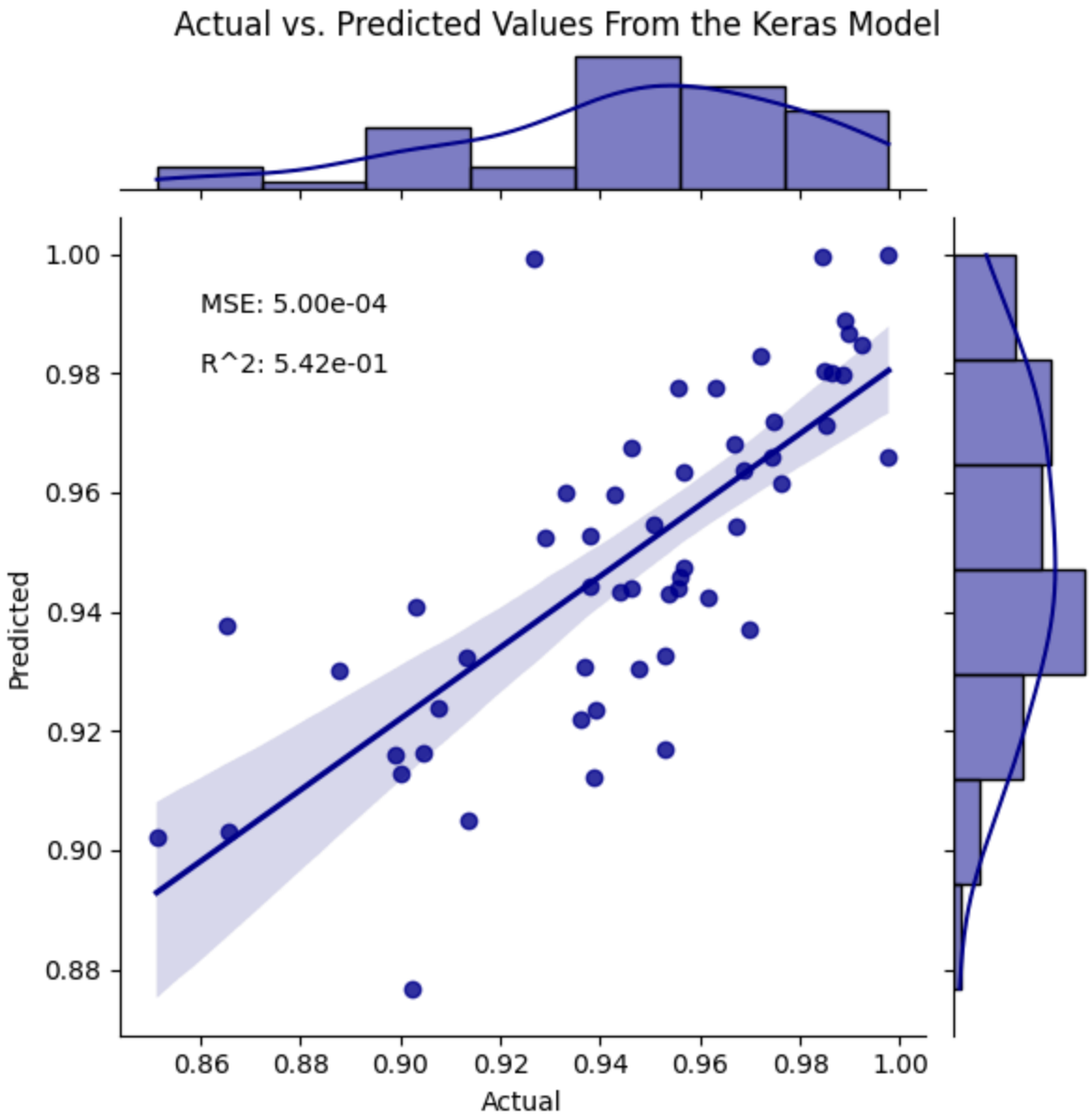

Figure 13. Actual versus predicted response values from the keras neural network model. This model performed much better than the beta regression model as the MSE here is much lower. This modeling type does not assume a linear relationship between the data, which may be the reason why this model performed better than the beta regression model.

Discussion

Our analysis confirms that demographics and clinical factors significantly affect transplant outcomes. Donor type, early survival, and organ type are all key variables that are predictive of 1 year survival after transplant. By integrating exploratory data analysis with both beta regression and neural network models, we identified consistent patterns that reinforce and expand upon existing knowledge in the transplant field. These findings correspond with previous clinical studies that have also identified similar patterns in existing transplant research (Bae et al., 2020).

Early post-transplant survival, specifically being alive at 30 days, was consistently a strong predictive factor of 1 year survival across all analytical approaches. In our exploratory data analysis, 30 day survival showed a moderately strong positive correlation with 1 year survival ($R^2 = 0.69$), indicating that patients who recover well early are more likely to experience longer term success. This finding is supported by prior work emphasizing the importance of early recovery (Yoo et al., 2017). This relationship was statistically significant in the beta regression model ($p < 0.0001$), further reinforcing its importance at the population level. Additionally, in the neural network model, 30 day survival emerged as the most influential feature contributing to the individual survival prediction, underscoring this variable as a promising candidate for a simplified, actionable early warning indicator in transplant monitoring and patient care (Badrouchi et. al, 2023).

Donor type, particularly whether the donor was living or deceased, was a strong and statistically significant predictor of survival. Our EDA showed that deceased donor recipients had worse 1 year survival and was reinforced by our beta regression model where it was a significant negative predictor ($p<0.0001$). Interestingly, it did not appear among the top contributors in the neural network's prediction output. This discrepancy is likely due to the case-by-case nature of feature importance in neural networks, unlike regression models, which estimate average effects across the entire dataset. As a result, feature contributions are localized to each specific case, meaning donor type may be an important predictor globally but might not significantly impact the survival prediction for a given case. Nonetheless, the consistent findings across EDA and regression supports the well-established clinical preference for living donor transplants when available, echoing findings from OPTN/SRTR (2020) and other large-scale transplant studies highlighting disparities in organ quality and clinical outcomes based on donor source (Javed et al., 2023). This result is likely due to improved organ preservation and overall better organ quality in living donations. Clinically, this underscores the importance of expanding living donor programs and prioritizing these transplants when appropriate.

Organ type had interesting and insightful findings across all three analytical approaches. Lung transplants were associated with the poorest clinical outcomes. In our EDA, they showed the lowest median values for both 30 day and 1 year survival. This trend was supported by the beta regression model, where lung transplants were significantly negatively associated with 1 year

survival ($p < 0.001$). However, lungs did not appear among the top contributors in the neural network model for the specific prediction shown. In contrast, kidney transplants had the most favorable outcomes in EDA, yet were not statistically significant in the regression model and contributed negatively in the neural network prediction. Liver transplants, which showed intermediate performance in EDA, were also not significant in the regression but appeared as a slight positive contributor in the neural network. Taken together, these mixed findings indicate that while lungs may predict poorer outcomes at the population level, the success rate of other organ transplants—like kidney and liver—are more nuanced and varied in predictive strength across models (Abdollahzade et al., 2025). Therefore, organ type alone may not universally predict transplant success but rather should be interpreted in the broader context of patient-specific clinical variables (Stephenson et al., 2020).

Age group composition also was an interesting factor in transplant outcomes. In our exploratory analysis, higher proportions of younger recipients (ages 18–30 and 31–40) were positively correlated with 1-year survival, while older age groups (51–60 and 61+) showed negative associations. These trends were statistically significant in regression models, emphasizing the potential impact of age distribution on center-level outcomes. However, age features did not rank among the top predictors in the beta regression model and contributed minimally in the neural network's prediction outputs. This suggests that while age may influence survival patterns across centers, its predictive power is diminished when accounting for other factors such as donor type, organ type, and early post-transplant survival (Ravindhran et al., 2023).

Conclusion

Despite the strengths of our dual-model approach, this study has several limitations that should be considered. Our dataset was center-level, limiting insight into individual-level interpretations, and focused only on short-term (1-year) survival. Notably, the model learning model like the Keras Neural Network provides only case-specific explanations and interpretations should carefully reflect its non-generalizability across populations. Moreover, our study has yet to address specific postoperative care variables - yet we are aware that there is prior literature that underscores that postoperative management plays a critical role in outcomes (Coombs, 2024). These gaps highlight the need for more integrative models that include care quality and institutional practices. Furthermore, while living donor transplant consistently outperformed deceased donor transplant, expanding living donor programs could be highly dependent on logistical or socioeconomic factors that were not captured in our dataset. Future studies should explore these implementation barriers to translate predictive insights into policy and practice (Raji et al., 2025). Integrating longitudinal data across multiple time points and incorporating real-time measures of postoperative care while applying interpretable machine learning techniques may enhance both model generalizability and utility. We are hopeful that continued collaborative efforts between clinicians, health policy makers and biological data scientists will help develop tools to predict transplant outcomes and drive improvements in patient care.

References

Abdollahzade, A., Rahimi, H., Ahmadzade, A. M., Khoshrounejad, F., Rahimi, A., Jamalirad, H., Eslami, S., Aliakbarian, M., & Khodashahi, R. (2025). *Predictive models for post-liver transplant survival using machine learning techniques in three critical time intervals*. Journal of Liver Transplantation, 17, 100253. https://doi.org/10.1016/j.liver.2024.100253

Bae, S., Massie, A. B., Caffo, B. S., Jackson, K. R., & Segev, D. L. (2020). *Machine learning to predict transplant outcomes: Helpful or hype?* Transplant International, 33(11), 1472–1480. https://doi.org/10.1111/tri.13695

Badrouchi, S., Bacha, M. M., Ahmed, A., Ben Abdallah, T., & Abderrahim, E. (2023). *Predicting long-term outcomes of kidney transplantation in the era of artificial intelligence*. Scientific Reports, 13, Article 21273. https://doi.org/10.1038/s41598-023-48645-w

Coombs, A. V. (2024). *Postoperative care*. In Merck Manual Professional Version. Retrieved from https://www.merckmanuals.com/professional/special-subjects/care-of-the-surgical-patient/postoperative-care

Ferrari, S., & Cribari-Neto, F. (2004). Beta regression for modelling rates and proportions. *Journal of Applied Statistics*, *31*(7), 799–815. https://doi.org/10.1080/0266476042000214501

Javed, H., Olanrewaju, O. A., Owusu, F. A., Saleem, A., Pavani, P., Tariq, H., Ortiz, B. S. V., Ram, R., & Varrassi, G. (2023). *Challenges and solutions in postoperative complications: A narrative review in general surgery*. Cureus, 15(12), e50942. https://doi.org/10.7759/cureus.50942

McElroy, L. M., Soranno, D. E., King, E. A., & Pittman, D. J. (2022). A scoping review of inequities in access to organ transplant in the United States. *International Journal for Equity in Health, 21*, Article 22. https://doi.org/10.1186/s12939-021-01616-x

Merola, J., Pei, K. Y., Rodriguez-Davalos, M. I., Gan, G., Deng, Y., Mulligan, D. C., & Davis, K. A. (2016). Attitudes toward organ donation among waitlisted transplant patients: results of a cross-sectional survey. *Clinical Transplantation*, *30*(11), 1449–1456. https://doi.org/10.1111/ctr.12839

Raji, C. G., Chandra, S. S. V., Gracious, N., Pillai, Y. R., & Sasidharan, A. (2025). *Advanced prognostic modeling with deep learning: Assessing long-term outcomes in liver*

transplant recipients from deceased and living donors. Journal of Translational Medicine, 23, Article 188. https://doi.org/10.1186/s12967-025-06183-1

Ravindhran, B., Chandak, P., Schafer, N., Kundalia, K., Hwang, W., Antoniadis, S., Haroon, U., & Zakri, R. H. (2023). *Machine learning models in predicting graft survival in kidney transplantation: Meta-analysis*. BJS Open, 7(2), zrad011. https://doi.org/10.1093/bjsopen/zrad011

Smithson, M., & Verkuilen, J. (2006). A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. Psychological Methods, 11(1), 54–71. https://doi.org/10.1037/1082-989X.11.1.54

Stephenson, C., Mohabbat, A., Raslau, D., Gilman, E., Wight, E., & Kashiwagi, D. (2020). *Management of common postoperative complications*. Mayo Clinic Proceedings, 95(11), 2540–2554. https://doi.org/10.1016/j.mayocp.2020.03.020

Yoo, K. D., Noh, J., Lee, H., Kim, D. K., Lim, C. S., Kim, Y. H., Pyo, J., Kim, G., & Kim, Y. S. (2017). *A machine learning approach using survival statistics to predict graft survival in kidney transplant recipients: A multicenter cohort study*. Scientific Reports, 7, Article 8904. https://doi.org/10.1038/s41598-017-08008-8

Packages and Learning Materials Used

API Documentation | TensorFlow v2.16.1. (n.d.). TensorFlow. Retrieved April 11, 2025, from https://www.tensorflow.org/api_docs

beeswarm plot—SHAP latest documentation. (n.d.). Retrieved April 11, 2025, from https://shap.readthedocs.io/en/latest/example_notebooks/api_examples/plots/beeswarm.html

Beta distribution. (2025). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Beta_distribution&oldid=1284886665

Boorman, G. (n.d.). *Supervised Learning with scikit-learn—DataCamp Learn.* Retrieved April11 ,2025, from https://app.datacamp.com/learn/courses/supervised-learning-with-scikit-learn

Esteban, Miguel. (n.d.). *Introduction to Deep Learning with Keras—DataCamp Learn*. Retrieved April 11, 2025, from https://app.datacamp.com/learn/courses/introduction-to-deep-learning-with-keras

Josef Perktold, Skipper Seabold, Kevin Sheppard, ChadFulton, Kerby Shedden, jbrockmendel, j-grana6, Peter Quackenbush, Vincent Arel-Bundock, Wes McKinney, Ian Langmore, Bart Baker, Ralf Gommers, yogabonito, s-scherrer, Yauhen Zhurko,

Matthew Brett, Enrico Giampieri, yl565, … Yaroslav Halchenko. (2024). *statsmodels/statsmodels: Release 0.14.2* (Version v0.14.2) [Computer software]. Zenodo. https://doi.org/10.5281/ZENODO.593847

Lundberg, S. (n.d.). *An introduction to explainable AI with Shapley values—SHAP latest documentation*. Retrieved April 11, 2025, from https://shap.readthedocs.io/en/latest/example_notebooks/overviews/An%20introduction%20to%20explainable%20AI%20with%20Shapley%20values.html

math—Mathematical functions. (n.d.). Python Documentation. Retrieved April 11, 2025, from https://docs.python.org/3/library/math.html

Matplotlib documentation—Matplotlib 3.10.1 documentation. (n.d.). Retrieved April 11, 2025, from https://matplotlib.org/stable/index.html

NumPy user guide—NumPy v2.2 Manual. (n.d.). Retrieved April 11, 2025, from https://numpy.org/doc/stable/user/index.html#user

os—Miscellaneous operating system interfaces. (n.d.). Python Documentation. Retrieved April 11, 2025, from https://docs.python.org/3/library/os.html

Patsy—Describing statistical models in Python—Patsy 0.5.1+dev documentation. (n.d.). Retrieved April 11, 2025, from https://patsy.readthedocs.io/en/latest/

Scikit-learn: Machine learning in Python—Scikit-learn 1.6.1 documentation. (n.d.). Retrieved April 11, 2025, from https://scikit-learn.org/stable/index.html

SciPy documentation—SciPy v1.15.2 Manual. (n.d.). Retrieved April 11, 2025, from https://docs.scipy.org/doc//scipy/index.html

Sigmoid function. (2025). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Sigmoid_function&oldid=1283662194

Smithson, M., & Verkuilen, J. (2006). A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychological Methods*, *11*(1), 54–71. https://doi.org/10.1037/1082-989X.11.1.54

StandardScaler. (n.d.). Scikit-Learn. Retrieved April 11, 2025, from https://scikit-learn/stable/modules/generated/sklearn.preprocessing.StandardScaler.html

statistics—Mathematical statistics functions. (n.d.). Python Documentation. Retrieved April 11, 2025, from https://docs.python.org/3/library/statistics.html

Team, K. (n.d.-a). *Keras documentation: Keras 3 API documentation*. Retrieved April 11, 2025, from https://keras.io/api/

Team, K. (n.d.-b). *Keras documentation: The Sequential model*. Retrieved April 11, 2025, from https://keras.io/guides/sequential_model/

User Guide—Pandas 2.2.3 documentation. (n.d.). Retrieved April 11, 2025, from https://pandas.pydata.org/pandas-docs/stable/user_guide/index.html#user-guide

warnings—Warning control. (n.d.). Python Documentation. Retrieved April 11, 2025, from https://docs.python.org/3/library/warnings.html

Waskom, M. (2021). seaborn: Statistical data visualization. *Journal of Open Source Software*, *6*(60), 3021. https://doi.org/10.21105/joss.03021

QR Code to GitHub