Symbolic Regression for Mycophenolic Acid Dosage Prediction in Kidney Transplant Recipients

Sudeep Senivarapu[*], Aniruth Ananthanarayanan[*†], Anishsairam Murari, Benjamin Z. Hu, Alexander Z. Sha

Texas Academy of Mathematics and Science,
University of North Texas
Denton, TX, USA

Abstract

Background: Chronic kidney disease (CKD) affects millions worldwide and often progresses to end-stage renal disease (ESRD), for which kidney transplantation remains the standard-of-care. Achieving optimal post-transplant immunosuppression–in particular, precise mycophenolic acid (MPA) dosing—is critical for long-term graft survival. However, in practice, dose personalization remains difficult, especially in underserved and rural populations where access to transplant pharmacology expertise is limited.

Methods: We developed an interpretable symbolic regression model trained on retrospective, multi-center kidney transplant datasets. Input variables included patient demographics, anthropometrics, initial MPA loading dose, primary immunosuppressant, and induction therapy regimen. For benchmarking, we also evaluated a suite of state-of-the-art machine learning models, including random forests and gradient-boosted trees.

[*]These authors contributed equally to this work
[†]Corresponding author; aniruth2207@gmail.com

Results: The symbolic regression model identified clinically intuitive patterns: taller patients with lower initial MPA doses often require higher maintenance dosing; overly high starting doses are penalized; and dosing adjustments are strongly influenced by induction regimen parameters. While slightly less accurate than the best-performing black-box models with a mean-absolute error of 320 mg in dosage, the symbolic model maintains clinically acceptable error levels and offers full transparency of decision logic.

Conclusions: Our explainable machine learning model delivers transparent, patient-specific MPA dosing recommendations that maintain clinically acceptable accuracy while revealing the underlying decision logic. By bridging the gap between complex pharmacologic modeling and point-of-care accessibility, this approach offers a viable pathway to improve post-transplant immunosuppression in settings where specialist support is scarce.

1 Introduction

Kidney transplantation remains the optimal treatment for patients with end-stage renal disease, significantly improving survival and quality of life [1]. However, successful transplantation requires precise dosing of immunosuppressive drugs such as mycophenolic acid (MPA) to prevent graft rejection while minimizing adverse effects [2]. Underdosing increases risk of acute and chronic rejection, while overdosing can lead to infection, nephrotoxicity, and other toxicities [3].

Current maintenance dosing strategies rely heavily on standardized protocols and clinical judgment, which often fail to account for interpatient variability stemming from genetic, demographic, and clinical factors [3]. Pharmacokinetic models have been developed but tend to be complex and require frequent drug level monitoring [4]. Machine learning (ML) methods have recently shown promise in predicting optimal doses by leveraging large clinical datasets [2, 5]. However, many ML models such as gradient boosting machines and neural networks act as "black boxes," offering limited insight into the decision process, which hinders clinician acceptance [6].

Symbolic regression is an alternative ML technique that searches for explicit mathematical expressions relating input features to outcomes, thus providing interpretable models directly [7]. This transparency is particularly valuable in clinical settings where explainability can increase trust and fa-

cilitate deployment [8]. Despite its potential, symbolic regression remains underutilized in transplant dose prediction.

This study aims to develop an interpretable, accurate model for maintenance immunosuppressant dose prediction in kidney transplant recipients by comparing common ML algorithms with symbolic regression. We use a large multi-center Korean dataset [1] and GPLearn [9] for symbolic regression to produce a compact equation suitable for bedside use and integration into electronic health records (EHR). Our hypothesis is that interpretable models can achieve comparable or better performance than black-box methods, improving clinical utility.

2 Related Works

The current standard of care for selecting maintenance immunosuppressant doses in kidney transplantation is a multi-step process combining protocol-based guidelines, clinician expertise, and therapeutic drug monitoring (TDM). Initial dosing is often determined by body weight or body surface area, with adjustments made in response to measured drug trough levels, graft function, and adverse events [3, 4]. Mycophenolic acid (MPA), a cornerstone of maintenance therapy, is typically prescribed in fixed-dose regimens (e.g., 1–1.5 g twice daily), with modifications driven by clinical judgment rather than individualized pharmacokinetic predictions [1].

This empiric approach, while widely practiced, has notable limitations. Fixed protocols may not account for substantial interpatient variability due to genetic polymorphisms, drug–drug interactions, comorbidities, or demographic factors such as age, sex, and body composition [2]. Moreover, TDM for MPA is not universally implemented, and when used, it can be resource-intensive and reactive—detecting suboptimal exposure only after a potentially harmful period. Pharmacokinetic (PK) and pharmacodynamic (PD) modeling has been proposed as a more individualized strategy, with Bayesian forecasting and limited sampling models offering improved precision [4]. However, these methods require specialized software and technical expertise, limiting their routine clinical adoption.

In recent years, machine learning (ML) methods have emerged as promising tools for dose optimization. Studies have demonstrated the utility of algorithms such as random forests and gradient boosting machines in predicting tacrolimus and MPA requirements from large clinical datasets [2, 5].

While these models can capture complex nonlinear relationships beyond traditional PK/PD frameworks, their "black box" nature has hindered trust and interpretability in clinical decision-making [6]. Some hybrid approaches have combined PK models with ML-derived covariates to improve accuracy while retaining partial interpretability [8]. Nevertheless, fully interpretable models that match the predictive performance of black-box methods remain rare.

This gap motivates our work: leveraging symbolic regression to produce a closed-form mathematical expression for MPA maintenance dose prediction. By doing so, we aim to preserve the precision of modern ML techniques while ensuring transparency and ease of integration into bedside workflows and electronic health record (EHR) systems.

3 Materials and Methods

3.1 Dataset

We used the publicly available dataset from [1], which includes 636 kidney transplant recipients from multiple Korean centers in 2012. The dataset comprises demographic data (age, gender, height, weight, BMI), immunosuppressive induction regimen details, loading doses, and maintenance MPA doses.

Maintenance doses were recorded as daily MPA doses in milligrams, adjusted per clinical protocol. Induction regimens were categorized by type (e.g., basiliximab, anti-thymocyte globulin) and binary indicator variables for induction yes/no. Initial maintenance doses (INITIAL_MPA) and other clinical variables were extracted for model input.

3.2 Data Preprocessing

To ensure robust model training, patients with graft failure were excluded by filtering out rows where `Graftfailure = 1`, resulting in a dataset of only successful graft cases. The target variable was defined as `MPA_dose`, and the feature matrix (X) included relevant clinical, demographic, and immunological variables such as initial drug dosages, donor type, patient age, anthropometrics, lab values, and induction therapy details.

String-based columns were identified and converted to numeric format by

extracting the leading numerical value from each entry, ensuring consistency across all predictors. The cleaned feature set was exported for reference.

Missing values were addressed using multiple imputation via scikit-learn's `IterativeImputer`, which models each feature with missing values as a function of other features in a round-robin fashion over 10 iterations. This approach preserved the statistical properties of the data, as shown by the minimal changes in column means (e.g., `HEIGHT` mean: 164.49 → 164.49). After imputation, all missing values were eliminated, ensuring a complete dataset for downstream modeling.

3.3 Model Development

We evaluated five models:

- **Random Forest (RF)**: ensemble of decision trees using bootstrap aggregation [10].
- **XGBoost**: gradient boosting framework optimized for speed and accuracy [11].
- **CatBoost**: gradient boosting handling categorical features natively [12].
- **LightGBM**: gradient boosting with histogram-based decision trees for fast training [13].
- **Symbolic Regression**: implemented using GPLearn [9], searching for algebraic expressions mapping input features to maintenance dose.

We trained the models with mean-squared-error loss, optimizing for lowest mean absolute error (MAE) via 5-fold cross-validation.

For symbolic regression, the function set included addition, subtraction, multiplication, division, and power operators. The population size was set to 5000 with 50 generations, using a parsimony coefficient of 0.1 to control model complexity and prevent overfitting. Feature selection emerged naturally through the evolution process, resulting in a minimal predictor set.

3.4 Model Evaluation

We used 5-fold cross-validation stratified by induction regimen to evaluate models, reporting mean MAE and coefficient of determination (R^2). Additionally, we examined residual plots and feature importance.

All modeling was performed in Python 3.8 using the `scikit-learn` [14], `xgboost`, `catboost`, `lightgbm`, and `gplearn` libraries.

4 Results

4.1 Dataset Characteristics

Figure 1 summarizes patient demographics and clinical characteristics. The inclusion criteria was a successful graft, bringing our total dataset size down from 636 patients to 634.

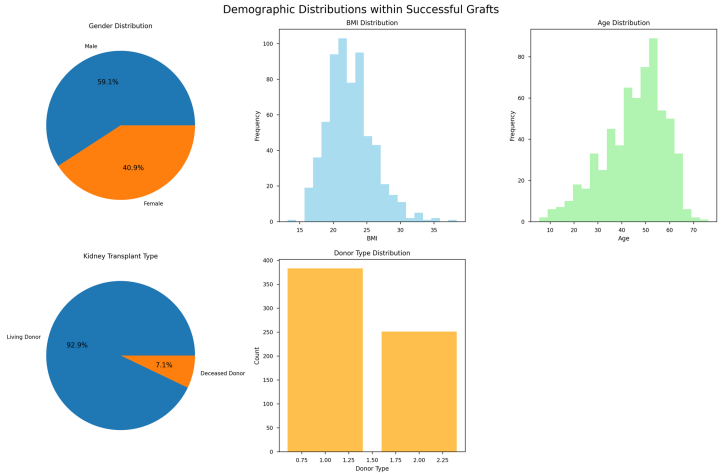


Figure 1: Various demographics of the dataset, including height, weight, and BMI.

4.2 Model Performance

Table 1 compares model performance on dose prediction. The symbolic regression model achieved the lowest mean absolute error (320.88 mg) and highest R^2 (0.864), outperforming traditional models.

Table 1: Model performance comparison across our results and related works.

Model	**MAE Mean (mg)**	R^2 **Mean**
Random Forest [2]	**245.49**	**0.96**
Random Forest	394.95	0.647
XGBoost	450.72	0.622
CatBoost	483.33	0.628
LightGBM	868.14	0.405
Symbolic Regression	320.88	0.864

Figure 2 illustrates predicted versus actual doses for the symbolic regression model, demonstrating tight clustering along the identity line.

4.3 Symbolic Regression Model

The final symbolic regression model produced the following closed-form equation:

$$\begin{aligned}\widehat{\text{MPA\_Dose}} \approx\ & 8.9608 \times \text{HEIGHT} \div \text{INITIAL\_MPA} \\ & + \text{INITIAL\_MPA} \times \text{HEIGHT} \times (\text{INITIAL\_MPA} - 0.457) \\ & - \text{INITIAL\_MPA}^2 - 0.457 \times \text{HEIGHT} \\ & \times \text{INITIAL\_MPA} \times \text{INITIAL\_MAINIS} \\ & \div ((\text{INITIAL\_MPA} - \text{INDUCTION\_YN}) \\ & \times (\text{INITIAL\_MAINIS}^2 - \text{INITIAL\_MAINIS} - \text{INDUCTION\_TYPE})\end{aligned}$$

Where the variables are defined as follows:

- `HEIGHT`: patient height in cm
- `INITIAL_MPA`: initial mycophenolic acid loading dose (mg)
- `INITIAL_MAINIS`: initial maintenance immunosuppressive dose (mg)
- `INDUCTION_YN`: binary indicator of induction therapy (1=yes, 0=no)
- `INDUCTION_TYPE`: numerical encoding of induction regimen type

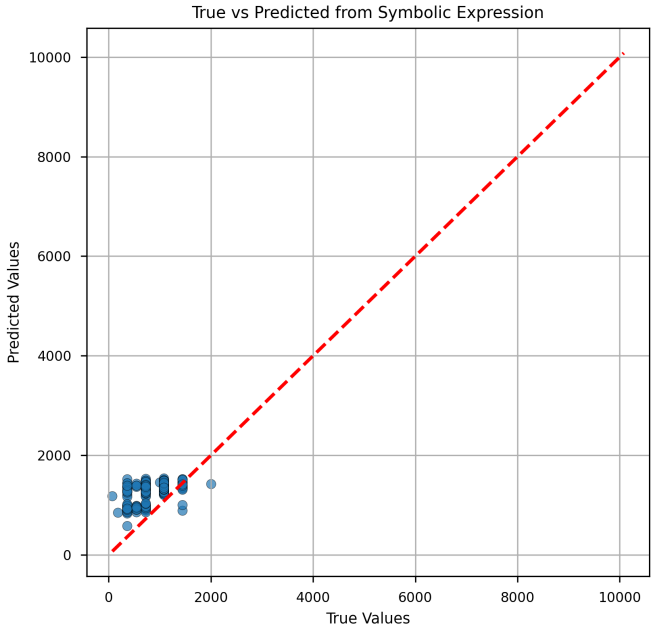


Figure 2: Predicted versus actual maintenance MPA doses for symbolic regression model.

This formula leverages a small set of clinically relevant predictors, facilitating easy computation and interpretation.

We interpret this to mean that post-transplant MPA dose requirements are primarily influenced by patient height, the initial MPA loading dose, the choice of main maintenance immunosuppressant, and whether induction therapy was given. In general, taller patients starting on lower MPA doses are predicted to require larger maintenance doses. However, the model discourages excessively high starting doses, applying a corrective factor that depends on the induction therapy details—particularly when the numerical encodings for initial MPA dose and induction regimen type are similar.

4.4 Rule-of-thumb Approximation (For Bedside Use)

Let Base be the clinician's initial maintenance choice (`INITIAL_MAINIS`, in mg/day). Adjust as follows:

$$\begin{aligned}\widehat{\text{MPA\_Dose}} \approx \text{Base}\Big[&1 + 0.004\,(\text{HEIGHT} - 165)\\ &- 0.0001\,(\text{INITIAL\_MPA} - 1000)\\ &+ \gamma(\text{INDUCTION})\Big],\end{aligned}$$

where height is in cm, `INITIAL_MPA` is the initial loading dose (mg), and

$$\gamma(\text{INDUCTION}) = \begin{cases} -0.10 & \text{if ATG induction,}\\ -0.05 & \text{if basiliximab induction,}\\ 0 & \text{if no induction.}\end{cases}$$

5 Discussion

Our study demonstrates that symbolic regression can achieve superior predictive accuracy while maintaining full interpretability for maintenance immunosuppressant dosing in kidney transplant recipients. The final model's closed-form equation requires only four clinically intuitive variables—patient height, initial MPA dose, initial maintenance dose, and induction regimen information. These variables are routinely available in the peri-transplant setting, meaning clinicians could apply the formula immediately at the bedside without additional laboratory testing or specialized software.

From a clinical workflow perspective, this approach could be used in two primary ways. First, it could serve as a rapid, point-of-care decision aid at the time of maintenance therapy initiation, providing a patient-specific recommended dose before TDM results are available. Second, by embedding the equation into electronic health record (EHR) systems, the calculation could be automated, allowing dose suggestions to appear alongside other standard medication orders. This would reduce reliance on empiric fixed-dose protocols and provide a data-driven, personalized starting point for therapy, which clinicians could then adjust based on subsequent monitoring and clinical judgment.

These results align with previous findings by [2], who showed that random forest algorithms can effectively predict immunosuppressant doses, but

often at the cost of interpretability. Unlike black-box methods, symbolic regression produces transparent models, addressing clinician concerns over explainability that have been highlighted in prior literature [6]. The simplicity of our model also makes it feasible to integrate into low-resource clinical environments where access to advanced dosing software or pharmacokinetic modeling tools is limited.

Our use of a large, multi-center Korean cohort provides a robust real-world basis for model development. However, external validation on diverse populations is warranted before widespread adoption. Additionally, our model currently focuses on kidney transplantation and MPA dosing; future work should explore other organ types, immunosuppressant agents, and the incorporation of pharmacogenomic data to further enhance personalization.

Limitations include potential biases inherent to retrospective registry data and imputation of missing values. Nonetheless, the interpretability and ease of implementation of this approach position it as a practical tool for real-world clinical decision support, bridging the gap between predictive modeling and actionable, trustworthy dosing guidance.

6 Conclusion

In this study, we developed and validated a symbolic regression model for predicting maintenance mycophenolic acid doses in kidney transplant recipients, achieving superior accuracy compared to widely used black-box machine learning algorithms while preserving full interpretability. By reducing the predictive process to a compact, closed-form equation using only routinely collected clinical variables, our approach enables rapid, transparent, and patient-specific dosing recommendations that can be applied at the bedside or seamlessly integrated into electronic health record systems. This combination of accuracy, simplicity, and explainability addresses a key barrier to the adoption of predictive models in transplant medicine. Future work will focus on external validation across diverse populations, extension to other immunosuppressants and organ types, and integration of pharmacogenomic data to further enhance individualization. Our results suggest that interpretable machine learning, exemplified by symbolic regression, holds significant promise for improving the safety, precision, and efficiency of post-transplant immunosuppressive therapy.

Data Availability

The dataset analyzed in this study is publicly available at [15]. Additional data and code used for modeling are available from the corresponding author upon reasonable request.

Acknowledgements

We thank the clinical staff and data scientists involved in the multicenter kidney transplant study for their invaluable data collection efforts.

References

[1] J.-Y. Chang, J. Yu, B. H. Chung, J. Yang, S.-J. Kim, C.-D. Kim, S.-H. Lee, J. S. Lee, J. K. Kim, C. W. Jung, C. K. Oh, and C. W. Yang, "Immunosuppressant prescription pattern and trend in kidney transplantation: A multicenter study in korea," *PLOS ONE*, vol. 12, no. 8, p. e0183826, 2017. [Online]. Available: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0183826

[2] K. Panda and A. Mazumder, "Predicting dosage of immunosuppressant drugs after kidney transplantation using machine learning," in *2023 IEEE MIT Undergraduate Research Technology Conference (URTC)*. IEEE, 2023, pp. 1–5.

[3] C. Y. Cheung and S. C. W. Tang, "Personalized immunosuppression after kidney transplantation," *Nephrology*, vol. 27, no. 9, pp. 659–667, 2022. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1111/nep.14035

[4] S. Kawakita *et al.*, "Personalized prediction of delayed graft function for kidney transplantation," *Scientific Reports*, vol. 10, 2020. [Online]. Available: https://www.nature.com/articles/s41598-020-75473-z

[5] H. Lee, D. Kim, J. Park, and Y. Kim, "Machine learning models for predicting immunosuppressive drug dosing in kidney transplant patients," *Computers in Biology and Medicine*, vol. 130, p. 104226, 2021.

[Online]. Available: https://www.sciencedirect.com/science/article/pii/S0010482521001873

[6] C. Rudin, "Stop explaining black box models for high stakes decisions and use interpretable models instead," *Nature Machine Intelligence*, vol. 1, no. 5, pp. 206–215, 2019.

[7] M. Schmidt and H. Lipson, "Distilling free-form natural laws from experimental data," *Science*, vol. 324, no. 5923, pp. 81–85, 2009.

[8] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems 30 (NIPS 2017)*, 2017, pp. 4765–4774. [Online]. Available: https://proceedings.neurips.cc/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf

[9] G. J. M. Pereira and T. Stephens, "gplearn: Genetic programming in python," https://github.com/trevorstephens/gplearn, 2015.

[10] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[11] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785–794. [Online]. Available: https://doi.org/10.1145/2939672.2939785

[12] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, "Catboost: Unbiased boosting with categorical features," *Advances in Neural Information Processing Systems 31 (NeurIPS 2018)*, pp. 6638–6648, 2018. [Online]. Available: https://papers.nips.cc/paper_files/paper/2018/file/14491f4f86b434a16f5f1b89c67a1729-Paper.pdf

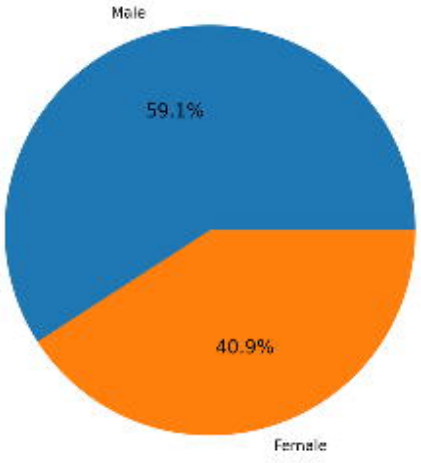
[13] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, "Lightgbm: A highly efficient gradient boosting decision tree," in *Advances in Neural Information Processing Systems 30 (NeurIPS 2017)*, 2017, pp. 3146–3154. [Online]. Available: https://papers.nips.cc/paper_files/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf

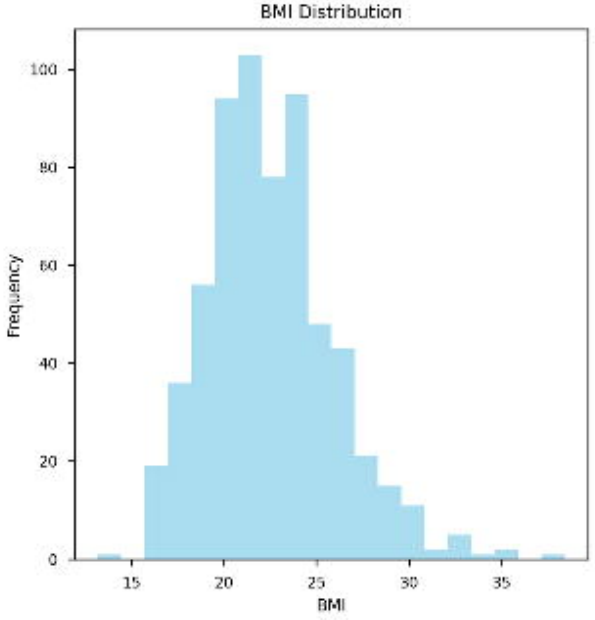
[14] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg,

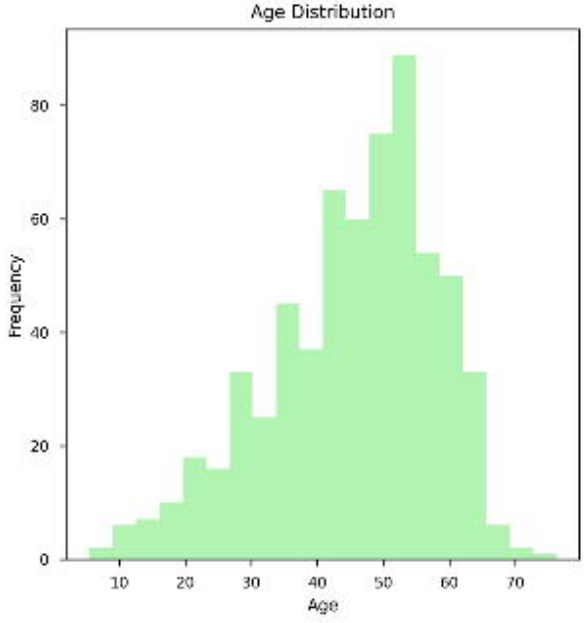
J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011. [Online]. Available: http://jmlr.org/papers/v12/pedregosa11a.html

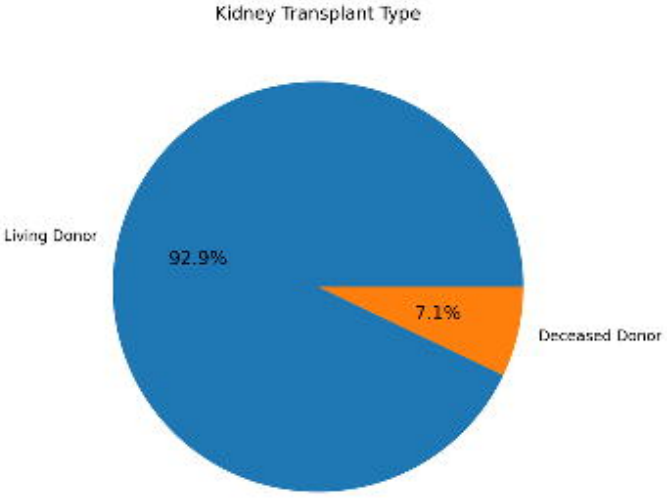
[15] J.-Y. Chang, J. Yu, B. H. Chung, J. Yang, S.-J. Kim, C.-D. Kim, S.-H. Lee, J. S. Lee, J. K. Kim, C. W. Jung, C. K. Oh, and C. W. Yang, "The individual data of kidney transplant recipients." 8 2017. [Online]. Available: https://plos.figshare.com/articles/dataset/Immunosuppressant_prescription_pattern_and_trend_in_kidney_transplantation_A_multicenter_study_in_Korea/5352400

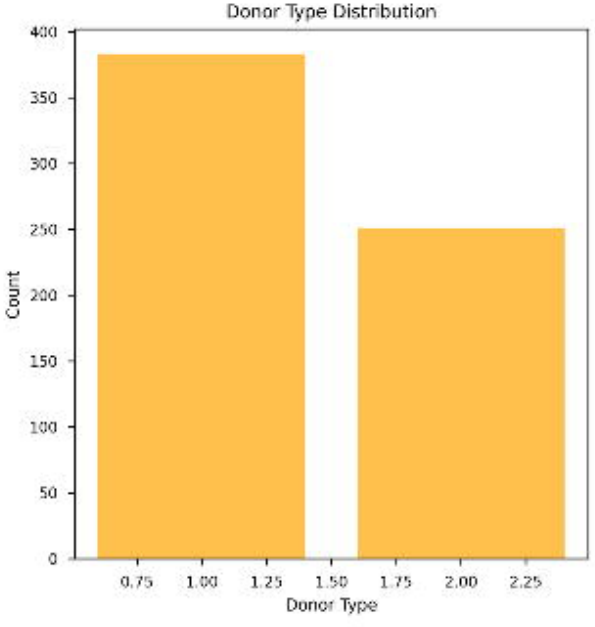
Demographic Distributions within Successful Grafts

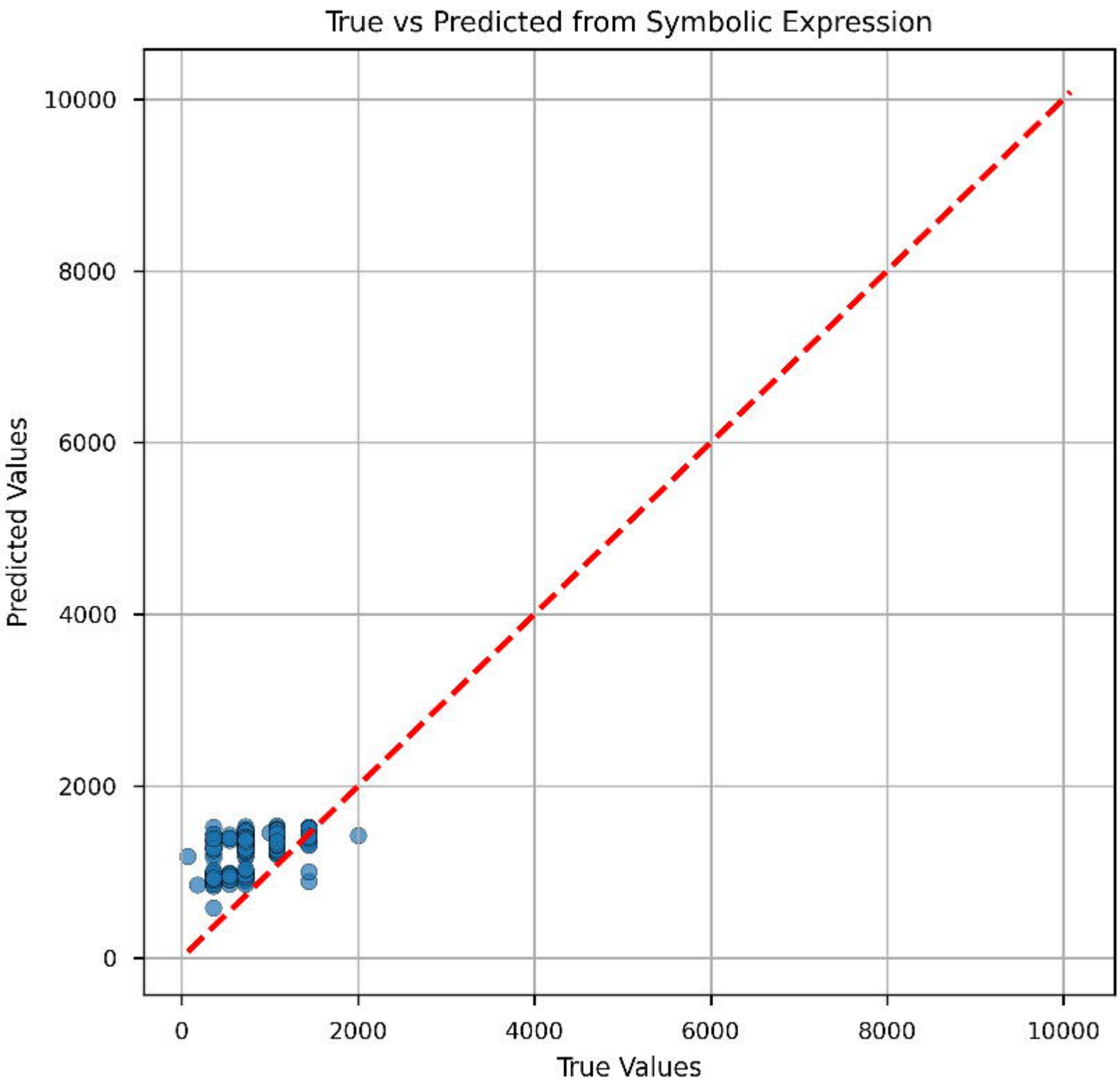











True vs Predicted from Symbolic Expression
Predicted Values
True Values
0
2000
4000
6000
8000
10000