

Original Paper

# Machine-Learning Predictive Tool for the Individualized Prediction of Outcomes of Hematopoietic Cell Transplantation for Sickle Cell Disease: Registry-Based Study

Rajagopal Subramaniam Chandrasekar<sup>1,2</sup>, MTech; Michael Kane<sup>3</sup>, PhD, MA, MS; Lakshmanan Krishnamurti<sup>1</sup>, MD

<sup>1</sup>Section of Pediatric Hematology/Oncology/BMT, Yale School of Medicine, New Haven, CT, United States

<sup>2</sup>School of Data Science and AI, Indian Institute of Technology, Madras, Chennai, India

<sup>3</sup>Department of Lymphoma-Myeloma, The University of Texas MD Anderson Cancer Center, Houston, TX, United States

## Corresponding Author:

Lakshmanan Krishnamurti, MD  
Section of Pediatric Hematology/Oncology/BMT  
Yale School of Medicine  
2073 A, LMP Building 330 Cedar Street  
New Haven, CT  
United States  
Phone: 1 412-612-4761  
Email: [lakshmanan.krishnamurti@yale.edu](mailto:lakshmanan.krishnamurti@yale.edu)

## Abstract

**Background:** Disease-modifying therapies ameliorate disease severity of sickle cell disease (SCD), but hematopoietic cell transplantation (HCT), and more recently, autologous gene therapy are the only treatments that have curative potential for SCD. While registry-based studies provide population-level estimates, they do not address the uncertainty regarding individual outcomes of HCT. Computational machine learning (ML) has the potential to identify generalizable predictive patterns and quantify uncertainty in estimates, thereby improving clinical decision-making. There is no existing ML model for SCD, and ML models for HCT for other diseases focus on single outcomes rather than all relevant outcomes.

**Objective:** This study aims to address the existing knowledge gap by developing and validating an individualized ML prediction model SPRIGHT (Sickle Cell Predicting Outcomes of Hematopoietic Cell Transplantation), incorporating multiple relevant pre-HCT features to make predictions of key post-HCT clinical outcomes.

**Methods:** We applied a supervised random forest ML model to clinical parameters in a deidentified Center for International Blood and Marrow Transplant Research (CIBMTR) dataset of 1641 patients who underwent HCT between 1991 and 2021 and were followed for a median of 42.5 (IQR 52.5; range 0.3-312.9) months. We applied forward and reverse feature selection methods to optimize a set of predictive variables. To counter the imbalance bias toward predicting positive outcomes due to the small number of negative outcomes, we constructed a training dataset, taking each outcome as variable of interest, and performed 2-times repeated 10-fold cross-validation. SPRIGHT is a web-based individualized prediction tool accessible by smartphone, tablet, or personal computer. It incorporates predictive variables of age, age group, Karnofsky or Lansky score, comorbidity index, recipient cytomegalovirus seropositivity, history of acute chest syndrome, need for exchange transfusion, occurrence and frequency of vaso-occlusive crisis (VOC) before HCT, and either a published or custom chemotherapy or radiation conditioning, serotherapy, and graft-versus-host disease prophylaxis. SPRIGHT makes individualized predictions of overall survival (OS), event-free survival, graft failure, acute graft-versus-host disease (AGVHD), chronic graft-versus-host disease (CGVHD), and occurrence of VOC or stroke post-HCT.

**Results:** The model's ability to distinguish between positive and negative classes, that is, discrimination, was evaluated using the area under the curve, accuracy, and balanced accuracy. Discrimination met or exceeded published predictive benchmarks with area under the curve for OS (0.7925), event-free survival (0.7900), graft failure (0.8024), acute graft-versus-host disease (0.6793), chronic graft-versus-host disease (0.7320), and VOC post-HCT (0.8779). SPRIGHT revealed good calibration with a slope of 0.87-0.96, with small negative intercepts (-0.01 to 0.03), for 4 out of the 5 outcomes. However, OS exhibits nonideal calibration, which may be reflective of the overall high OS in all subgroups.

**Conclusions:** A web-based ML prediction tool incorporating multiple clinically relevant variables predicts key clinical outcomes with a high level of discrimination and calibration and has potential in shared decision-making

**Keywords:** sickle cell disease; SCD; prediction algorithms; hematopoietic stem cell transplantation; machine learning; ML; predictive tool; prediction; hematopoietic cell transplantation; HCT; hematopoietic cell; registry-based study; clinical decision-making; prediction model; clinical outcomes; gene therapy; shared decision-making

## Introduction

The complications of sickle cell disease (SCD) can be prevented or ameliorated by disease-modifying therapies [1], but hematopoietic cell transplantation (HCT), and more recently, gene therapy remain the only therapeutic options with curative intent [2-12]. Population-level studies demonstrate the association of outcomes of HCT with age, type of donor, type of conditioning, and graft-versus-host disease (GVHD) prophylaxis [5,6], but do not address the uncertainty regarding individualized outcomes of HCT. Such uncertainty contributes to the decisional dilemma and is a barrier to shared decision-making. An individualized prediction model that incorporates all predictive variables and provides individualized estimates of key outcomes of HCT of interest to patients and their physicians has the potential to inform shared decision-making [13-15]. Brazauskas et al [16] have proposed a predictive model based on the age of the recipient and the type of donor. However, their model does not incorporate other clinically relevant patient, HCT, and disease characteristics and does not include all key outcomes. Computational machine learning (ML) has the potential to determine generalizable predictive patterns and quantify uncertainty, but published ML predictive models for HCT are limited to predicting single clinical outcomes [17-30]. To address the knowledge gap, we developed and described the initial validation of SPRIGHT (Sickle Cell Predicting Outcomes of Hematopoietic Cell Transplantation), an individualized ML prediction model for outcomes of HCT for SCD, incorporating multiple relevant features to make predictions of key clinical outcomes.

## Methods

### Dataset

We developed SPRIGHT using an anonymized HCT for the SCD dataset [31] derived from data submitted to the Center for International Bone Marrow Transplant Research (CIBMTR) registry on children and adults undergoing HCT for SCD between 1991 and 2021 in the United States. The dataset was obtained through the NHLBI (National Heart Lung and Blood Institute) Biologic Specimen and Data Repository Information Coordinating Center (BIOLINCC) [32].

The CIBMTR maintains a research database to serve as a comprehensive data source that can be used to study cellular therapies, including HCT. All US transplant centers are required to submit outcomes data on all allogeneic transplants when either the stem cell donation or the transplant occurs within the United States. CIBMTR assigns patients to either a Transplant Essential Data (TED) track, which collects core transplant data, or a Comprehensive Report Form (CRF)

track that captures detailed disease- and treatment-related data [31]. Assignment to each track is based on submission of the initial pretransplantation TED 2000032419 form and uses a weighted randomization algorithm designed to produce a cohort representative of current clinical practice. All centers submit a Pre-HCT TED Form (Form 2400) for each allogeneic (related or unrelated) HCT.

Of 1641 patients undergoing HCT for SCD between 1991 and 2021, on whom data were submitted to CIBMTR, detailed CRFs were submitted on 763 patients. Of the patients in the dataset, 84% (1377/1641) had undergone HCT after 2007. We performed the imputation of missing data using MissForest, an ML data imputation algorithm that operates on random forest (RF).

### Feature Selection

We identified overall survival (OS), event-free survival (EFS), graft failure (GF), acute graft-versus-host disease (AGVHD), and chronic graft-versus-host disease (CGVHD) as key outcomes. We used wrapper methods of backward feature elimination (BFE) and forward sequential selection (FSS) to select and optimize the input variables [33,34]. The BFE procedure begins with a complete set of features and a chosen ML model. The model is trained, and the importance of each feature is evaluated based on the model's coefficients or feature importance scores, and then the least important feature is discarded, and the model is retrained on the remaining features. This process is repeated until a predetermined number of features is reached or until further removal of features leads to a significant decrease in model performance. Using FSS, we incrementally built a feature set starting with an empty model, sequentially adding a feature that most improves the model performance at each step, as evaluated through a predefined metric like cross-validation score. We continued this stepwise addition until new features no longer significantly enhanced the model or a specified number of features was reached. We have included the detailed descriptive statistics and missingness for each of the selected features and outcomes of interest in Table S3 in [Multimedia Appendix 1](#) and Table S4 in [Multimedia Appendix 2](#).

### Model Design

The model was designed by subsampling the majority class to create subtraining datasets, followed by pooling and thresholding to obtain the final prediction [35]. We determined the discriminative performance of the model, which refers to how well the predictions can separate between 2 groups of participants, that is, those with or without an outcome. Discrimination was quantified by the concordance (c) statistic (index), which for binary outcomes, is equivalent to the area under the curve (AUC). We also assessed accuracy, that is, the percentage of correct predictions out

of all predictions correct or incorrect in the model. We also assessed balanced accuracy, which is an accuracy adjusted for imbalance and is derived by averaging sensitivity and specificity, so that each class's importance is equal. We compared RF, extreme gradient boosting, logistic regression, Naive Bayes, AdaBoost, and support vector classification algorithms (Table 1). The HCT for SCD dataset spans children and adults undergoing HCT between 1991 and 2021. During this time, there have been many changes in conditioning regimens and improvements in supportive

care. Gluckman et al [6] reported that EFS was higher in patients who underwent HCT in or after 2007 as compared to those who underwent HCT in or before 2006 (HR [hazard ratio] 0.95, CI 0.90-0.99;  $P=.01$ ). To determine if the model performed consistently across eras, we tested model performance in patients with the year of HCT <2007 versus HCT >2007. We also tested model performance across age at HCT <10, <18, and >18 years, respectively, including outcomes at 1- and 3-year post-HCT (Table 2).

**Table 1.** Model performance: comparison of area under the curve of different algorithms for each outcome.

Model	Outcome of interest				
	EFS <sup>a</sup>	OS <sup>b</sup>	GF <sup>c</sup>	AGVHD <sup>d</sup>	CGVHD <sup>e</sup>
Random forest entire dataset	0.7900	0.7925	0.8024	0.6793	0.7320
XGBoost <sup>f</sup>	0.7754	0.7785	0.7948	0.6731	0.7230
Logistic regression	0.7464	0.7835	0.7578	0.6925	0.7019
Naïve Bayes	0.6930	0.7111	0.7107	0.6386	0.6384
Adaboost	0.7452	0.7806	0.7561	0.6934	0.7005
Support vector classifier	0.7357	0.7810	0.7561	0.6841	0.7061

<sup>a</sup>EFS: event-free survival.

<sup>b</sup>OS: overall survival.

<sup>c</sup>GF: graft failure.

<sup>d</sup>AGVHD: acute graft-versus-host disease.

<sup>e</sup>CGVHD: chronic graft-versus-host disease.

<sup>f</sup>XGBoost: extreme gradient boosting.

**Table 2.** Model performance: Comparison of area under the curve across different paradigms and data time periods.

Paradigm	Outcome of interest				
	EFS <sup>a</sup>	OS <sup>b</sup>	GF <sup>c</sup>	AGVHD <sup>d</sup>	CGVHD <sup>e</sup>
All data	0.790	0.793	0.802	0.679	0.732
Post 2007 data	0.787	0.775	0.783	0.702	0.729
1 Year outcome analysis on post 2007 data	0.801	0.788	0.807	0.741	0.705
3 Year outcome analysis on post 2007	0.792	0.771	0.82	0.730	0.721

<sup>a</sup>EFS: event-free survival.

<sup>b</sup>OS: overall survival.

<sup>c</sup>GF: graft failure.

<sup>d</sup>AGVHD: acute graft-versus-host disease.

<sup>e</sup>CGVHD: chronic graft-versus-host disease.

The accuracy of risk estimates, relating to the agreement between the estimated and observed number of events, is called “calibration [36,37].” Calibration is crucial in predictive algorithms because it ensures the accuracy of risk estimates, which directly affects clinical decision-making and patient expectations. Poor calibration can lead to systematic overestimation or underestimation of risk, resulting in false expectations and potentially harmful decisions [36]. We performed causal isotonic calibration, a novel nonparametric

method for calibrating predictors of heterogeneous treatment effect [37]. We performed a 5-fold internal cross-validation on the training set to determine the optimal calibration. The resultant calibration model was then applied to the predictions during evaluation. To adjust for the bias caused due to undersampling, we recalibrate the probabilities according to the method by Pozollo et al [38]. We evaluated the calibration curve, slope, and intercept across different outcomes of interest in the post-2007 data (Table 3).

**Table 3.** Calibration analysis: slope and intercept across different outcomes of interest for post-2007 data.

Outcome of interest	Calibration property	
	Slope	Intercept
EFS <sup>a</sup>	0.93	−0.02
OS <sup>b</sup>	0.75	−0.07
GF <sup>c</sup>	0.9	−0.02
AGVHD <sup>d</sup>	0.96	−0.03
CGVHD <sup>e</sup>	0.87	−0.01

<sup>a</sup>EFS: event-free survival.<sup>b</sup>OS: overall survival.<sup>c</sup>GF: graft failure.<sup>d</sup>AGVHD: acute graft-versus-host disease.<sup>e</sup>CGVHD: chronic graft-versus-host disease.

To understand the contributions of each feature to the predictive model, we use the Shapley additive explanations (SHAP) scores ([Multimedia Appendices 3–7](#)). SHAP scores are based on game theory's Shapley values, quantifying each feature's marginal contribution to individual predictions. They are calculated by measuring how each feature affects the model output when included or excluded from all possible feature combinations.

### The Problem of Imbalance

The outcomes data for HCT for SCD is imbalanced, with very few negative outcomes. This imbalance has the potential to lead to a prediction bias, where an uncorrected model default may be skewed toward predicting positive outcomes. To address the problem of imbalance, we constructed a training dataset taking each outcome as variable of interest. We included randomly sampled positive outcomes, typically 1.5-3 times the total instances of the variable of interest. To address the issue of class imbalance, we used a 2-step approach involving bootstrapping and consensus-based decision-making. Initially, we generated 20 bootstrapped datasets by undersampling the majority class to achieve a

(2-3):1 ratio with the minority class. These datasets served as the training sets for our predictive models, ensuring a balanced representation of classes during model training. Once the models were trained, each was then tested on a consistent test dataset to obtain a series of predictions. These individual predictions were subsequently pooled across all models. A final prediction model for each test instance was determined based on a consensus threshold. If the majority of the models exceeded a predetermined threshold agreed on a particular class, that class was assigned as the outcome for the instance. We ran the test dataset and used a RF algorithm on a 2-times repeated 10-fold cross-validation to demonstrate our model's versatility and response to unknown data ([Figure 1](#)). We assigned value 1 for a negative outcome prediction and −1 for a positive outcome prediction and found the average sum across the 20 trials for each element.

Throughout the paper, we are guided by the CREMLS (Consolidated Reporting of Machine Learning Studies) guidelines [39], to ensure transparency and rigor in reporting. We have attached a completed author CREMLS checklist.

**Figure 1.** Data partitioning, imputation, cross-validation, and item optimization A) Dataset preprocessing, imputation, and feature selection. (B) Model finalization and test of predictive performance. BFE: backward feature elimination; FSS: forward sequential selection.

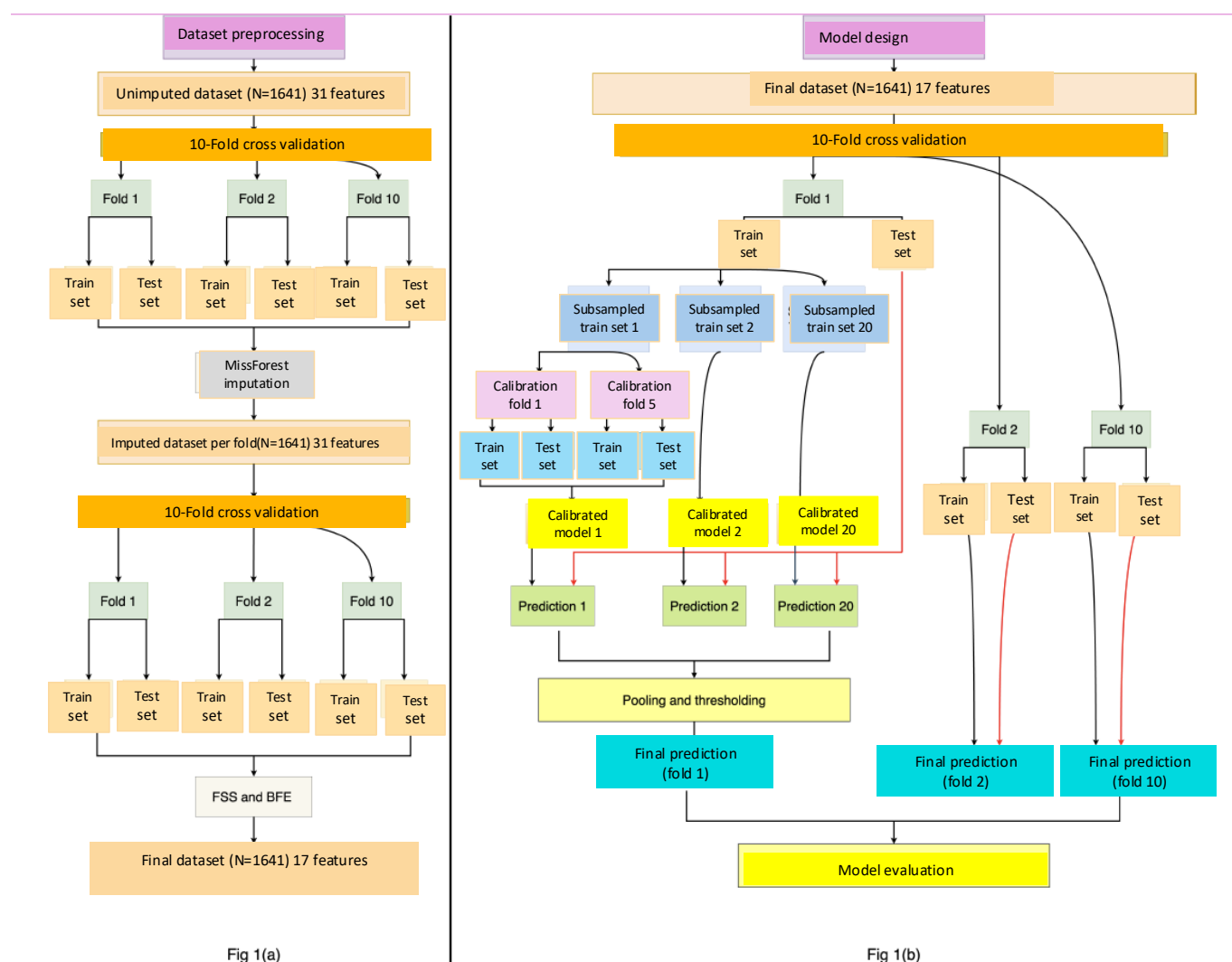


Fig 1(a)

Fig 1(b)

## Ethical Considerations

The Institutional Review Board at Yale University determined on March 3, 2023, that this study did not constitute human participants research per IRB protocol number 2000032419.

## Results

### Feature Selection and Optimization

Of the 160 variables in the dataset, we selected 31 potentially clinically relevant predictive variables. Through FSS and BFE processes, we selected a final set of 17 predictive variables grouped into 3 categories. Patient Data variables included age at transplant, age group at transplant, sex, Karnofsky or Lansky score, HCT-comorbidity index, and recipient CMV serostatus. SCD variables included Number of ACS syndromes within 2 years pre-HCT, Required exchange transfusion, Vaso-occlusive crisis needing hospitalization in 2 years pre-HCT, and Hospitalization frequency for vaso-occlusive crises. Transplant Data variables included Donor type, Graft type, Conditioning intensity, Conditioning regimen,

Serotherapy (ATG or Alemtuzumab), GVHD prophylaxis, and Donor-recipient HLA matching.

### Evaluation Outcomes and Model Performance

Discrimination in predictive performance is evaluated using accuracy, the percentage of correct predictions out of all predictions, balanced accuracy, the average sensitivity and specificity, and each class's importance is equal, and AUC, the measure of a model's true positive rate against a false positive rate, indicates the ability to differentiate classes. AUC is the metric used in published literature. The RF model achieved the highest predictive AUC (Table 1) across multiple clinical outcomes. We measure our model's performance using the benchmark established in the literature. To determine the statistical validity, we implemented the method proposed by Bouckaert and Frank [40]. We first performed 2x10 repeated cross-validation and obtained the performance of each of the models for each fold. We then applied Nadeau and Bengio's [41] correction, which accounts for training set overlap in the variance estimation, to check whether the mean AUC of the RF is greater than the mean AUC of the other models across folds. The differences in the

mean AUC were statistically significant for EFS, OS, GF, AGVHD, and CGVHD ( $P<.05$ ). We performed hyperparameter tuning using grid search cross-validation for the RF model. The ideal hyperparameters are described in [Table 4](#). AUC, accuracy, and balanced accuracy equaled or exceeded the benchmarks of the ML predictive tools in the published literature [16,42]. The RF model has been previously reported to have the best AUC in predicting survival following HCT

[24]. Brazauskas et al [16] published benchmark AUC for EFS of 0.72, and Taheriyani et al [42] reported benchmarks AUC 0.82 for AGVHD post-HCT. Accuracy and balanced accuracy were excellent for EFS (0.76, 0.69), OS (0.82, 0.68), GF (0.8, 0.71), vaso-occlusive pain post-HCT (0.9, 0.78), stroke post-HCT (0.92, 0.65), acute GVHD (0.71, 0.60), CGVHD (0.72, 0.63).

**Table 4.** Final tuned hyperparameters of the random forest model for each outcome.

Outcome of interest	Random forest hyperparameters			Criterion	CCP_alpha <sup>a</sup>	Max features
	Max depth	Min samples split	Min samples leaf			
EFS <sup>b</sup>	20	10	16	entropy	0.0	0.7
OS <sup>c</sup>	20	10	8	entropy	0.01	0.125
GF <sup>d</sup>	20	20	16	entropy	0.0	0.5
AGVHD <sup>e</sup>	5	20	8	entropy	0.0	0.5
CGVHD <sup>f</sup>	15	10	8	entropy	0.0	0.5

<sup>a</sup>CCP\_alpha: cost complexity pruning alpha.  
<sup>b</sup>EFS: event-free survival.  
<sup>c</sup>OS: overall survival.  
<sup>d</sup>GF: graft failure.  
<sup>e</sup>AGVHD: acute graft-versus-host disease.  
<sup>f</sup>CGVHD: chronic graft-versus-host disease.

SPRIGHT retained high AUC in subpopulations, including patients  $\leq 10$ ,  $\leq 18$ ,  $>18$  of age in undergoing HCT after 2007, as well as 1- and 3-year survival analysis ([Table 1](#)).

Calibration is the agreement between the estimated and observed number of events, for major outcomes. A calibration slope of 1 and an intercept close to zero are associated with good calibration. SPRIGHT revealed good calibration with a slope range of 0.87-0.96, with small negative intercepts ( $-0.01$  to  $0.03$ ), for 4 out of the 5 outcomes. However, OS exhibits nonideal calibration and may be reflective of the overall high OS in all subgroups ([Figure 2](#)).

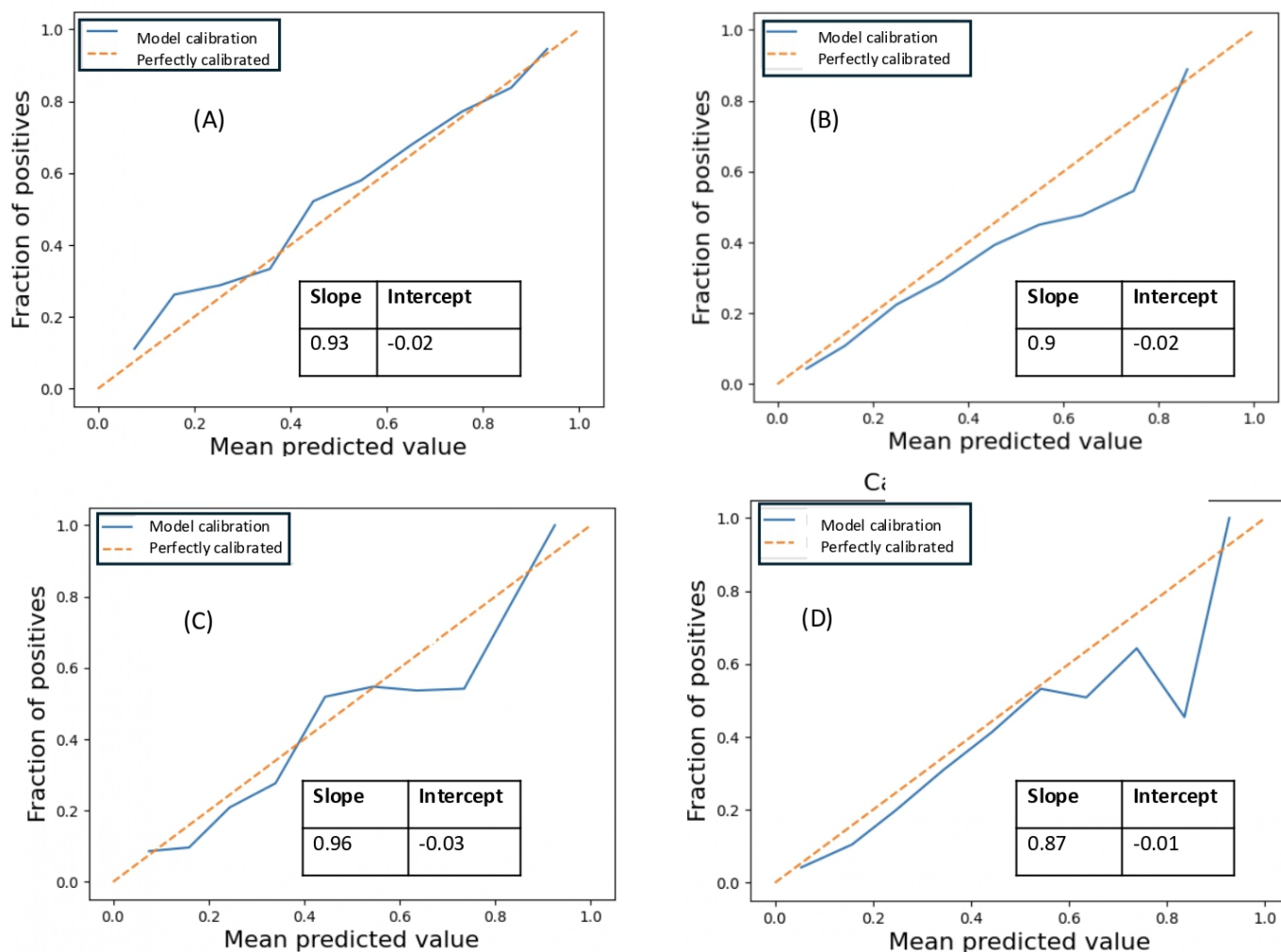
Feature importance analysis using SHAP values revealed consistent patterns across all outcomes. Age at transplantation and donor type emerged as the most influential predictors, corroborating previous findings by Brazauskas et al [16] and Eapen et al [5]. Disease severity indicators, most importantly the frequency of acute chest syndrome episodes in the 2 years preceding HCT, were identified as another critical

predictor. The frequency of VOC requiring hospitalization and the need for exchange transfusions also demonstrated substantial predictive importance. This suggests the utility of including pretransplant disease characteristics for predicting outcomes. The comprehensive SHAP analysis, including feature importance rankings and their relative contributions to model predictions, is presented in [Multimedia Appendices 3–7](#).

To demonstrate the clinical utility of our model, we analyzed predictions across 3 distinct hypothetical patient scenarios. Case-specific patient characteristics and their corresponding predicted outcomes are detailed in [Table 5](#) [Table S3](#) in [Multimedia Appendix 8](#) respectively. The model’s predictions aligned with established clinical observations, showing less favorable outcomes in cases involving non-HLA identical donors and in older patients with more severe disease characteristics, which were consistent with previous studies [5,16].



**Figure 2.** Calibration of SPRIGHT for various outcomes. Calibration curve, slope, and intercept for (A). Event-free survival, (B) graft failure, (C) acute graft-versus-host disease (GVHD), and (D) chronic GVHD. Overall survival (OS) shows a nonideal calibration slope of 0.75 and intercept of  $-0.07$  (data not shown) and may be reflective of the overall high OS in all subgroups.



**Table 5.** Hypothetical patient profiles with varying age, donor type, and disease severity.

Data	Patient 1	Patient 2	Patient 3
<b>Patient data</b>			
Age at transplant (year)	6	6	16
Age group (year)	$\leq 10$	$\leq 10$	11-17
Sex	Male	Male	Male
KPS or Lansky <sup>a</sup> score	$< 90$	$< 90$	$< 90$
HCT <sup>b</sup> -Comorbidity index (in range)	0-2	0-2	0-2
Recipient CMV <sup>c</sup> serostatus	Negative	Negative	Negative
<b>Transplant data</b>			
Donor	HLA <sup>d</sup> identical sibling	HLA mismatch relative	HLA identical sibling
Graft type	Bone marrow	Bone marrow	Bone marrow
Donor-recipient HLA matching	8/8	7/8	8/8
Conditioning intensity	Myeloablative	Nonmyeloablative	Myeloablative
Conditioning regimen	Flu/Bu <sup>e</sup>	TBI/Cy/Flu/TT <sup>f</sup>	Flu/Bu
Serotherapy	ATG <sup>g</sup>	ATG	ATG
GVHD <sup>h</sup> prophylaxis	CNI+MTX <sup>i</sup>	Post-Cy+Siro $\pm$ MMF <sup>j</sup>	CNI+MTX
<b>SCD<sup>k</sup> data</b>			

Data	Patient 1	Patient 2	Patient 3
Number of ACS <sup>1</sup> syndromes within 2 year pre-HCT	0	0	2
Require exchange transfusion	No	No	No
VOC <sup>m</sup> requiring hospitalization within 2 year pre-HCT	No	No	Yes
Frequency of hospitalizations for VOC	<3 per year	<3 per year	<3 per year

<sup>a</sup>KPS: Karnofsky performance status.

<sup>b</sup>HCT: hematopoietic cell transplantation.

<sup>c</sup>CMV: cytomegalovirus.

<sup>d</sup>HLA: human leukocyte antigen.

<sup>e</sup>Flu/Bu: fludarabine + busulfan.

<sup>f</sup>TBI/Cy/Flu/TT: total body irradiation + cyclophosphamide + fludarabine + thiotepa.

<sup>g</sup>ATG: anti-thymocyte globulin.

<sup>h</sup>GVHD: graft-versus-host disease.

<sup>i</sup>CNI+MTX: calcineurin inhibitor + methotrexate.

<sup>j</sup>Post-Cy+Siro± MMF: post-transplant cyclophosphamide + sirolimus + mycophenolate mofetil.

<sup>k</sup>SCD: sickle cell disease.

<sup>l</sup>ACS: acute chest syndrome.

<sup>m</sup>VOC: vaso-occlusive crisis.

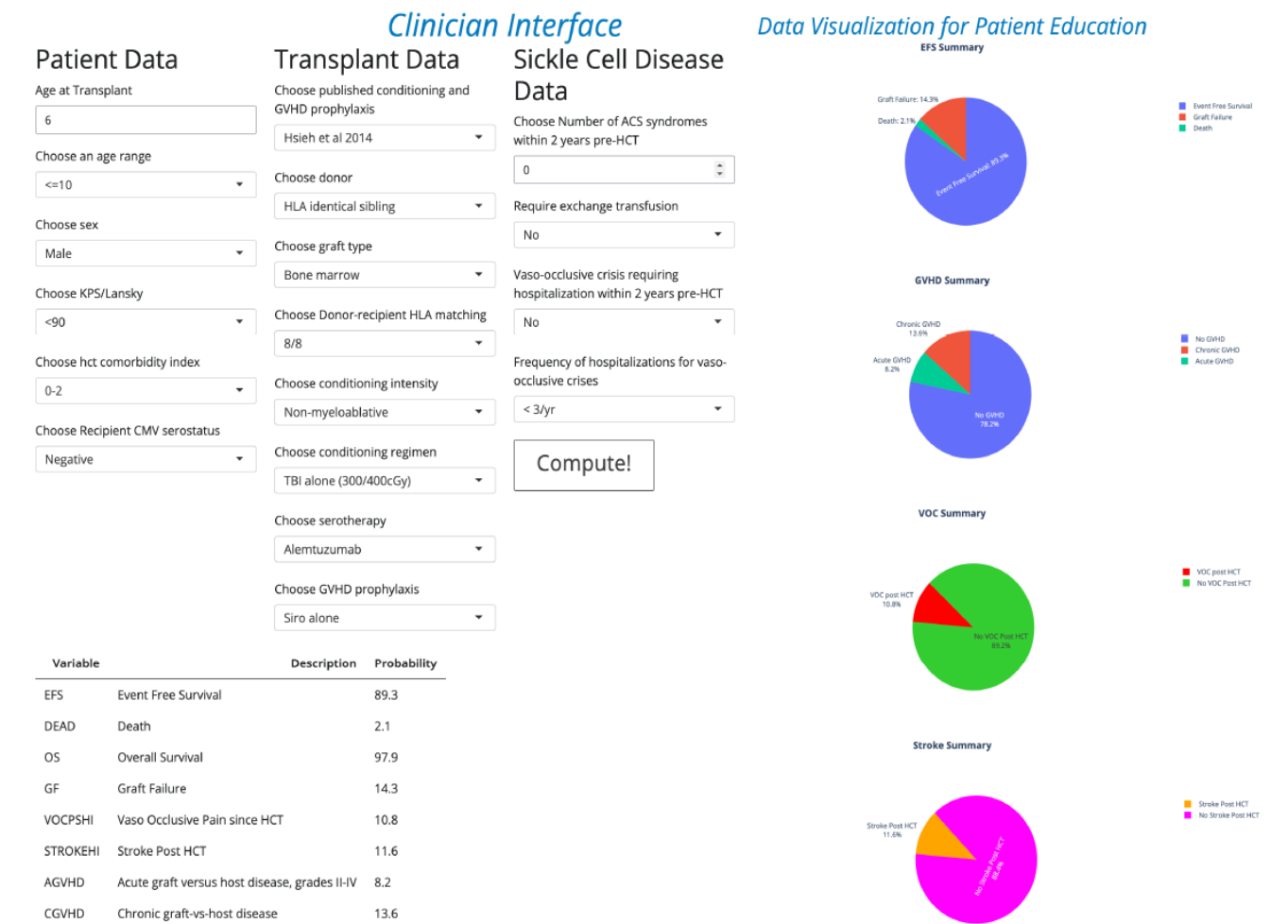
## SPRIGHT User Interface

Age is the only numeric feature that is entered manually. The rest of the inputs are categorical and are entered by selecting an option from a drop-down menu. Patient-specific, disease-specific data, and treatment donor, one of the major published treatment regimens or a customized conditioning regimen, conditioning intensity, and GVHD prophylaxis ATG/Alemtuzumab data can be entered. The tabular output describes predicted OS, EFS, GF, Death, AGVHD, CGVHD, VOC, and stroke. The predicted outcomes are also pictorially represented in pie charts. One unique feature is that the user has the option of selecting a published HCT regimen

or selecting a custom regimen by combining conditioning, GVHD prophylaxis, and serotherapy (Figure 3; Table S4 in Multimedia Appendix 9). This feature decreases keystrokes, improves ease of use of the app, and facilitates a comparative analysis across donor types and treatment regimens for different donor types. SPRIGHT can be accessed on any smartphone, tablet, or personal computer using a shortened URL or a QR code. The individualized outcomes are also represented as pie charts displaying individualized estimates (Figure 3). The pie charts can be downloaded and shared with the patient or added to the electronic medical record as an image.



**Figure 3.** Clinician user interface of sickle cell predicting outcomes of hematopoietic cell transplantation



## Limitations

There are several limitations to this study. The SPRIGHT predictive tool is based on the HCT for the SCD dataset derived from data reported to CIBMTR. The use of the CIBMTR dataset allows us to access the entire US experience reported, but it is also subject to limitations of registry studies, including bias, loss to follow-up, and a lack of generalizability across era, center volume, and expertise. While SPRIGHT uses multiple predictive variables, it is limited to those variables collected by CIBMTR. In 2019, Bolanos-Meade et al [44] reported an improved EFS following mismatched relative donor HCT as compared to their previous report in 2009 following the escalation of the dose of total body irradiation from 200 to 400 cGy [6,45]. However, CIBMTR form 2400 and consequently the HCT for the SCD dataset do not distinguish between patients who received the 2 dose levels of total body irradiation. Further, since pain crisis post-HCT is included as a discrete variable in the HCT for the SCD dataset, it is not possible to discern its timing, frequency, or severity. In the CIBMTR dataset, an analysis of predictor completeness reveals that 10/13 TED variables chosen have a completion rate exceeding 98%. However, only approximately 46.5% (763/1641) of patients are on the CRF track, with comprehensive disease-specific data available, a category under which 4 of our predictor variables fall. While the missing data could be a source of bias, the missingness of data was only a function of whether the institution was designated as a TED-only or CRF and whether the CIBMTR algorithm assigned an individual patient to the CRF track. Thus, the missing data may be missing at random. The model we used for imputation, MissForest, has been shown to outperform all other algorithms in all metrics [46,47]. However, these imputation algorithms can produce severely biased regression coefficients and require a careful critique of the missing data mechanism and the interrelationships between the variables in the data [48]. Overall, we acknowledge the concerns regarding the lack of details of SCD-related complications and the completeness of reporting of SCD-related clinical outcomes in the CIBMTR dataset. We, however, also recognize that the CIBMTR registry, with federally mandated data submission, contains the most complete data available. We support ongoing efforts to refine the data collection measures and training of data collection staff. One of the limitations of our model is the rarity of death events in the dataset, which limits the precision of calibration for OS. As a result, the model tends to slightly overestimate OS risk, particularly in subgroups with fewer events. This calibration limitation should be considered when interpreting OS predictions. Across all outcomes, despite mitigatory efforts to combat the inherent bias due to class imbalance, the bias may not be fully eradicated.

The tool incorporates a set of treatment regimens that have been carefully selected from peer-reviewed studies, ensuring that they are backed by sufficient clinical data. These regimens provide reliable and evidence-based predictions. However, custom combinations that are entered by users may fall outside of the dataset's training and may not be

fully supported by the underlying data. In such cases, it is essential to rely on clinical expertise and user discretion when interpreting the predictions.

We recognize a limitation in the registry data due to inherent selection bias, with limited insight into the clinical reasoning behind regimen choices. While the data detail the regimens administered, the underlying clinical judgment is often not recorded. This limitation is particularly notable in HCT, where the absence of universally established standards of care presents additional challenges. Thus, the model and its results should be interpreted within this context.

## Comparison With Prior Work

Brazauskas et al [16] developed a risk score using age and donor type as discrete variables. They reported that patients aged  $\leq 12$  years with an HLA-matched sibling donor were at the lowest risk. Patients aged  $\geq 13$  years with an HLA-matched sibling donor or aged  $\leq 12$  years with an HLA-matched unrelated donor were at intermediate risk. All other groups were at high risk. This simple risk score has good predictive performance but has certain gaps that limit its utility in the individualized prediction of outcomes. Gluckman et al [6] used age as a continuous variable and observed that for every 1-year increment in age, there was a 9% increase in the HR for treatment failure (graft failure or death) and a 10% increase in the HR for death. Thus, the Brazauskas model does not include the potential predictive value of increasing age from 5 to 13 years. Further, Brazauskas et al [16] do not incorporate other patient, disease, and conditioning regimen characteristics in the prediction model because they considered these factors to be dynamic and subject to change. The Brazauskas model limits the predicted post-HCT outcomes to death, OS, and EFS and does not include other outcomes that are important to physicians and patients in shared decision-making, such as AGVHD, CGVHD, recurrent pain crisis, or stroke after HCT [13,49-51]. SPRIGHT predicts these outcomes with high predictive performance. Other published ML predictive models for HCT for other diseases are of limited clinical relevance in decision-making since they limit themselves to predicting single outcomes, such as death, overall survival, disease relapse, GVHD, busulfan exposure, kidney injury, or reactivation of Epstein-Barr virus [17-30].

In developing SPRIGHT, we addressed the gaps in the knowledge in individualized prediction of outcomes of HCT. We included 7 clinically relevant outcomes, including rates of OS, EFS, GF, death, AGVHD, CGVHD, and VOC. Each of these 7 clinical outcomes initially required a distinct set of 10-11 pre-HCT predictive features for optimal performance, leading to incomplete overlap and potential model fragmentation. For addressing this, we adopted a unified approach by selecting a comprehensive set of 17 pre-HCT patient, HCT, and disease characteristics. By applying robust feature selection techniques to optimize predictive performance and improve the model's overall clinical applicability, we demonstrated the predictive value of these features. Further, patient and disease features inform patient selection and HCT features inform regimen selection. Thus, these

additional pre-HCT features are important, clinically relevant considerations in decision-making.

## Future Directions

We present here an initial in-lab validation of the predictive model. We recognize that the use of a US-based dataset may limit the generalizability of our findings to other regions or health care systems due to potential differences in population characteristics and treatment protocols. For further external validation across different geographic locations to establish the model's generalizability and clinical utility, we propose to use the European Bone Marrow Transplantation Registry (EBMT) dataset, a completely independent dataset, in collaboration with European investigators. Of note, Gluckman et al [6] have previously combined CIBMTR and EBMT registry data on SCD for analysis and included similar numbers of children, adults, and donor types from the 2 registries and do not report differences in predictive factors or outcomes in the 2 registries. We have demonstrated that

the predictive performance remains equivalent whether we use the entire dataset or the more recent data after 2007, which represents 84% (1378/1641) of participants in the dataset. Acknowledging the evolving nature of supportive care practices in HCT, we propose further temporal validation with future years of data being added to this dataset. We also propose to continue to enhance SPRIGHT by incorporating expert opinion, adapting to patient health literacy, values, and preferences, and using patient-friendly data visualization to support shared decision-making [52-54].

## Conclusions

In conclusion, the SPRIGHT prediction model integrates individual-specific patient and disease characteristics, conditioning regimens, GVHD prophylaxis, and donor characteristics and predicts key clinical outcomes. It exhibits superior predictive performance across multiple measures of discrimination and calibration as compared to logistic regression and other ensemble ML methods.

## Data Availability

The National Institutes of Health (NIH) Hematopoietic Cell Transplantation for Sickle Cell Disease dataset used to develop the clinical decision tool was obtained from the NHLBI Biologic Specimen and Data Repository Information Coordinating Center (BIOLINCC) [31]. For additional questions on the algorithms used for the decision tool, investigators may contact the corresponding author by email.

## Conflicts of Interest

None declared.

## Multimedia Appendix 1

Feature distribution of the dataset.

[\[DOCX File \(Microsoft Word File\), 21 KB-Multimedia Appendix 1\]](#)

## Multimedia Appendix 2

Outcome distribution.

[\[DOCX File \(Microsoft Word File\), 15 KB-Multimedia Appendix 2\]](#)

## Multimedia Appendix 3

Mean absolute Shapley additive explanations values quantifying predictor importance for acute graft-vs-host disease. Values averaged across cross-validation folds and bootstraps.

[\[PDF File \(Adobe File\), 77 KB-Multimedia Appendix 3\]](#)

## Multimedia Appendix 4

Mean absolute Shapley additive explanations values quantifying predictor importance for graft failure. Values averaged across cross-validation folds and bootstraps.

[\[PDF File \(Adobe File\), 169 KB-Multimedia Appendix 4\]](#)

## Multimedia Appendix 5

Mean absolute Shapley additive explanations values quantifying predictor importance for Event Free Survival. Values averaged across cross-validation folds and bootstraps.

[\[PDF File \(Adobe File\), 163 KB-Multimedia Appendix 5\]](#)

## Multimedia Appendix 6

Mean absolute Shapley additive explanations values quantifying predictor importance for CGVHD. Values averaged across cross-validation folds and bootstraps.

[\[PDF File \(Adobe File\), 157 KB-Multimedia Appendix 6\]](#)

## Multimedia Appendix 7

Mean absolute Shapley additive explanations values quantifying predictor importance for overall survival. Values averaged across cross-validation folds and bootstraps.

[\[PDF File \(Adobe File\), 58499 KB-Multimedia Appendix 7\]](#)

### Multimedia Appendix 8

Predicted probability percentage outcomes for each of the hypothetical patient profiles.

[\[DOCX File \(Microsoft Word File\), 13 KB-Multimedia Appendix 8\]](#)

### Multimedia Appendix 9

Combinations of conditioning regimen, serotherapy, and graft-versus-host disease prophylaxis used in published case series, which are included in the SPRIGHT.

[\[DOCX File \(Microsoft Word File\), 14 KB-Multimedia Appendix 9\]](#)

### Checklist 1

Consolidated Reporting of Machine Learning Studies checklist.

[\[PDF File \(Adobe File\), 116 KB-Checklist 1\]](#)

### References

1. Kato GJ, Piel FB, Reid CD, et al. Sickle cell disease. *Nat Rev Dis Primers*. Mar 15, 2018;4(1):18010. [doi: [10.1038/nrdp.2018.10](https://doi.org/10.1038/nrdp.2018.10)] [Medline: [29542687](https://pubmed.ncbi.nlm.nih.gov/29542687/)]
2. Walters MC, De Castro LM, Sullivan KM, et al. Indications and results of HLA-identical sibling hematopoietic cell transplantation for sickle cell disease. *Biol Blood Marrow Transplant*. Feb 2016;22(2):207-211. [doi: [10.1016/j.bbmt.2015.10.017](https://doi.org/10.1016/j.bbmt.2015.10.017)]
3. Walters MC, Patience M, Leisenring W, et al. Bone marrow transplantation for sickle cell disease. *N Engl J Med*. Aug 8, 1996;335(6):369-376. [doi: [10.1056/NEJM199608083350601](https://doi.org/10.1056/NEJM199608083350601)]
4. Walters MC, Patience M, Leisenring W, et al. Collaborative multicenter investigation of marrow transplantation for sickle cell disease: current results and future directions. *Biol Blood Marrow Transplant*. Dec 1997;3(6):310-315. [Medline: [9502298](https://pubmed.ncbi.nlm.nih.gov/9502298/)]
5. Eapen M, Brazauskas R, Walters MC, et al. Effect of donor type and conditioning regimen intensity on allogeneic transplantation outcomes in patients with sickle cell disease: a retrospective multicentre, cohort study. *Lancet Haematol*. Nov 2019;6(11):e585-e596. [doi: [10.1016/S2352-3026\(19\)30154-1](https://doi.org/10.1016/S2352-3026(19)30154-1)]
6. Gluckman E, Cappelli B, Bernaudin F, et al. Sickle cell disease: an international survey of results of HLA-identical sibling hematopoietic stem cell transplantation. *Blood*. Mar 16, 2017;129(11):1548-1556. [doi: [10.1182/blood-2016-10-745711](https://doi.org/10.1182/blood-2016-10-745711)] [Medline: [27965196](https://pubmed.ncbi.nlm.nih.gov/27965196/)]
7. Hsieh MM, Fitzhugh CD, Weitzel RP, et al. Nonmyeloablative HLA-matched sibling allogeneic hematopoietic stem cell transplantation for severe sickle cell phenotype. *JAMA*. Jul 2, 2014;312(1):48-56. [doi: [10.1001/jama.2014.7192](https://doi.org/10.1001/jama.2014.7192)] [Medline: [25058217](https://pubmed.ncbi.nlm.nih.gov/25058217/)]
8. Krishnamurti L, Neuberg DS, Sullivan KM, et al. Bone marrow transplantation for adolescents and young adults with sickle cell disease: results of a prospective multicenter pilot study. *Am J Hematol*. Apr 2019;94(4):446-454. URL: <https://onlinelibrary.wiley.com/toc/10968652/94/4> [doi: [10.1002/ajh.25401](https://doi.org/10.1002/ajh.25401)] [Medline: [30637784](https://pubmed.ncbi.nlm.nih.gov/30637784/)]
9. Shenoy S, Eapen M, Panepinto JA, et al. A trial of unrelated donor marrow transplantation for children with severe sickle cell disease. *Blood*. Nov 24, 2016;128(21):2561-2567. [doi: [10.1182/blood-2016-05-715870](https://doi.org/10.1182/blood-2016-05-715870)] [Medline: [27625358](https://pubmed.ncbi.nlm.nih.gov/27625358/)]
10. King AA, Kamani N, Bunin N, et al. Successful matched sibling donor marrow transplantation following reduced intensity conditioning in children with hemoglobinopathies. *Am J Hematol*. Dec 2015;90(12):1093-1098. [doi: [10.1002/ajh.24183](https://doi.org/10.1002/ajh.24183)] [Medline: [26348869](https://pubmed.ncbi.nlm.nih.gov/26348869/)]
11. Gluckman E, Cappelli B, Scigliuolo GM, Fuente JD la, Corbacioglu S. Alternative donor hematopoietic stem cell transplantation for sickle cell disease in Europe. *Hematol Oncol Stem Cell Ther*. 2020;13(4):181-188. [doi: [10.1016/j.hemonc.2019.12.011](https://doi.org/10.1016/j.hemonc.2019.12.011)]
12. Patel DA, Akinsete AM, Fuente J de la, Kassim AA. Haploidentical bone marrow transplant with posttransplant cyclophosphamide for sickle cell disease: an update. *Hematol Oncol Stem Cell Ther*. 2020;13(2):91-97. [doi: [10.1016/j.hemonc.2020.01.002](https://doi.org/10.1016/j.hemonc.2020.01.002)] [Medline: [32202252](https://pubmed.ncbi.nlm.nih.gov/32202252/)]
13. Sinha CB, Bakshi N, Ross D, Loewenstein G, Krishnamurti L. Primary caregiver decision-making in hematopoietic cell transplantation and gene therapy for sickle cell disease. *Pediatr Blood Cancer*. Jan 2021;68(1):e28749. [doi: [10.1002/pbc.28749](https://doi.org/10.1002/pbc.28749)] [Medline: [33034129](https://pubmed.ncbi.nlm.nih.gov/33034129/)]
14. Bakshi N, Katoch D, Sinha CB, et al. Assessment of patient and caregiver attitudes and approaches to decision-making regarding bone marrow transplant for sickle cell disease: a qualitative study. *JAMA Netw Open*. May 1, 2020;3(5):e206742. [doi: [10.1001/jamanetworkopen.2020.6742](https://doi.org/10.1001/jamanetworkopen.2020.6742)] [Medline: [32469414](https://pubmed.ncbi.nlm.nih.gov/32469414/)]



15. Veludhandi A, Ross D, Sinha CB, McCracken C, Bakshi N, Krishnamurti L. A decision support tool for allogeneic hematopoietic stem cell transplantation for children with sickle cell disease: acceptability and usability study. *JMIR Form Res.* Oct 28, 2021;5(10):e30093. [doi: [10.2196/30093](https://doi.org/10.2196/30093)] [Medline: [34709190](https://pubmed.ncbi.nlm.nih.gov/34709190/)]
16. Brazauskas R, Scigliuolo GM, Wang HL, et al. Risk score to predict event-free survival after hematopoietic cell transplant for sickle cell disease. *Blood.* Jul 30, 2020;136(5):623-626. [doi: [10.1182/blood.2020005687](https://doi.org/10.1182/blood.2020005687)] [Medline: [32518950](https://pubmed.ncbi.nlm.nih.gov/32518950/)]
17. Afanaseva KS, Bakin EA, Smirnova AG, et al. A pilot study of implication of machine learning for relapse prediction after allogeneic stem cell transplantation in adults with Ph-positive acute lymphoblastic leukemia. *Sci Rep.* Oct 5, 2023;13(1):16790. [doi: [10.1038/s41598-023-43950-w](https://doi.org/10.1038/s41598-023-43950-w)] [Medline: [37798335](https://pubmed.ncbi.nlm.nih.gov/37798335/)]
18. Al-Riyami AZ, Maryamchik E, Hanna RS, et al. A machine-learning model that incorporates CD45 surface expression predicts hematopoietic progenitor cell recovery after freeze-thaw. *Cytotherapy.* Oct 2023;25(10):1048-1056. [doi: [10.1016/j.jcyt.2023.05.007](https://doi.org/10.1016/j.jcyt.2023.05.007)] [Medline: [37318396](https://pubmed.ncbi.nlm.nih.gov/37318396/)]
19. Fan S, Hong HY, Dong XY, et al. Machine learning algorithm as a prognostic tool for Epstein-Barr virus reactivation after haploidentical hematopoietic stem cell transplantation. *Blood Sci.* Jan 2023;5(1):51-59. [doi: [10.1097/BS9.000000000000143](https://doi.org/10.1097/BS9.000000000000143)] [Medline: [36742189](https://pubmed.ncbi.nlm.nih.gov/36742189/)]
20. Keret S, Rimar D, Lansiaux P, et al. Differentially expressed genes in systemic sclerosis: towards predictive medicine with new molecular tools for clinicians. *Autoimmun Rev.* Jun 2023;22(6):103314. [doi: [10.1016/j.autrev.2023.103314](https://doi.org/10.1016/j.autrev.2023.103314)] [Medline: [36918090](https://pubmed.ncbi.nlm.nih.gov/36918090/)]
21. Li D, Zhao J, Xu B, et al. Predicting busulfan exposure in patients undergoing hematopoietic stem cell transplantation using machine learning techniques. *Expert Rev Clin Pharmacol.* 2023;16(8):751-761. [doi: [10.1080/17512433.2023.2226866](https://doi.org/10.1080/17512433.2023.2226866)] [Medline: [37326641](https://pubmed.ncbi.nlm.nih.gov/37326641/)]
22. Mushtaq AH, Shafqat A, Salah HT, Hashmi SK, Muhsen IN. Machine learning applications and challenges in graft-versus-host disease: a scoping review. *Curr Opin Oncol.* Nov 1, 2023;35(6):594-600. [doi: [10.1097/CCO.0000000000000996](https://doi.org/10.1097/CCO.0000000000000996)] [Medline: [37820094](https://pubmed.ncbi.nlm.nih.gov/37820094/)]
23. Shourabizadeh H, Aleman DM, Rousseau LM, Law AD, Viswabandya A, Michelis FV. Machine learning for the prediction of survival post-allogeneic hematopoietic cell transplantation: a single-center experience. *Acta Haematol.* May 29, 2024;147(3):280-291. [doi: [10.1159/000533665](https://doi.org/10.1159/000533665)] [Medline: [37769635](https://pubmed.ncbi.nlm.nih.gov/37769635/)]
24. Sobrino S, Magnani A, Semeraro M, et al. Severe hematopoietic stem cell inflammation compromises chronic granulomatous disease gene therapy. *Cell Rep Med.* Feb 21, 2023;4(2):100919. [doi: [10.1016/j.xcrm.2023.100919](https://doi.org/10.1016/j.xcrm.2023.100919)] [Medline: [36706754](https://pubmed.ncbi.nlm.nih.gov/36706754/)]
25. Sorror ML. The use of prognostic models in allogeneic transplants: a perspective guide for clinicians and investigators. *Blood.* May 4, 2023;141(18):2173-2186. [doi: [10.1182/blood.2022017999](https://doi.org/10.1182/blood.2022017999)] [Medline: [36800564](https://pubmed.ncbi.nlm.nih.gov/36800564/)]
26. Sparapani RA, Logan BR, Maiers MJ, Laud PW, McCulloch RE. Nonparametric failure time: Time-to-event machine learning with heteroskedastic Bayesian additive regression trees and low information omnibus Dirichlet process mixtures. *Biometrics.* Dec 2023;79(4):3023-3037. [doi: [10.1111/biom.13857](https://doi.org/10.1111/biom.13857)] [Medline: [36932826](https://pubmed.ncbi.nlm.nih.gov/36932826/)]
27. von Asmuth EGJ, Neven B, Albert MH, et al. Predicting Patient Death after Allogeneic Stem Cell Transplantation for Inborn Errors Using Machine Learning (PREPAD): A European Society for Blood and Marrow Transplantation Inborn Errors Working Party Study. *Transplant Cell Ther.* Dec 2023;29(12):775. [doi: [10.1016/j.jtct.2023.09.007](https://doi.org/10.1016/j.jtct.2023.09.007)] [Medline: [37709203](https://pubmed.ncbi.nlm.nih.gov/37709203/)]
28. Wang P, Liu C, Wei Z, et al. Nomogram for predicting early mortality after umbilical cord blood transplantation in children with inborn errors of immunity. *J Clin Immunol.* Aug 2023;43(6):1379-1392. [doi: [10.1007/s10875-023-01505-8](https://doi.org/10.1007/s10875-023-01505-8)] [Medline: [37155023](https://pubmed.ncbi.nlm.nih.gov/37155023/)]
29. Zhou Y, Smith J, Keerthi D, et al. Longitudinal clinical data improve survival prediction after hematopoietic cell transplantation using machine learning. *Blood Adv.* Feb 13, 2024;8(3):686-698. [doi: [10.1182/bloodadvances.2023011752](https://doi.org/10.1182/bloodadvances.2023011752)] [Medline: [37991991](https://pubmed.ncbi.nlm.nih.gov/37991991/)]
30. Krishnamurti L, Liang J, He Z, et al. Incidence and risk factors of pain crisis after hematopoietic cell transplantation for sickle cell disease. *Blood Adv.* Apr 23, 2024;8(8):1908-1919. [doi: [10.1182/bloodadvances.2023010749](https://doi.org/10.1182/bloodadvances.2023010749)] [Medline: [38324722](https://pubmed.ncbi.nlm.nih.gov/38324722/)]
31. Friedman D, Dozor AJ, Milner J, et al. Stable to improved cardiac and pulmonary function in children with high-risk sickle cell disease following haploidentical stem cell transplantation. *Bone Marrow Transplant.* Sep 2021;56(9):2221-2230. [doi: [10.1038/s41409-021-01298-7](https://doi.org/10.1038/s41409-021-01298-7)] [Medline: [33958740](https://pubmed.ncbi.nlm.nih.gov/33958740/)]
32. Giffen CA, Carroll LE, Adams JT, Brennan SP, Coady SA, Wagner EL. Providing contemporary access to historical biospecimen collections: Development of the NHLBI Biologic Specimen and Data Repository Information Coordinating Center (BioLINCC). *Biopreserv Biobank.* Aug 2015;13(4):271-279. [doi: [10.1089/bio.2014.0050](https://doi.org/10.1089/bio.2014.0050)] [Medline: [26186276](https://pubmed.ncbi.nlm.nih.gov/26186276/)]

33. Hong F, Tian L, Devanarayan V. Improving the robustness of variable selection and predictive performance of regularized generalized linear models and Cox proportional hazard models. *Mathematics*. Feb 2023;11(3):557. [doi: [10.3390/math11030557](https://doi.org/10.3390/math11030557)] [Medline: [37990696](https://pubmed.ncbi.nlm.nih.gov/37990696/)]
34. Jain R, Xu W. Artificial Intelligence based wrapper for high dimensional feature selection. *BMC Bioinformatics*. Oct 18, 2023;24(1):392. [doi: [10.1186/s12859-023-05502-x](https://doi.org/10.1186/s12859-023-05502-x)] [Medline: [37853338](https://pubmed.ncbi.nlm.nih.gov/37853338/)]
35. Ruisen L, Songyi D, Chen W, et al. Bagging of XGBoost classifiers with random under-sampling and Tomek link for noisy label-imbalanced data. *IOP Conf Ser: Mater Sci Eng.* ;428(1):012004. [doi: [10.1088/1757-899X/428/1/012004](https://doi.org/10.1088/1757-899X/428/1/012004)]
36. Van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW, Topic Group 'Evaluating diagnostic tests and prediction models' of the STRATOS initiative. Calibration: the Achilles heel of predictive analytics. *BMC Med*. Dec 16, 2019;17(1):230. [doi: [10.1186/s12916-019-1466-7](https://doi.org/10.1186/s12916-019-1466-7)] [Medline: [31842878](https://pubmed.ncbi.nlm.nih.gov/31842878/)]
37. van der Laan L, Ulloa-Pérez E, Carone M, Luedtke A. Causal isotonic calibration for heterogeneous treatment effects. *Proc Mach Learn Res*. Jul 2023;202:34831-34854. [Medline: [37575467](https://pubmed.ncbi.nlm.nih.gov/37575467/)]
38. Pozzolo AD, Caelen O, Johnson RA, Bontempi G. Calibrating probability with undersampling for unbalanced classification. Presented at: 2015 IEEE Symposium Series on Computational Intelligence (SSCI); Dec 7-10, 2015; Cape Town, South Africa. [doi: [10.1109/SSCI.2015.33](https://doi.org/10.1109/SSCI.2015.33)]
39. El Emam K, Leung TI, Malin B, Klement W, Eysenbach G. Consolidated Reporting Guidelines for Prognostic and Diagnostic Machine Learning Models (CREMLS). *J Med Internet Res*. May 2, 2024;26:e52508. [doi: [10.2196/52508](https://doi.org/10.2196/52508)] [Medline: [38696776](https://pubmed.ncbi.nlm.nih.gov/38696776/)]
40. Bouckaert RR, Frank E. *Evaluating the Replicability of Significance Tests for Comparing Learning Algorithms*. Berlin, Heidelberg: Springer; 2004.
41. Nadeau C, Bengio Y. Inference for the generalization error. *Mach Learn*. Sep 2003;52(3):239-281. [doi: [10.1023/A:1024068626366](https://doi.org/10.1023/A:1024068626366)]
42. Taheriyani M, Safaei Nodehi S, Niakan Kalhori SR, Mohammadzadeh N. A systematic review of the predicted outcomes related to hematopoietic stem cell transplantation: focus on applied machine learning methods' performance. *Expert Rev Hematol*. Feb 1, 2022;15(2):137-156. [doi: [10.1080/17474086.2022.2042248](https://doi.org/10.1080/17474086.2022.2042248)]
43. Cappelli B, Volt F, Tozatto-Maio K, et al. Risk factors and outcomes according to age at transplantation with an HLA-identical sibling for sickle cell disease. *Haematologica*. Dec 2019;104(12):e543-e546. [doi: [10.3324/haematol.2019.216788](https://doi.org/10.3324/haematol.2019.216788)] [Medline: [31018975](https://pubmed.ncbi.nlm.nih.gov/31018975/)]
44. Bolaños-Meade J, Cooke KR, Gamper CJ, et al. Effect of increased dose of total body irradiation on graft failure associated with HLA-haploidentical transplantation in patients with severe haemoglobinopathies: a prospective clinical trial. *Lancet Haematol*. Apr 2019;6(4):e183-e193. [doi: [10.1016/S2352-3026\(19\)30031-6](https://doi.org/10.1016/S2352-3026(19)30031-6)]
45. Bolaños-Meade J, Brodsky RA. Blood and marrow transplantation for sickle cell disease: is less more? *Blood Rev*. Nov 2014;28(6):243-248. [doi: [10.1016/j.blre.2014.08.001](https://doi.org/10.1016/j.blre.2014.08.001)] [Medline: [25217413](https://pubmed.ncbi.nlm.nih.gov/25217413/)]
46. Buczak P, Chen JJ, Pauly M. Analyzing the effect of imputation on classification performance under MCAR and MAR missing mechanisms. *Entropy (Basel)*. Mar 17, 2023;25(3):521. [doi: [10.3390/e25030521](https://doi.org/10.3390/e25030521)] [Medline: [36981409](https://pubmed.ncbi.nlm.nih.gov/36981409/)]
47. Tang F, Ishwaran H. Random forest missing data algorithms. *Stat Anal Data Min*. Dec 2017;10(6):363-377. URL: <https://onlinelibrary.wiley.com/toc/19321872/10/6> [doi: [10.1002/sam.11348](https://doi.org/10.1002/sam.11348)] [Medline: [29403567](https://pubmed.ncbi.nlm.nih.gov/29403567/)]
48. Hong S, Lynn HS. Accuracy of random-forest-based imputation of missing data in the presence of non-normality, non-linearity, and interaction. *BMC Med Res Methodol*. Jul 25, 2020;20(1):199. [doi: [10.1186/s12874-020-01080-1](https://doi.org/10.1186/s12874-020-01080-1)] [Medline: [32711455](https://pubmed.ncbi.nlm.nih.gov/32711455/)]
49. Gluckman E, Fuente J de la, Cappelli B, et al. The role of HLA matching in unrelated donor hematopoietic stem cell transplantation for sickle cell disease in Europe. *Bone Marrow Transplant*. Oct 2020;55(10):1946-1954. [doi: [10.1038/s41409-020-0847-z](https://doi.org/10.1038/s41409-020-0847-z)] [Medline: [32157246](https://pubmed.ncbi.nlm.nih.gov/32157246/)]
50. Meier ER, Dioguardi JV, Kamani N. Current attitudes of parents and patients toward hematopoietic stem cell transplantation for sickle cell anemia. *Pediatr Blood Cancer*. Jul 2015;62(7):1277-1284. [doi: [10.1002/pbc.25446](https://doi.org/10.1002/pbc.25446)] [Medline: [25809231](https://pubmed.ncbi.nlm.nih.gov/25809231/)]
51. Sinha CB, Meacham LR, Bakshi N, Ross D, Krishnamurti L. Parental perspective on the risk of infertility and fertility preservation options for children and adolescents with sickle cell disease considering hematopoietic stem cell transplantation. *Pediatr Blood Cancer*. Jul 2023;70(7):e30276. [doi: [10.1002/pbc.30276](https://doi.org/10.1002/pbc.30276)] [Medline: [37051746](https://pubmed.ncbi.nlm.nih.gov/37051746/)]
52. Greenes RA, Bates DW, Kawamoto K, Middleton B, Osheroff J, Shahar Y. Clinical decision support models and frameworks: seeking to address research issues underlying implementation successes and failures. *J Biomed Inform*. Feb 2018;78(134-43):134-143. [doi: [10.1016/j.jbi.2017.12.005](https://doi.org/10.1016/j.jbi.2017.12.005)] [Medline: [29246790](https://pubmed.ncbi.nlm.nih.gov/29246790/)]
53. Marcial LH, Richardson JE, Lasater B, et al. The imperative for patient-centered clinical decision support. *EGEMS (Wash DC)*. May 30, 2018;6(1):12. [doi: [10.5334/egems.259](https://doi.org/10.5334/egems.259)] [Medline: [30094284](https://pubmed.ncbi.nlm.nih.gov/30094284/)]



54. Krishnamurti L, Ross D, Sinha C, et al. Comparative effectiveness of a web-based patient decision aid for therapeutic options for sickle cell disease: randomized controlled trial. J Med Internet Res. Dec 4, 2019;21(12):e14462. [doi: [10.2196/14462](https://doi.org/10.2196/14462)] [Medline: [31799940](https://pubmed.ncbi.nlm.nih.gov/31799940/)]

## Abbreviations

**AGVHD:** acute graft-versus-host disease  
**AUC:** area under the curve  
**BFE:** backward feature elimination  
**BIOLINCC:** Biologic Specimen and Data Repository Information Coordinating Center  
**CGVHD:** chronic graft-versus-host disease  
**CIBMTR:** Center for International Bone Marrow Transplant Research  
**CREMLS:** Consolidated Reporting of Machine Learning Studies  
**CRF:** Comprehensive Report Form  
**EBMT:** European Bone Marrow Transplantation Registry  
**EFS:** event-free survival  
**FSS:** forward sequential selection  
**GF:** graft-free survival  
**GVHD:** graft-versus-host disease  
**HCT:** hematopoietic cell transplantation  
**HR:** hazard ratio  
**ML:** machine learning  
**NHLBI :** National Heart, Lung, and Blood Institute  
**OS:** overall survival  
**SCD:** sickle cell disease  
**SHAP :** Shapley additive explanations  
**SPRIGHT:** Sickle Cell Predicting Outcomes of Hematopoietic Cell Transplantation  
**VOC:** vasocclusive crisis

*Edited by Khaled El Emam; peer-reviewed by Deepak Singh, Takuto Takahashi; submitted 21.07.2024; final revised version received 20.04.2025; accepted 09.06.2025; published 15.09.2025*

### *Please cite as:*

Chandrasekar RS, Kane M, Krishnamurti L

Machine-Learning Predictive Tool for the Individualized Prediction of Outcomes of Hematopoietic Cell Transplantation for Sickle Cell Disease: Registry-Based Study

JMIR AI 2025;4:e64519

URL: <https://ai.jmir.org/2025/1/e64519>

doi: [10.2196/64519](https://doi.org/10.2196/64519)

© Rajagopal Subramaniam Chandrasekar, Michael Kane, Lakshmanan Krishnamurti. Originally published in JMIR AI (<https://ai.jmir.org>), 15.09.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR AI, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.ai.jmir.org/>, as well as this copyright and license information must be included.