

Using T-lymphocyte subsets at engraftment to predict the risk of acute graft-versus-host disease in patients with thalassemia major: development of a new predictive nomogram

Hongwen Xiao*¹, Gaohui Yang*, Qiulin Huang, Zhenbin Wei, Zhaoping Gan, Meiqing Wu, Zeyan Shi, Huicheng Huang, Zhaofang Pan, Lianjin Liu, Lingling Shi, Zhongming Zhang, Rongrong Liu and Yongrong Lai²

Abstract

Background: Acute graft-versus-host disease (aGvHD) is the primary cause of mortality following allogeneic hematopoietic cell transplantation (HCT).

Objectives: This study aimed to predict the risk of aGvHD after HCT in patients with thalassemia major using a novel predictive nomogram.

Design: A retrospective study was used to develop the prediction model.

Methods: We performed retrospective analyses on 402 consecutive thalassemia patients who underwent HCT. Risk factors for aGvHD were analyzed using Cox proportional regression models. T-lymphocyte subsets were collected from 240 patients at the time of neutrophil engraftment. Least Absolute Shrinkage and Selection Operator regression was utilized to screen the indices, with cut-off values established through restricted cubic spline (RCS) regression. The predictive model was developed by integrating these T-lymphocyte subsets with clinical features, aiming to enhance the accuracy of aGvHD risk prediction.

Results: Among 402 thalassemia patients analyzed post-transplantation, significant independent risk factors for aGvHD included matched unrelated donors, haploid-related donors, peripheral blood stem cell infusions, and donor age older than 40 years. Our RCS analysis indicated a marked increase in aGvHD risk when CD4+ T-cell counts exceeded 36 cells/ μ L and CD8+ T-cell counts exceeded 43 cells/ μ L during neutrophil engraftment. The integration of T-lymphocyte subsets with clinical risk factors into a Cox regression model demonstrated good predictive performance for assessing aGvHD risk.

Conclusion: This study presents a novel model designed to predict aGvHD in thalassemia patients post-transplantation by utilizing T-lymphocyte data at the time of engraftment. The model facilitates the creation of personalized treatment plans, aiming to minimize the incidence of aGvHD and improve patient outcomes.

Keywords: acute graft-versus-host disease, nomogram, risk factor, T lymphocyte, thalassemia major

Received: 28 May 2024; revised manuscript accepted: 7 October 2024.

Introduction

Thalassemia is one of the most prevalent monogenic disorders worldwide. In regions with limited resources, fewer than 15% of individuals with transfusion-dependent thalassemia (TDT) live

beyond the age of 35, significantly adding to the global disease burden.^{1,2} Gene editing and allogeneic hematopoietic cell transplantation (HCT) provide curative options for thalassemia, yet the high costs of gene editing limit its accessibility.^{1,2}

Ther Adv Hematol

2024, Vol. 15: 1–17

DOI: 10.1177/
20406207241294054

© The Author(s), 2024.
Article reuse guidelines:
[sagepub.com/journals-](https://sagepub.com/journals-permissions)
permissions

Correspondence to:

Yongrong Lai
Rongrong Liu
Department of
Hematology, The First
Affiliated Hospital
of Guangxi Medical
University, No. 6
Shuangyong Road,
Nanning, Guangxi 530021,
China

NHC Key Laboratory of
Thalassemia Medicine,
Nanning, Guangxi, China

Guangxi Key Laboratory of
Thalassemia Research,
Nanning, Guangxi, China
laiyongrong@hotmail.com
rong2liu@hotmail.com

Hongwen Xiao
Gaohui Yang
Qiulin Huang
Zhenbin Wei
Zhaoping Gan
Meiqing Wu
Zeyan Shi
Huicheng Huang
Zhaofang Pan
Lianjin Liu
Lingling Shi
Zhongming Zhang
Department of
Hematology, The First
Affiliated Hospital
of Guangxi Medical
University, Nanning,
Guangxi, China

*These authors
contributed equally

Currently, the overall survival and thalassemia-free survival (TFS) rates post-transplantation exceed 90% for TDT patients, with the efficacy of alternative donor transplants approaching that of matched sibling transplants in specialized centers.³⁻⁵

Transplant-related complications are the primary cause of mortality in thalassemia post-transplantation, with acute graft-versus-host disease (aGvHD) being particularly significant. The incidence of aGvHD in thalassemia patients who receive HCT from matched sibling donors is relatively low, with a total incidence reported at 15.3% and grade 2-4 aGvHD at 13.1%.⁴ In contrast, rates of aGvHD from alternative donors remain higher, ranging from 21.5% to 42%.^{3,6,7} The pathogenesis of aGvHD involves a complex immune response, initiated by T-cell receptors recognizing host major histocompatibility complex and minor histocompatibility antigens, triggering cytokine storms that lead to sustained immune activation. Both CD4+ T cell and CD8+ T cell play crucial roles in this process: CD4+ T cell primarily regulate the immune response through cytokine secretion, while CD8+ T cell directly exert cytotoxic effects against recipient cells.^{8,9} These actions result in inflammation and injury to host tissues, mainly affecting the skin, liver, and digestive tract. Clinically, the severity of aGvHD is categorized by the involvement of these organ systems,⁸ and the onset of aGvHD often predicts a poorer prognosis for transplant recipients due to significant tissue and organ damage and increased susceptibility to secondary infections.^{10,11}

Preventing and treating aGvHD is vital in transplantation, yet current prediction models based on clinical features often provide suboptimal results and require additional validation.¹²⁻¹⁶ While proteomic and gut microbiome-based models show potential, their integration into clinical practice remains limited.¹⁷⁻¹⁹ Notably, variations in gut microbiota at the time of engraftment can influence aGvHD outcomes by modulating T-lymphocyte populations,^{17,18} and the CD4+ T-cell count at day 28 post-transplantation has been identified as a potential marker for aGvHD.²⁰ Few studies have examined the correlation between lymphocyte subsets at engraftment and aGvHD. This gap highlights the need for innovative approaches that incorporate T-lymphocyte data at engraftment with clinical features to

improve aGvHD prediction models. This study focuses on aGvHD risk factors in thalassemia major (TM) patients undergoing HCT and assesses T lymphocytes at engraftment as potential biomarkers. By integrating these identified risk factors, we aim to develop a refined aGvHD prediction model that facilitates personalized intervention strategies, enhancing the precision of patient management.

Methods

Source of data

This retrospective cohort study adhered to the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) statement.²¹ The study focused on patients with thalassemia who underwent HCT at the Department of Hematology, The First Affiliated Hospital of Guangxi Medical University, with data retrieved from the Hospital's Electronic System. This article has been approved by the ethics committee of The First Affiliated Hospital of Guangxi Medical University. The data were intended to evaluate aGvHD risk factors in patients with thalassemia. A total of 407 patients with thalassemia who underwent HCT between August 1, 2019, and December 31, 2023, were included in the study, with follow-up until March 31, 2024.

Study patients

Inclusion criteria were as follows: (1) Confirmed thalassemia by genetic diagnosis, clinically transfusion-dependent; and (2) Allogeneic hematopoietic stem cell transplantation (HSCT) performed using the "GX-07-TM" protocol (3). Donor types included matched sibling donors, matched unrelated donors, and haploidentical-related donors. Exclusion criteria were: (1) Secondary transplantation; (2) Patients without aGvHD and death from other causes within 100 days; (3) Graft failure or relapse within 100 days; (4) Patients had aGvHD early before the engraftment period. The primary endpoint was whether aGvHD occurred within 100 days. aGvHD was diagnosed and classified by at least two doctors according to Glucksberg grading.²²

Predictors

Based on previous literature and the expertise of transplant physicians regarding aGvHD risk

factors post-transplantation, relevant factors were collected. While predictive factors with data loss exceeding 15% were excluded (e.g., B cells and NK cells count of stem cells, peripheral blood B cells, and NK cells count at engraftment). Clinical variables recorded for risk factor analysis and model development included age, gender, TM gene type, splenectomy status, liver and spleen size, ABO blood type, and ferritin levels. Donor information such as age, gender, and ABO blood type was also collected. Transplantation protocol details encompassed human leukocyte antigen (HLA) match, graft source, graft-versus-host disease (GvHD) prophylaxis (e.g., cyclosporine, tacrolimus), total mononuclear cells (MNC) count, and CD34 count of infused stem cells. Graft samples were extracted for lymphocyte subset analysis (CD3+T cells, CD4+T cells, and CD8+T cells). The first-day post-infusion and the patient's neutrophil and platelet engraftment times were recorded. Neutrophil engraftment recovery was defined as the first of three consecutive days with an absolute neutrophil count of $\geq 0.5 \times 10^9/L$ in peripheral blood. Upon achieving this threshold, peripheral blood was drawn within 24h for lymphocyte subsets analysis. Platelet engraftment recovery was defined as the first of seven consecutive days with an untransfused platelet count of $\geq 20 \times 10^9/L$.

Detection of T-lymphocyte subsets

Samples were collected in EDTA-coated tubes, stained with the respective antibodies (BD Multitest TM IMK Kit, PerCP Mouse Anti-Human CD45, FITC Mouse Anti-Human CD3, APC Mouse Anti-Human CD4, PE Mouse Anti-Human CD8, PE Mouse Anti-Human CD16, PE Mouse Anti-Human CD56, APC Mouse Anti-Human CD19 were used). Samples were divided and labeled in 12 mm \times 75 mm test tubes. Added 20 μ L of BD Multitest reagents to each tube: CD3-FITC/CD8-PE/CD45-PerCP/CD4 APC and CD3-FITC/CD16+CD56-PE/CD45-PerCP/CD19-APC. Mixed with 50 μ L of anticoagulant-treated blood, vortex, and incubate in the dark at 20°C–25°C for 15 min. Added 450 μ L of 1 \times BD Multitest IMK kit lysing solution, vortex, and incubate again for 15 min. Analyzed samples within 1h for accuracy. FACSCanto flow cytometer (BD, Franklin Lakes, New Jersey, USA) was for the test. Data were analyzed using BD Diva Software. Lymphocytes were gated using CD45 PerCP and SSC gates. CD3+ cells and

CD3- cells were then distinguished, and CD3+ CD4+ and CD3+ CD8+ cells were delineated accordingly. The proportions of CD3+, CD3+ CD4+, and CD3+ CD8+ cells in lymphocytes were determined, and the absolute numbers of these subsets were calculated by combining them with the total white blood cell count obtained from routine blood examination.

Transplantation protocol

The GX-07-TM regimen comprised busulfan (Bu, 1 mg/kg, administered i.v. four times daily for 4 days, from days -9 to -6), cyclophosphamide (Cy, 50 mg/kg/day, administered i.v. for 4 days, from days -5 to -2), fludarabine (Flu, 50 mg/m²/day, administered i.v. for 3 days, from days -12 to -10), and anti-thymocyte globulin (ATG, 2.5 mg/kg/day, administered i.v. for 4 days, from days -4 to -1). Prior to transplantation, all patients received 30 mg/kg hydroxyurea orally once daily for 2 months.

The GvHD prophylaxis regimen included a calcineurin inhibitor (cyclosporine A mostly for patients with matched sibling donors, and tacrolimus mostly for other transplant patients), methotrexate, low-dose, short-course mycophenolate mofetil.⁴ For sibling and haploid transplantations, a preference was given to granulocyte colony-stimulating factor (G-CSF) mobilized bone marrow combined with G-CSF mobilized peripheral blood stem cells (G-PBSCs) infusion after evaluating donor and recipient conditions, while G-PBSCs infusion was utilized for unrelated transplantations.

Model development and validation

Based on previous studies, the liver was categorized into two groups using a cut-off value of 5 cm, the spleen at 3 cm, ferritin at 2500 ng/mL, and donor age at 40 years old.^{1,11,23} To improve the accuracy and statistical power of the estimators, we conducted Cox regression analysis on all patients with complete clinical indicators to screen for aGvHD risk factors. Final independent clinical risk factors were determined through backward stepwise Cox multivariate regression with *p* value <0.1. T-lymphocyte subsets data were completely missing at random, with no statistical differences between missing and complete data (Supplementary Table 1). There were 240 cases meeting inclusion criteria, which were then

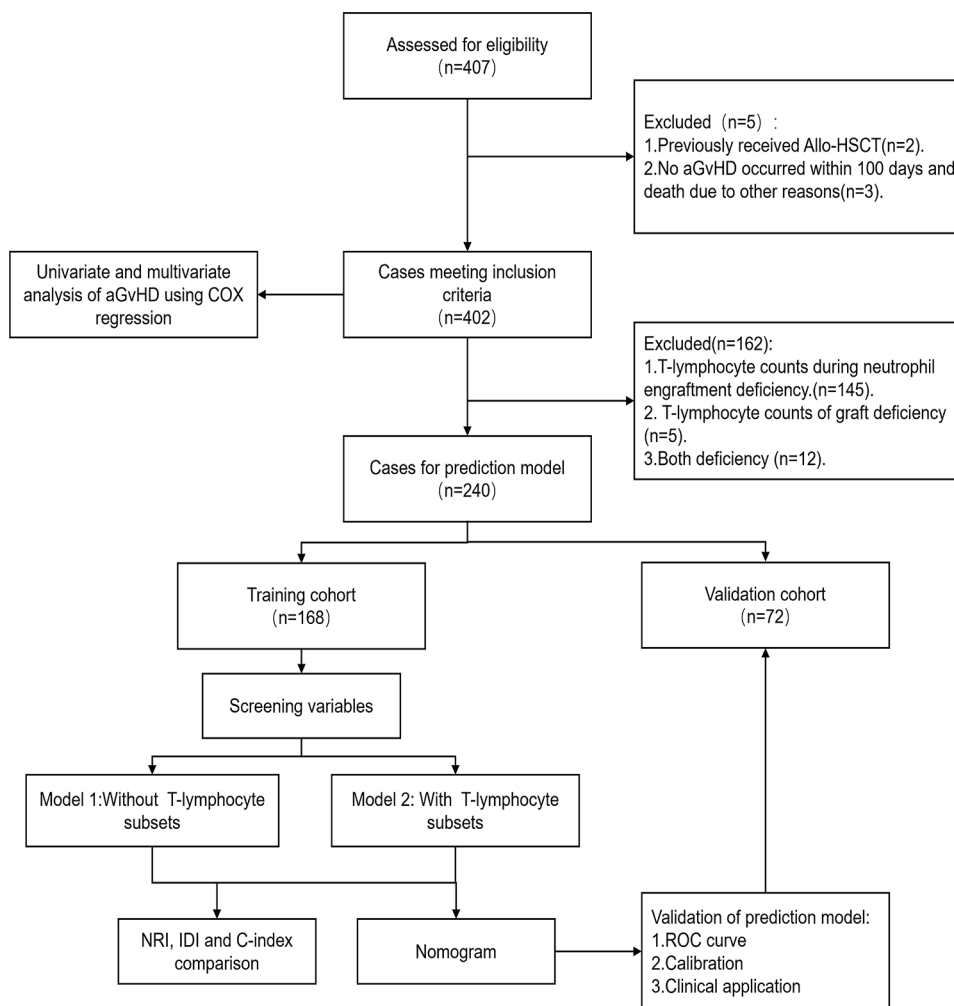


Figure 1. Patients' flow diagram.

aGvHD, acute graft-versus-host disease; Allo-HSCT, allogeneic hematopoietic stem cell transplantation; IDI, Integrated Discrimination Improvement Index; NRI, Net Reclassification Improvement Index; ROC, receiver operating characteristic curve

randomly allocated into training and validation cohorts at a ratio of 7:3 (Figure 1). The training cohort was utilized for model development, while the validation cohort was employed for model validation. The Least Absolute Shrinkage and Selection Operator (LASSO) method, utilizing the “glmnet” package, was employed to select parameters with nonzero coefficients. Two Cox regression models were constructed. Performance improvement due to T-lymphocyte subsets was evaluated through a comparison of the C-index, net reclassification improvement index (NRI), and integrated discrimination improvement index (IDI). Receiver operating characteristic (ROC) curve, C-index, calibration curve, and decision curve analysis (DCA) were employed in both

training and validation sets to assess the performance of model.

Statistical analysis

The categorical variables are presented as *n* (%) and continuous variables as median (IQR). Statistical comparisons for categorical variables were conducted using the Pearson's Chi-squared test and Fisher's exact test, while continuous variables were compared using the nonparametric Mann–Whitney *U* test. Hazard ratios (HR) and 95% confidence intervals (CI) were calculated for aGvHD risk factor analysis using the Cox proportional hazards model. The cut-off value for T-lymphocyte subsets was determined through

restricted cubic spline (RCS) regression, adjusting for type of transplantation, G-PBSCs infusion, and donor age. Nonlinear continuous variables were categorized based on the cut-off values during model construction. Data analysis and graph creation were performed using SPSS 22.0 (IBM Corp., Armonk, NY, USA) and R version 4.2.1 (R Core Team, Vienna, Austria). A p -value of <0.05 was considered statistically significant.

Results

Patients

During the study period, 407 patients underwent allogeneic HSCT. Excluding two patients with prior HSCT and three patients who died within 100 days without an outcome event (Figure 1), the analysis included 402 patients (Table 1). The median age of the patients was 8 years (range 2–20). Among the 402 patients, 146 (36%) developed aGvHD, with 95 (24%) cases classified as grade 2–4 aGvHD and 38 (9.5%) as grade 3–4 aGvHD. The median time to aGvHD onset was 29 days (range 9–142). The follow-up time is 30.8 (range 0.8–56.8) months. There were 15 deaths, all linked to transplantation, and one case of secondary graft failure. Among 15 deceased patients, 11 had aGvHD, and 8 deaths were aGvHD-related. Cox proportional regression analysis demonstrated that aGvHD correlated with adverse outcomes in terms of overall survival (HR 5.04, 95% CI: 1.61–15.84, $p=0.006$) and TFS (HR 4.03, 95% CI: 1.40–11.60, $p=0.01$).

Clinical risk factors of aGvHD

In the analysis of 402 patients, 146 developed aGvHD (Table 1). Univariate Cox regression analysis identified significant differences in transplantation type, HLA mismatch points, donor and recipient age, graft source, and calcineurin inhibitor (CNI) type (Table 2). Factors like CNI type and bone marrow stem cell infusion were closely linked to transplantation type and thus excluded from multivariate analysis. Ultimately, patient age, donor age, ferritin levels, transplantation type, HLA mismatch points, ABO matching, CD34+ cells count, G-PBSCs infusion, and umbilical cord blood stem cell infusion were included in backward stepwise Cox multivariate regression. Independent risk factors for aGvHD comprised unrelated donor (HR 3.63, 95% CI: 2.17–6.05, $p<0.001$), haploid-related donor

(HR 4.03, 95% CI: 2.40–6.77, $p<0.001$), peripheral blood stem cell infusion (HR 2.11, 95% CI: 1.09–4.11, $p=0.03$), and donor age older than 40 years (HR 1.53, 95% CI: 1.00–2.33, $p=0.049$; Table 2).

Predictors selection

Cases lacking T-lymphocyte subsets were excluded, leaving 240 datasets meeting inclusion criteria, randomly divided into training and verification cohorts at a 7:3 ratio (Figure 1). The training cohort comprised 168 cases, while the verification cohort had 72, with no significant baseline characteristic differences (Table 3). In the training cohort, LASSO coefficient curves were generated for 10 characteristics, including transplant type, donor age group, G-PBSCs, and T-lymphocyte subsets (CD3+ T cells, CD4+ T cells, and CD8+ T cells) in graft and peripheral blood during neutrophil engraftment, using 10-fold cross-validation. At Lambda.1se value of 0.05286106, nonzero variables were identified: matched unrelated donor, haploid-related donor, peripheral blood stem cell infusion, donor age group, CD4+ T-cell, and CD8+ T-cell count during neutrophil engraftment (Figure 2).

T-lymphocyte subsets

In 240 patients, we utilized RCS to ascertain the cut-off values for CD4+ T-cell and CD8+ T-cell counts at engraftment. Following adjustment for transplantation type, peripheral blood stem cells infusion, and donor age, the distribution of CD4+ T cells was found to be nonlinear. The risk of aGvHD escalated when the cell count exceeded 36 cells/ μ L. Conversely, the distribution of CD8+ T cells exhibited linearity, with a heightened risk of aGvHD observed when the count surpassed 43 cells/ μ L (Figure 3(a) and (b)). Cox proportional regression analysis indicated that elevated T-lymphocyte counts amplified the risk of total aGvHD, grade 2–4 aGvHD, and grade 3–4 aGvHD (Figure 3(c)).

Nomogram model development

The primary endpoint was the occurrence of aGvHD within 100 days. Model 1 was constructed using Cox regression, incorporating the type of transplantation, peripheral blood stem cell infusion, and donor age. Based on Model 1, Model 2

Table 1. Patient characteristics.

Variable	Total (N= 402)	Non-aGvHD (N= 256)	aGvHD (N= 146)
Total aGvHD, n (%)	146 (36)	—	146 (100)
Grade 2–4 aGvHD, n (%)	95 (24)	—	95 (65)
Grade 3–4 aGvHD, n (%)	38 (9.5)	—	38 (26)
Sex, n (%)			
Female	160 (40)	108 (42)	52 (36)
Male	242 (60)	148 (58)	94 (64)
Age, years, median (IQR)	8.38 (5.86, 11.24)	7.97 (5.48, 10.97)	9.45 (6.62, 11.74)
Genotype, n (%)			
α	15 (3.7)	10 (3.9)	5 (3.4)
β	387 (96)	246 (96)	141 (97)
Serum ferritin, ng/mL, n (%)			
Median (IQR)	3555.32 (2347.56, 5414.97)	3323.69 (2230.80, 5058.93)	3942.03 (2606.86, 6121.28)
≥ 2500	288 (72)	175 (68)	113 (77)
< 2500	114 (28)	81 (32)	33 (23)
Spleen group, cm, n (%)			
Splenomegaly < 3	293 (73)	183 (71)	110 (75)
Splenomegaly ≥ 3	57 (14)	35 (14)	22 (15)
Splenectomy	52 (13)	38 (15)	14 (9.6)
Hepatomegaly, cm, n (%)			
Liver < 5	380 (95)	242 (95)	138 (95)
Liver ≥ 5	22 (5.5)	14 (5.5)	8 (5.5)
Type of Allo-HCT, n (%)			
Matched sibling donor	140 (35)	120 (47)	20 (14)
Matched unrelated donor	137 (34)	70 (27)	67 (46)
Haploidentical-related donor	125 (31)	66 (26)	59 (40)
Number of HLA mismatch, n (%)			
0–1 point	281 (70)	193 (75)	88 (60)
2–3 points	22 (5.5)	13 (5.1)	9 (6.2)
≥ 4 points	99 (25)	50 (20)	49 (34)

(Continued)

Table 1. (Continued)

Variable	Total (N=402)	Non-aGvHD (N=256)	aGvHD (N=146)
Donor age, n (%)			
Age <40	358 (89)	240 (94)	118 (81)
Age ≥40	44 (11)	16 (6.3)	28 (19)
Female donor for male recipients, n (%)			
No	317 (79)	201 (79)	116 (79)
Yes	85 (21)	55 (21)	30 (21)
Donor/patient ABO-match, n (%)			
Match	206 (51)	141 (55)	65 (45)
Mismatch	196 (49)	115 (45)	81 (55)
G-PBSCs, n (%)	348 (87)	212 (83)	136 (93)
G-BM, n (%)	264 (66)	186 (73)	78 (53)
CB, n (%)	33 (8.2)	30 (12)	3 (2.1)
CSA-based GvHD prophylaxis, n (%)	140 (35)	119 (46)	21 (14)
TAC-based GvHD prophylaxis, n (%)	262 (65)	137 (54%)	125 (86)
Total MNCs, ×10 ⁸ /kg, median (IQR)	11.77 (9.08, 14.93)	11.67 (9.03, 14.85)	11.81 (9.51, 15.32)
Total CD34+ cells, ×10 ⁶ /kg, median (IQR)	8.44 (6.87, 11.74)	8.73 (7.01, 11.98)	8.15 (6.48, 10.64)
Neutropenia, days, median (IQR)	13.00 (11.00, 14.00)	13.00 (11.00, 14.00)	12.00 (11.00, 15.00)
Neutrophil engraftment, days, median (IQR)	12.00 (11.00, 14.00)	12.00 (11.00, 14.00)	11.50 (11.00, 14.00)
Platelet engraftment, days, median (IQR)	13.00 (11.00, 18.00)	13.00 (12.00, 18.00)	13.00 (11.00, 17.75)
aGvHD, acute graft-versus-host disease; IQR, interquartile range; Allo-HCT, allogeneic hematopoietic stem cell transplantation; HLA, human leukocyte antigen; G-PBSCs, granulocyte colony-stimulating factor mobilized peripheral blood stem cells; G-BM, granulocyte colony-stimulating factor mobilized bone marrow; CB, umbilical cord blood; CSA, cyclosporine A; TAC, tacrolimus; GvHD, graft-versus-host disease; MNC, mononuclear cell.			

was developed by adding the CD4+ T-cell group and CD8+ T-cell at the time of engraftment (Table 4). In the training set, a comparison of the C-index between the two models (Model 1 vs Model 2: 0.69 vs 0.73, $p=0.02$) revealed a significantly higher C-index for Model 2. The NRI for comparison of models was 0.24 (95% CI 0.09–0.39, $p=0.04$), and the IDI was 0.08 (95% CI 0.02–0.16, $p<0.001$). In Model 2, illustrated in Figure 4, the patient's type of transplantation, peripheral blood stem cell infusion status, CD4+

T-cell group, and CD8+ T-cell count were assigned corresponding points. The total score derived reflected the patient's risk score, indicating the likelihood of aGvHD-free occurrence within 100 days.

Model performance

ROC curves and C-index assessed Model 2's predictive performance. Area under the curves for total aGvHD, grade 2–4 aGvHD, and grade 3–4

Table 2. Univariable and multivariable Cox regression analysis of aGvHD in 402 patients.

Characteristic	Univariable analysis			Multivariable analysis		
	HR	95% CI	p	HR	95% CI	p
Age	1.05	1.00–1.09	0.032			
Serum ferritin \geq 2500 ng/mL	1.42	0.97–2.10	0.075			
Spleen group, cm						
Splenomegaly $<$ 3	Ref	—	—			
Splenomegaly \geq 3	0.99	0.63–1.57	0.969			
Splenectomy	0.67	0.38–1.17	0.157			
Hepatomegaly, cm						
Liver $<$ 5	Ref	—	—			
Liver \geq 5	0.97	0.47–1.97	0.928			
Type of Allo-HSCT						
Matched sibling donor	Ref	—	—	Ref	—	—
Matched unrelated donor	4.36	2.64–7.18	$<$ 0.001	3.63	2.17–6.05	$<$ 0.001
Haploidentical related donor	4.26	2.56–7.07	$<$ 0.001	4.03	2.40–6.77	$<$ 0.001
Number of HLA mismatch						
0–1 point	Ref	—	—			
2–3 points	1.45	0.73–2.88	0.289			
\geq 4 points	1.86	1.31–2.64	$<$ 0.001			
Donor age \geq 40 years	2.38	1.57–3.59	$<$ 0.001	1.53	1.00–2.33	0.049
Female donor for male recipient	0.97	0.65–1.45	0.893			
Donor/patient ABO-mismatch	1.36	0.98–1.88	0.067			
G-PBSCs	2.46	1.29–4.67	0.006	2.11	1.09–4.11	0.03
G-BM	0.53	0.38–0.73	$<$ 0.001			
CB	0.19	0.06–0.60	0.005			
TAC-based GvHD prophylaxis	4.07	2.56–6.47	$<$ 0.001			
Total MNCs, $\times 10^8$ /kg	1.02	0.99–1.05	0.202			
Toal CD34+ cells, $\times 10^6$ /kg	0.96	0.92–1.00	0.054			
Neutropenia, days	0.98	0.93–1.03	0.492			
Neutrophil engraftment, days	0.98	0.92–1.05	0.588			
Platelet engraftment, days	1.00	0.98–1.02	0.896			

Allo-HSCT, allogeneic hematopoietic stem cell transplantation; CB, umbilical cord blood; CI, confidence interval; HLA, human leukocyte antigen; HR, hazard ratio; G-PBSCs, granulocyte colony-stimulating factor mobilized peripheral blood stem cells; G-BM, granulocyte colony-stimulating factor mobilized bone marrow; TAC, tacrolimus; GvHD, graft versus host disease; MNC, mononuclear cell.

Table 3. Patient characteristics stratified by training cohort and validation cohort.

Variable	Training cohort (N=168)	Validation cohort, (N=72)	<i>p</i> ²
Total aGvHD, <i>n</i> (%)	61 (36)	33 (46)	0.166
Grade 2–4 aGvHD, <i>n</i> (%)	36 (21)	19 (26)	0.402
Grade 3–4 aGvHD, <i>n</i> (%)	16 (9.5)	8 (11)	0.707
Type of Allo-HSCT, <i>n</i> (%)			0.878
Matched sibling donor	57 (34)	24 (33)	
Matched unrelated donor	58 (35)	23 (32)	
Haploidentical-related donor	53 (32)	25 (35)	
Donor age, <i>n</i> (%)			0.562
Age <40	147 (88)	61 (85)	
Age ≥40	21 (13)	11 (15)	
G-PBSCs, <i>n</i> (%)	143 (85)	65 (90)	0.281
G-BM, <i>n</i> (%)	110 (65)	48 (67)	0.859
CB, <i>n</i> (%)	14 (8.3)	5 (6.9)	0.715
CD3+ T cells of graft source, ×10 ⁸ /kg, median (IQR)	2.08 (1.44, 3.36)	1.99 (1.54, 2.94)	0.873
CD4+ T cells of graft source, ×10 ⁸ /kg, median (IQR)	0.99 (0.55, 1.67)	0.99 (0.75, 1.36)	0.859
CD8+ T cells of graft source, ×10 ⁸ /kg, median (IQR)	0.94 (0.64, 1.48)	0.86 (0.62, 1.29)	0.462
CD4+/CD8+ ratio of graft source, median (IQR)	1.07(0.79, 1.41)	1.09(0.77, 1.37)	0.955
CD3+ T cells at neutrophil engraftment, cells/μL, median (IQR)	66.00 (17.75, 175.00)	73.50 (21.50, 140.25)	0.938
CD4+ T cells at neutrophil engraftment, cells/μL, median (IQR)	11.50 (2.00, 42.25)	13.50 (3.75, 28.00)	0.941
CD8+ T cells at neutrophil engraftment, cells/μL, median (IQR)	41.50 (8.75, 95.00)	45.00 (12.75, 94.50)	0.746
CD4+/CD8+ ratio at neutrophil engraftment, median (IQR)	0.32(0.17, 0.54)	0.29(0.13, 0.50)	0.290

aGvHD, acute graft-versus-host disease; Allo-HSCT, allogeneic hematopoietic stem cell transplantation; G-PBSCs, granulocyte colony-stimulating factor mobilized peripheral blood stem cells; G-BM, granulocyte colony-stimulating factor mobilized bone marrow; CB, umbilical cord blood; IQR, interquartile range.

aGvHD in the training cohort were 0.802, 0.809, and 0.846, respectively (Figure 5(a)), with corresponding C-index values of 0.736 (95% CI: 0.736–0.870), 0.755 (95% CI: 0.681–0.850), and 0.772 (95% CI: 0.6655–0.859), indicating robust predictive ability. In the validation cohort, AUCs were 0.777, 0.757, and 0.746 (Figure 5(d)), with corresponding C-index values of 0.691 (95% CI: 0.668–0.882), 0.672 (95% CI:

0.564–0.824), and 0.660 (95% CI: 0.498–0.902). Calibration curves for both cohorts showed good fit (Figure 5(b) and (e)), with total Brier scores below 0.25. The DCA curve illustrated the effective prediction of aGvHD occurrence within 100 days by the nomogram, yielding positive net benefits across different grades in the training cohort (Figure 5(c)). Similarly, the DCA curve for the validation cohort indicated positive net

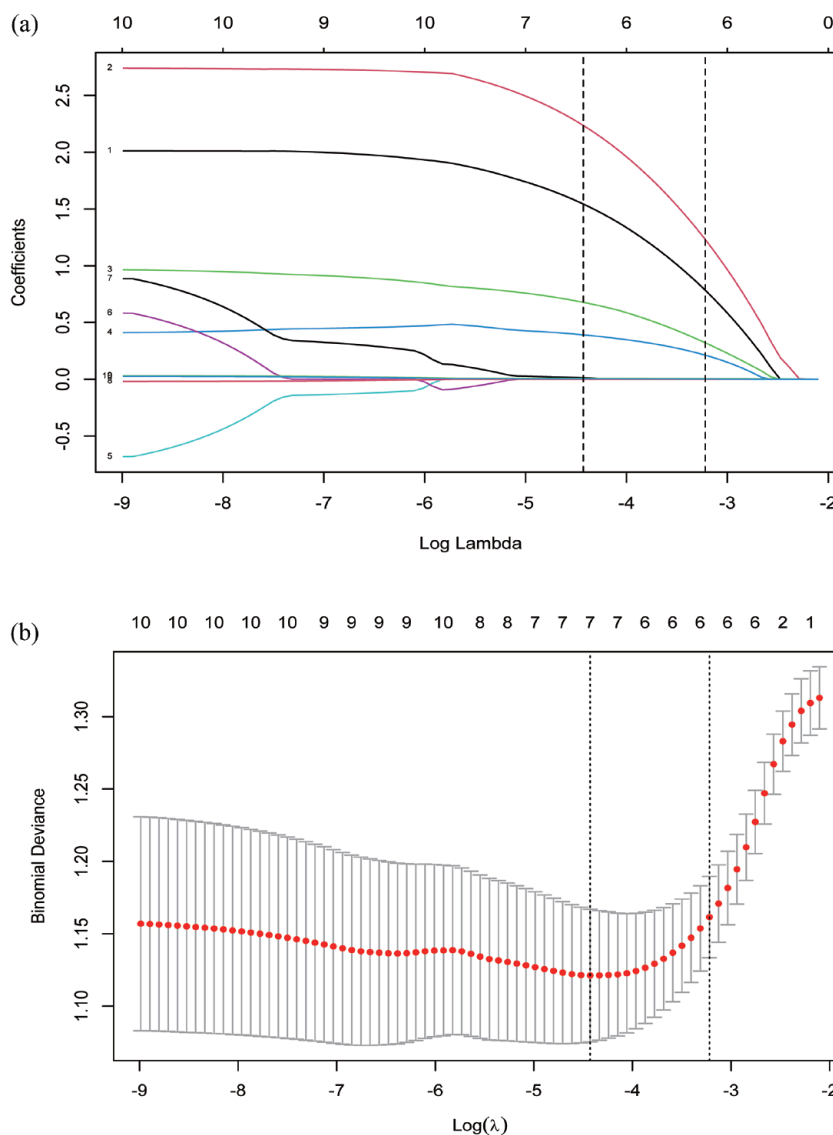


Figure 2. Perioperative variables selection using the LASSO regression. (a) The features with nonzero coefficients were selected by optimal lambda. (b) LASSO coefficient curves of the 10 variables. A coefficient profile plot was produced against the log (λ) sequence. Vertical line was drawn at the value selected using 10-fold cross-validation, where lambda.min resulted in seven nonzero coefficients, and Lambda.1se resulted in six nonzero coefficients. LASSO, Least Absolute Shrinkage and Selection Operator.

benefits for aGvHD of different grades within a certain range (Figure 5(f)).

Discussion

aGvHD is a major adverse factor impacting post-transplantation survival.¹¹ Our study found that patients with elevated counts of CD4+ and CD8+ T lymphocytes at the time of engraftment are at a

higher risk of developing aGvHD. Through Cox multivariate regression analysis, we also identified alternative donors, peripheral blood stem cell infusion, and donor age older than 40 years as aGvHD risk factors. We successfully developed an aGvHD prediction model that integrates clinical characteristics with T-cell data. The NRI and IDI metrics show that the inclusion of T-lymphocyte subsets significantly enhances the model's performance.

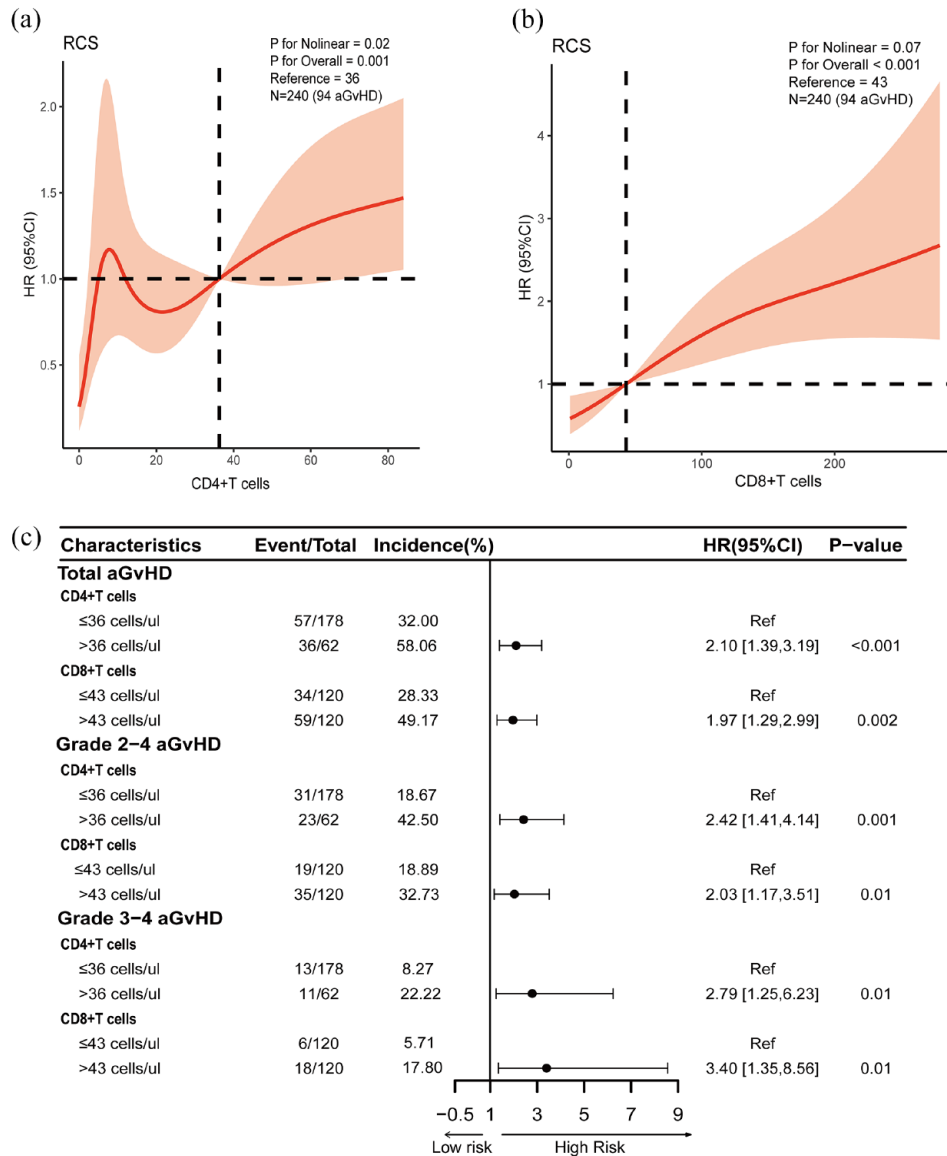


Figure 3. Potential nonlinear for the levels of CD4+ T cells and CD8+ T cells during neutrophil engraftment with aGvHD and measured by restricted cubic spline regression. The red lines and shadows represent the estimated HR and the 95% CI, respectively. CD4+ T cells 36 was selected as the reference level. Forty-three was selected as the CD8+ T cells reference level. (a) CD4+ T cells and total aGvHD in 240 patients (knots=5). (b) CD8+ T cells and total aGvHD in 240 patients (knots=3). (c) Comparison of different grades of aGvHD at 100 days post-transplantation using Cox proportional regression based on reference values of CD4+ T cells and CD8+ T cells. aGvHD, acute graft-versus-host disease.

This is the first aGvHD model that combines T-lymphocyte data and clinical characteristics, specifically for thalassemia patients, ensuring disease consistency and reducing heterogeneity risks. The model is easy to implement in clinical practice, providing clinicians with a practical tool to predict and manage patients.

Both CD4+ and CD8+ T cells play pivotal roles in the development of aGvHD through immune regulation, cytokine secretion, and cytotoxicity.⁸ Our study shows that elevated counts of CD4+ and CD8+ T cells at the time of engraftment are related to aGvHD occurrence. Although there was no statistically significant difference in

Table 4. Multivariate Cox regression analysis of aGVHD in two models based on data in the training cohort.

Variables	Model 1		Model 2	
	HR (95% CI)	p	HR (95% CI)	p
Matched unrelated donor	4.13 (1.78–9.59)	<0.001	4.76 (1.98–11.42)	<0.001
Haploid related donor	5.90 (2.51–13.84)	<0.001	9.29 (3.60–24.01)	<0.001
G-PBSCs infusion	1.04 (0.54–1.99)	0.92	1.27 (0.65–2.48)	0.48
Donor age over 40years	3.13 (1.09–9.04)	0.03	2.22 (0.75–6.58)	0.15
CD4+ T cells over 36 cells/ μ L	—	—	1.40 (0.72–2.73)	0.32
CD8+ T cells count	—	—	1.00 (1.00–1.01)	0.007

aGVHD, acute graft-versus-host disease; HR, hazard ratio; CI, confidence interval; G-PBSCs, granulocyte colony-stimulating factor mobilized peripheral blood stem cells.

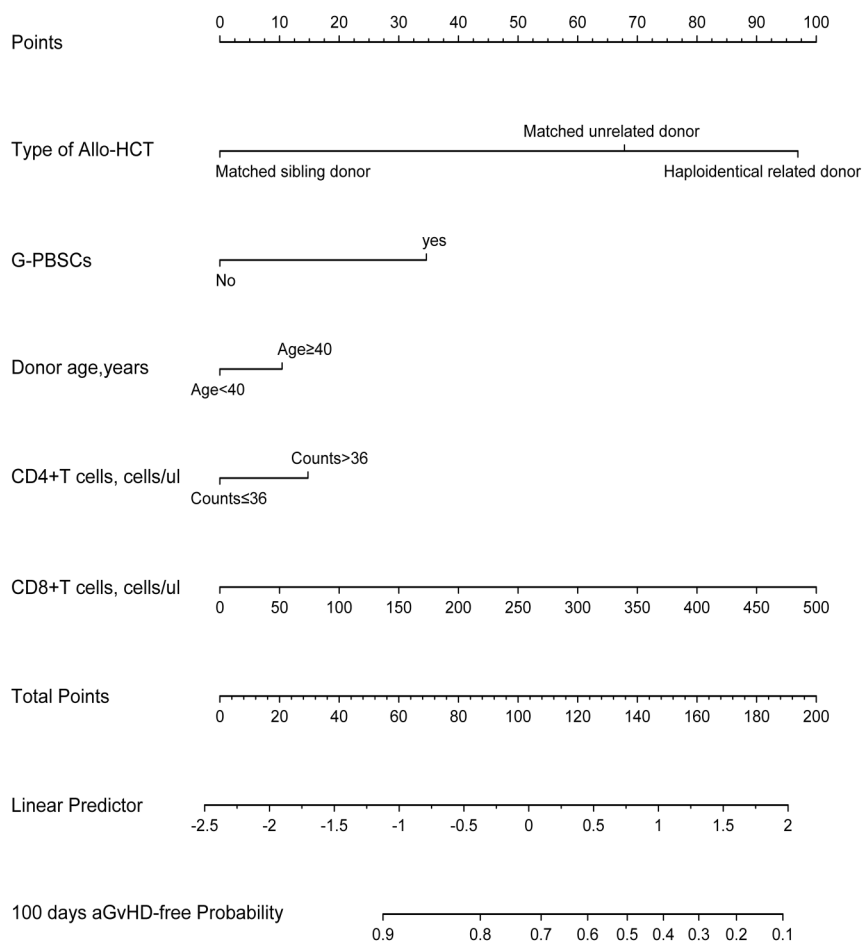


Figure 4. Prediction nomogram of aGVHD after 100 days of transplantation. aGVHD, acute graft-versus-host disease.

CD4+T-cell classification in Model 2, considering the correlation between CD4+ and CD8+ T-cells ($r=0.70$, $p<0.001$). So, we still included both variables in the model construction. Furthermore, RCS regression adjusted by clinical risk factors revealed that aGvHD risk increases when CD4+ T cells exceed 36 cells/ μ L and CD8+ T cells surpass 43 cells/ μ L. This is consistent with previous studies showing that patients with aGvHD had higher CD4+ and CD8+ T-cell counts 32 days post-transplantation.²⁴ Additionally, CD4+ T cells exceeding 38 cells/ μ L at day 28 post-transplantation have been identified as a risk factor for aGvHD.²⁰ In patients with thalassemia, research indicates that a rise in CD4+ T cells on day 14 post-transplantation serves as a valuable biomarker for predicting aGvHD. Furthermore, an increase in CD8+ T cells could be indicative of a more severe case of aGvHD.²⁵ However, some conflicting findings exist. There are studies showing that the number of T cells after transplantation is not parallel to the occurrence of aGvHD,²⁶ and patients lacking CD4+ T-cell reconstitution before the onset of aGvHD suggest a worse prognosis.²⁷ This potentially due to differences in aGvHD prophylaxis regimens, such as anti-T-lymphocyte globulin versus post-transplant cyclophosphamide.²⁸ Immune reconstitution and engraftment are dynamic processes, with neutrophil engraftment serving as a key early transplantation event. We focused on thalassemia patients at the time of neutrophil engraftment, aiming to reduce physiological variation. The relationship between T-lymphocyte counts and aGvHD is still under-explored, though our study shows that elevated T-cell counts at engraftment can predict aGvHD occurrence. Thus, vigilant monitoring of immunosuppressive drugs in high-risk patients is crucial, with adjustments like prolonged glucocorticoid administration, while balancing infection risks.

Clinical features play a pivotal role in the onset of aGvHD, including transplantation type, HLA match, graft source, donor and recipient age, disease remission status, and preconditioning regimen.^{9,11,29} Our study conducted under standardized conditions to ensure uniformity in underlying diseases and protocols, identified matched unrelated donors, haploid-related donors, peripheral blood stem cell infusion, and older donors as independent risk factors for aGvHD. Although

peripheral blood stem cell infusion had no statistically significant difference in the training cohort, considering the intimate association with aGvHD³⁰ and successful screening in LASSO regression, we still used this indicator to construct the model. While matched sibling donor and bone marrow transplantation remain preferred for TM due to high success rates and low GvHD incidence,^{1,4} scarcity of sibling donors and improved efficacy in alternative donor transplantation are leading more patients to choose alternatives. For alternative donors, haploid transplantation often entails HLA mismatch and unrelated donor grafts typically comprise peripheral blood stem cells, contributing to elevated GvHD incidence. Despite advancements in GvHD prophylaxis, grade 2–4 aGvHD incidence remains high in alternative donors.^{3,6,7} Thus, establishing more effective prophylaxis and enhancing post-transplantation monitoring are crucial. Our model empowers clinicians to evaluate the risk of aGvHD in thalassemia patients following transplantation, thereby optimizing management strategies and enhancing their quality of life.

We developed a nomogram model for aGvHD prediction by integrating clinical characteristics. Additionally, LASSO regression was employed to screen for T-lymphocyte subsets from graft sources and at the time of engraftment. This integration was expected to enhance the model's performance, addressing variable collinearity issues.³¹ LASSO regression screening revealed CD8+ T cells in the stem cell at Lambda.min, and CD4+ and CD8+ T cells at neutrophil engraftment as nonzero variables. However, RCS regression, adjusted for transplantation type, peripheral blood stem cell infusion, and donor age, found no statistically significant relationship between CD8+ T lymphocytes in stem cells and aGvHD risk. The impact of T lymphocytes in the graft on aGvHD risk is unclear, with conflicting conclusions among studies.^{32,33} The higher T-lymphocyte content in peripheral blood stem cells compared to bone marrow by tenfold³⁴ may explain the lack of significance in CD8+ T cells post-adjustment. At neutrophil engraftment, CD4+ and CD8+ T cells continued to indicate aGvHD risk after RCS regression adjustment, leading to their incorporation into the model. CD4+ T cells were treated as a categorical variable based on their nonlinear nature. Both CD8+

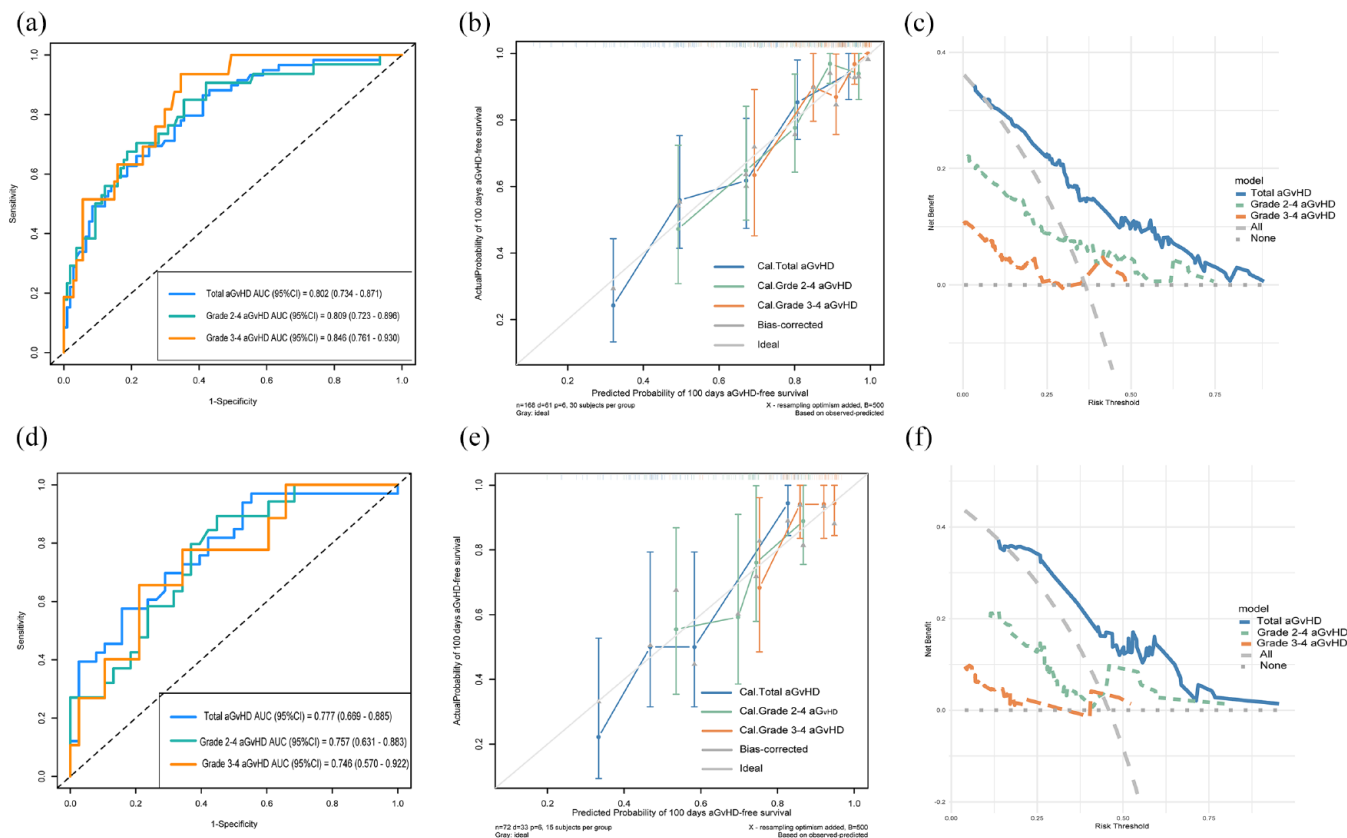


Figure 5. The performance of the nomogram in total aGvHD, grade 2–4 aGvHD, and grade 3–4 aGvHD: (a) receiver operating characteristic curve in the training cohort, (b) calibration curve in the training cohort, (c) decision curve analysis in the training cohort, (d) receiver operating characteristic curve in the validation cohort, (e) calibration curve in the validation cohort, and (f) decision curve analysis in the validation cohort. aGvHD, acute graft-versus-host disease.

and CD4+ T cells improved the model’s performance, which demonstrated good overall performance with easily obtainable indicators, suggesting clinical utility. Despite the model’s robust performance, there are limitations. Prospective, multicenter studies are needed to further validate it. Additionally, ongoing research aims to delve deeper into lymphocyte subsets, including Th1, Th2, Th17, and Treg cells, and analyzes lymphocyte cytokines, to both enhance the model’s predictive capabilities and shed light on the mechanisms behind aGvHD.

Conclusion

We developed the first successful aGvHD model for thalassemia transplant patients using T-lymphocyte subsets at engraftment. The model demonstrates excellent performance, with predictive indicators that are straightforward to obtain,

aiding clinicians in the early identification of high-risk aGvHD patients. This enables clinicians to optimize patient management, intervene promptly to halt disease advancement, thereby reducing mortality rates following transplantation in individuals with thalassemia.

Declarations

Ethics approval and consent to participate

The ethics committee of The First Affiliated Hospital of Guangxi Medical University approved this paper (Approval Number: No. 2023-S574-01). Written informed consent to participate in this study was obtained from the patients. Adults (≥ 18) finished the informed consent by themselves. For minors (< 18), informed consent was provided by the guardians on behalf of the minors involved in this study.

Consent for publication

All authors accept the terms and conditions of the editorial for publication.

Author contributions

Hongwen Xiao: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Software; Validation; Writing – original draft.

Gaohui Yang: Data curation; Formal analysis; Investigation; Methodology; Software; Validation.

Qiulin Huang: Data curation; Software.

Zhenbin Wei: Data curation.

Zhaoping Gan: Data curation.

Meiqing Wu: Data curation.

Zeyan Shi: Data curation.

Huicheng Huang: Investigation.

Zhaofang Pan: Investigation.

Lianjin Liu: Data curation.

Lingling Shi: Data curation.

Zhongming Zhang: Data curation; Formal analysis; Investigation.

Rongrong Liu: Conceptualization; Investigation; Methodology; Project administration; Writing – review & editing.

Yongrong Lai: Conceptualization; Investigation; Methodology; Project administration; Writing – review & editing.

Acknowledgements

The authors of this study would like to thank all the study participants for their contributions and Dr. Qi Zhou at the Gaobo Clinical Research Center, Shanghai, China for her statistical guidance.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by National Natural Science Foundation of China (Grant No. 82060029), Advanced Innovation Teams and Xinghu Scholars Program of Guangxi Medical University, and Open Project of NHC Key Laboratory of Thalassemia Medicine (GJWJWDP202304).

Competing interests

The authors declare that there is no conflict of interest.

Availability of data and materials

The datasets used and analyzed during the current study are available from the corresponding author on a reasonable request. Requests to access these datasets should be directed to laiyoungrong@hotmail.com.

ORCID iDs

Hongwen Xiao  <https://orcid.org/0000-0003-2363-0294>

Yongrong Lai  <https://orcid.org/0000-0002-0371-1719>

Supplemental material

Supplemental material for this article is available online.

References

1. Kattamis A, Kwiatkowski JL and Aydinok Y. Thalassemia. *Lancet* 2022; 399: 2310–2324.
2. Taher AT, Musallam KM and Cappellini MD. β -Thalassemias. *N Engl J Med* 2021; 384: 727–743.
3. He Y, Jiang H, Li C, et al. Long-term results of the NF-08-TM protocol in stem cell transplant for patients with thalassemia major: a multi-center large-sample study. *Am J Hematol* 2020; 95: E297–E299.
4. Li Q, Luo J, Zhang Z, et al. G-CSF-mobilized blood and bone marrow grafts as the source of stem cells for hla-identical sibling transplantation in patients with thalassemia major. *Biol Blood Marrow Transplant* 2019; 25: 2040–2044.
5. Li C, Wu X, Feng X, et al. A novel conditioning regimen improves outcomes in β -thalassemia major patients using unrelated donor peripheral blood stem cell transplantation. *Blood* 2012; 120: 3875–3881.
6. Li C, Mathews V, Kim S, et al. Related and unrelated donor transplantation for β -thalassemia major: results of an international survey. *Blood Adv* 2019; 3: 9.
7. Anurathapan U, Pakakasama S, Songdej D, et al. Haploidentical hematopoietic stem cell transplantation in thalassemia. *Hemoglobin* 2022; 46: 2–6.

8. Matte-Martone C, Liu J, Jain D, et al. CD8+ but not CD4+ T cells require cognate interactions with target tissues to mediate GVHD across only minor H antigens, whereas both CD4+ and CD8+ T cells require direct leukemic contact to mediate GVL. *Blood* 2008; 111: 3884–3892.
9. Zeiser R and Blazar BR. Acute graft-versus-host disease—biologic process, prevention, and therapy. *N Engl J Med* 2017; 377: 2167–2179.
10. Gooptu M and Koreth J. Translational and clinical advances in acute graft-versus-host disease. *Haematologica* 2020; 105: 2550–2560.
11. Lai Y-R, Chen Y-H, Hu D-M, et al. Multicenter phase II study of a combination of cyclosporine a, methotrexate and mycophenolate mofetil for GVHD prophylaxis: results of the Chinese Bone Marrow Transplant Cooperative Group (CBMTCG). *J Hematol Oncol* 2014; 7: 59.
12. Tang S, Chappell GT, Mazzoli A, et al. Predicting acute graft-versus-host disease using machine learning and longitudinal vital sign data from electronic health records. *JCO Clin Cancer Inform* 2020; 4: 128–135.
13. Zhang C, Huang W, Zhang P, et al. Dynamic changes in serum cytokine levels and their clinical significance in predicting acute GVHD. *Oncotarget* 2017; 8: 53691–53700.
14. Lee C, Haneuse S, Wang H-L, et al. Prediction of absolute risk of acute graft-versus-host disease following hematopoietic cell transplantation. *PLoS One* 2018; 13: e0190610.
15. Arai Y, Kondo T, Fuse K, et al. Using a machine learning algorithm to predict acute graft-versus-host disease following allogeneic transplantation. *Blood Adv* 2019; 3: 3626–3634.
16. Jo T, Arai Y, Kanda J, et al. A convolutional neural network-based model that predicts acute graft-versus-host disease after allogeneic hematopoietic stem cell transplantation. *Commun Med (Lond)* 2023; 3: 67.
17. Han L, Zhao K, Li Y, et al. A gut microbiota score predicting acute graft-versus-host disease following myeloablative allogeneic hematopoietic stem cell transplantation. *Am J Transplant* 2020; 20: 1014–1027.
18. Han L, Jin H, Zhou L, et al. Intestinal microbiota at engraftment influence acute graft-versus-host disease via the Treg/Th17 balance in Allo-HSCT recipients. *Front Immunol* 2018; 9: 669.
19. Paczesny S, Krijanovski OI, Braun TM, et al. A biomarker panel for acute graft-versus-host disease. *Blood* 2009; 113: 273–278.
20. McCurdy SR, Radojcic V, Tsai H-L, et al. Signatures of GVHD and relapse after posttransplant cyclophosphamide revealed by immune profiling and machine learning. *Blood* 2022; 139: 608–623.
21. Collins GS, Reitsma JB, Altman DG, et al. Transparent reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med* 2015; 162: 55–63.
22. Przepiorka D, Weisdorf D, Martin P, et al. 1994 consensus conference on acute GVHD grading. *Bone Marrow Transplant* 1995; 15(6): 825–828.
23. Ramprakash S, Raghuram CP, Marwah P, et al. Splenomegaly may increase the risk of rejection in low-risk matched related donor transplant for thalassemia, this risk can be partially overcome by additional immunosuppression during conditioning. *Biol Blood Marrow Transplant* 2020; 26(10): 1886–1893.
24. Huttunen P, Taskinen M, Siitonen S, et al. Impact of very early CD4(+)/CD8(+) T cell counts on the occurrence of acute graft-versus-host disease and NK cell counts on outcome after pediatric allogeneic hematopoietic stem cell transplantation. *Pediatr Blood Cancer* 2015; 62: 522–528.
25. Huang K and Luo J. Activated CD4+ T lymphocyte is a potential biomarker for acute graft-vs.-host disease after hematopoietic stem cell transplantation in children with transfusion-dependent β -thalassemia. *Front Pediatr* 2022; 10: 985306.
26. Waller EK, Logan BR, Fei M, et al. Kinetics of immune cell reconstitution predict survival in allogeneic bone marrow and G-CSF-mobilized stem cell transplantation. *Blood Adv* 2019; 3: 2250–2263.
27. de Koning C, Prockop S, van Roessel I, et al. CD4+ T-cell reconstitution predicts survival outcomes after acute graft-versus-host-disease: a dual-center validation. *Blood* 2021; 137: 848–855.
28. Massoud R, Gagelmann N, Fritzsche-Friedland U, et al. Comparison of immune reconstitution between anti-T-lymphocyte globulin and posttransplant cyclophosphamide as acute graft-versus-host disease prophylaxis in allogeneic myeloablative peripheral blood stem cell transplantation. *Haematologica* 2021; 107: 857–867.
29. Hahn T, McCarthy PL, Zhang M-J, et al. Risk factors for acute graft-versus-host disease

- after human leukocyte antigen-identical sibling transplants for adults with leukemia. *J Clin Oncol* 2008; 26: 5728–5734.
30. Yu X, Liu L, Xie Z, et al. Bone marrow versus peripheral blood as a graft source for haploidentical donor transplantation in adults using post-transplant cyclophosphamide-A systematic review and meta-analysis. *Crit Rev Oncol Hematol* 2019; 133: 120–128.
31. Tibshirani R. The lasso method for variable selection in the Cox model. *Stat Med* 1997; 16: 385–395.
32. Jiang P, Yu F, Xu X, et al. Impact of lymphocyte subsets of grafts on the outcome of haploidentical peripheral blood stem cell transplantation. *Cell Transplant* 2023; 32: 9636897231157054.
33. Körbling M, Huh YO, Durett A, et al. Allogeneic blood stem cell transplantation: peripheralization and yield of donor-derived primitive hematopoietic progenitor cells (CD34+ Thy-1dim) and lymphoid subsets, and possible predictors of engraftment and graft-versus-host disease. *Blood* 1995; 86: 2842–2848.
34. Waller EK, Logan BR, Harris WAC, et al. Improved survival after transplantation of more donor plasmacytoid dendritic or naïve T cells from unrelated-donor marrow grafts: results from BMTCTN 0201. *J Clin Oncol* 2014; 32: 2365–2372.

Visit Sage journals online
[journals.sagepub.com/
home/tah](https://journals.sagepub.com/home/tah)

 Sage journals