REPUBLIC OF TÜRKİYE
ANKARA UNIVERSITY
GRADUATE SCHOOL OF HEALTH SCIENCES

PREDICTION OF SURVIVAL AND GRAFT VERSUS HOST DISEASE IN ALLOGENEIC HEMATOPOIETIC STEM CELL TRANSPLANT PATIENTS USING MACHINE LEARNING

Hikmet Can ÇUBUKÇU

DEPARTMENT OF STEM CELL AND REGENERATIVE MEDICINE

DOCTORAL THESIS

ADVISOR

Prof. Dr. Muhit ÖZCAN

ANKARA

2025

REPUBLIC OF TÜRKİYE
ANKARA UNIVERSITY
GRADUATE SCHOOL OF HEALTH SCIENCES

PREDICTION OF SURVIVAL AND GRAFT VERSUS HOST DISEASE IN ALLOGENEIC HEMATOPOIETIC STEM CELL TRANSPLANT PATIENTS USING MACHINE LEARNING

Hikmet Can ÇUBUKÇU

DEPARTMENT OF STEM CELL AND REGENERATIVE MEDICINE

DOCTORAL THESIS

ADVISOR

Prof. Dr. Muhit ÖZCAN

ANKARA

2025

ETHICAL STATEMENT

To Ankara University
Directorate of the Graduate School of Health Sciences,

The dissertation I prepared and presented for my doctoral degree, titled "Prediction Of Survival And Graft Versus Host Disease In Allogeneic Hematopoietic Stem Cell Transplant Patients Using Machine Learning," has been written by me in accordance with scientific ethics and values. The hypothesis of my thesis belongs entirely to my thesis advisor and me. The study presented in the thesis was conducted by me, and all the sentences and comments are my own.
I hereby declare the accuracy of the above-mentioned statements.

Student's Name and Surname: Hikmet Can ÇUBUKÇU
Date: 22.01.2025
Signature:

ACCEPTANCE AND APPROVAL

Ankara University Institute of Health Sciences
In the Department of Stem Cell and Regenerative Medicine
Prepared by Hikmet Can Çubukçu,
the thesis titled
"Prediction of Survival and Graft Versus Host Disease in Allogeneic Hematopoietic Stem Cell Transplant Patients Using Machine Learning"
has been accepted as a DOCTORAL THESIS by majority vote of the jury listed below.

Thesis Defense Date: 22.01.2025

Signature
Prof. Dr. Muhit ÖZCAN
Ankara University
Jury Chair

Signature

İstinye University
Rapporteur

Signature

Ankara University
Member

Signature

Ankara University
Member

Signature

Atılım University
Member

The decision regarding the thesis has been approved by the Ankara University Institute of Health Sciences Board of Directors.

Signature

Title, Name, and Surname

Director of the Institute of Health Sciences

ÖZET

Makine Öğrenmesi Kullanılarak Allojenik Hematopoetik Kök Hücre Nakli Yapılan Hastalarda Sağkalım ve Graft Versus Host Hastalığı Tahmini

Allojeneik hematopoetik kök hücre nakli (HKHN), çeşitli hematolojik kanserler için potansiyel şifa sağlayıcı kritik bir tedavi seçeneğidir. Bununla birlikte HKHN, hasta yönetimini zorlaştıran ve sağkalımı etkileyen greft-versus-host hastalığı (GVHD), enfeksiyonlar ve nüks gibi önemli riskler taşır. Yapay zeka alanındaki gelişmelerle birlikte, makine öğrenmesi tıpta etkili bir araç olarak ortaya çıkmıştır. Makine öğrenmesi son zamanlarda, klinik sonuçların tahmin edilmesinde ve karar verme süreçlerine yardımcı olması amacıyla kullanılmaktadır. Bu çalışmada, hastanın sağkalımı, akut GVHD'nin ortaya çıkması, şiddeti ve kronik GVHD gelişimi dahil olmak üzere HKHN sonrası kritik sonuçların tahmin edilmesi için hasta verilerinin kullanılarak makine öğrenmesi modellerinin oluşturulması amaçlanmıştır.

Bu çalışmada, Ankara Üniversitesi Tıp Fakültesi'nden 1988-2023 yılları arasında 1.313 hastayı kapsayan retrospektif bir veri seti kullanılmıştır. Çalışmaya allojenik HKHN yapılmış yetişkin hastalar dahil edilmiş, solid organ nakli, otolog HKHN yapılan veya GVHD ve sağkalım durumu gibi temel sonuçları eksik olan hastalar hariç tutulmuştur. Veriler titizlikle ön işleme tabi tutulmuş, sağkalım, akut GVHD varlığı ve şiddeti ile kronik GVHD dahil olmak üzere nakil sonrası sonuçları tahmin etmek için az miktarda kod ile çalışan bir makine öğrenmesi kütüphanesi olan PyCaret kullanılarak makine öğrenmesi modelleri geliştirilmiştir. Çeşitli algoritmalar değerlendirilmiş ve performans için optimize edilmiş, her bir özelliğin tahminlere katkısını göstermek için shapley eklemeli açıklamalar analizi ile modeller yorumlanmıştır. Ayrıca, tahmini sonuçların kullanıcılara kolayca sunulabilmesi için geliştirilen modeller bir web tabanlı uygulamaya entegre edilmiştir.

CatBoost modeli, allojenik HKHN'nin ardından sağkalım ve akut GVHD'nin ortaya çıkması çıktıları için güçlü bir tahmin performansı sergilemiştir. Sağkalım tahmini için model %76,82'lik bir doğruluk ve 0,82'lik bir AUC değeri elde etmiştir. Sağkalım tahminin temel pozitif öngördürücüleri arasında pozitif engraftman ve siklosporin A + metotreksat ile GVHD profilaksisi yer alırken, yüksek nüks sayıları ve şiddetli akut GVHD sağkalımı olumsuz etkilemiştir. Akut GVHD'yi tahmin modeli %64,71 doğruluğa ve 0,68'lik bir AUC değerine sahip olup, CMV enfeksiyonu ana öngördürücü olarak ortaya çıkmıştır. Akut GVHD şiddetini tahmin etmek için oluşturulan Rastgele Orman modeli %71,93 doğruluğa ve 0,78'lük bir AUC değerine ulaşmış, uyumsuz lokus sayısı ve tüm vücut ışınlaması gibi faktörler daha şiddetli GVHD ile ilişkilendirilmiştir. Son olarak, kronik GVHD tahmin modeli %78,30 doğruluğa ve 0,75'lük bir AUC değerine ulaşmış, akut GVHD varlığı ve miyeloablatif hazırlık rejimi tedavileri güçlü öngördürücüler olurken, defibrotid profilaksisi ve düşük yoğunluklu hazırlık rejimi tedavisi daha düşük kronik GVHD riski ile ilişkilendirilmiştir.

Bu çalışmada geliştirilen makine öğrenimi modelleri, klinik karar vermede değerli araçlar olarak potansiyellerini güçlü tahmin performansları ile göstermiştir. Geliştirilen web tabanlı uygulama, modellerin potansiyel olarak klinik karar destek sürecine entegre edilebileceğini göstermektedir. Bununla birlikte, modellerin genellenebilirliklerini ve güvenilirliklerini artırmak için, dış doğrulama ve daha büyük, daha çeşitli veri kümelerinin kullanılması gerekmektedir.

Anahtar Kelimeler: graft versus host hastalığı, hematopoietik kök hücre nakli, makine öğrenmesi, sağkalım

SUMMARY

Prediction of Survival and Graft Versus Host Disease in Allogeneic Hematopoietic Stem Cell Transplant Patients Using Machine Learning

Allogeneic hematopoietic stem cell transplantation (HSCT) is an essential therapeutic approach for diverse hematologic cancers, providing a chance for a cure. However, this procedure is linked to substantial adverse events, such as graft-versus-host disease (GVHD), relapse, and infections, that complicate patient management, as well as impact survival. With advances in artificial intelligence, machine learning (ML) has demonstrated significant potential as a valuable tool in medicine, enabling the prediction of clinical outcomes and aiding in decision-making. This study aimed to build machine learning models utilizing our patient data to predict critical post-transplant outcomes, including survival, the onset and severity of acute GVHD, and the occurrence of chronic GVHD.

This study employed a retrospective dataset of 1,313 patients from the Ankara University Faculty of Medicine, spanning 1988 to 2023. We included adult allogeneic HSCT recipients, excluding patients who had undergone solid organ transplantation, autologous HSCT, or for whom data on key outcomes such as GVHD and survival status were unavailable. Data were meticulously preprocessed, and ML models were developed utilizing PyCaret, a low-code ML library to predict post-transplant outcomes, including survival, acute GVHD presence and severity, and chronic GVHD. Various algorithms were evaluated and optimized for performance, with model interpretability enhanced through shapley additive explanations to elucidate the feature-wise contributions to predictions. Furthermore, the finalized models were integrated into a web-based application, enabling users to input patient data and generate outcome predictions interactively.

The CatBoost model demonstrated strong predictive performance for survival and emergence of acute GVHD following allogeneic HSCT. For survival prediction, the model achieved a 76.82% accuracy and an AUC of 0.82. Main positive predictors were positive engraftment and GVHD prophylaxis with cyclosporine A + methotrexate, while higher relapse counts and severe acute GVHD negatively impacted survival. The model for predicting acute GVHD showed a 64.71% accuracy and an AUC of 0.68, with CMV infection emerging as key predictors. For predicting acute GVHD severity, the Random Forest model attained an accuracy of 71.93% and an AUC of 0.78, with factors like mismatched locus count and total body irradiation being related to more severe GVHD. Lastly, the chronic GVHD prediction model reached a 78.30% accuracy and an AUC of 0.75, with acute GVHD presence and myeloablative conditioning therapies being strong predictors, while defibrotide prophylaxis and reduced-intensity conditioning related to lower chronic GVHD risk.

ML models developed within this research exhibited robust predictive capabilities, highlighting their potential as valuable tools in clinical decision-making. The developed web-based application demonstrates the potential for models to be integrated into clinical decision-making processes. However, to enhance their generalizability and reliability, further external validation and the use of larger, more diverse datasets are necessary.

Keywords: graft versus host disease, hematopoietic stem cell transplantation, machine learning, survival

CONTENTS

PREFACE

I began my doctoral education in 2020 under the guidance of Prof. Meral Beksaç. During the first three years of my journey, she directed me to explore various fields, including graft-versus-host disease (GVHD) and its relationship with the microbiota, methods for isolating and expanding natural killer cells and their application in immunotherapies, and the isolation of stem cells from cord blood. Before her retirement, we focused on implementing machine learning methods to predict the outcomes of patients who had undergone allogeneic hematopoietic stem cell transplantation. Ultimately, we decided to concentrate on this latter topic for my doctoral thesis.

After the retirement of Prof. Meral Beksaç, Prof. Muhit Özcan guided me in completing my doctoral thesis. I am deeply grateful to both Prof. Meral Beksaç and Prof. Muhit Özcan for their invaluable support and for sharing their decades of expertise in the field of hematology.

I would also like to extend my heartfelt thanks to Assoc. Prof. Güldane Cengiz Seval for providing retrospective data on patients who had undergone allogeneic hematopoietic stem cell transplantation and for her insights into transplantation-related information. Additionally, I am sincerely grateful to all the medical doctors at the Department of Hematology, Ankara University Faculty of Medicine, for their tireless efforts in caring for patients over decades, from whom the data for this research was derived.

My special thanks go to my thesis monitoring committee members, Prof. Asım Egemen Yılmaz and Assoc. Prof. Klara Dalva, for their invaluable guidance throughout this journey. Lastly, I am profoundly thankful to my mother, father, and sister for their unwavering support, care, and love from the very beginning of my life.

SYMBOLS AND ABBREVIATIONS

ADV	Adenovirus
AI	Artificial Intelligence
ALL	Acute Lymphoblastic Leukemia
AML	Acute Myeloid Leukemia
AUC	Area Under the Curve
CMV	Cytomegalovirus
CML	Chronic Myeloid Leukemia
EBMT	European Group for Blood and Marrow Transplantation
GVHD	Graft-Versus-Host Disease
HBV	Hepatitis B Virus
HCT-CI	Hematopoietic Cell Transplantation-Specific Comorbidity Index
HLA	Human Leukocyte Antigen
HSCT	Hematopoietic Stem Cell Transplantation
HSV	Herpes Simplex Virus
IBMTR	International Bone Marrow Transplant Registry
KNN	K-Nearest Neighbors
MHC	Major Histocompatibility Complex
PBSCT	Peripheral Blood Stem Cell Transplantation
ROC	Receiver Operating Characteristic
SHAP	SHapley Additive exPlanations
TBI	Total Body Irradiation
VOD	Veno-Occlusive Disease
VZV	Varicella-Zoster Virus

FIGURES

TABLES

1. INTRODUCTION

1.1. Allogeneic hematopoietic stem cell transplantation

Hematopoietic stem cell transplantation (HSCT) is a pivotal therapeutic intervention for individuals suffering from hematologic diseases, involving the hematopoietic progenitor cell infusion to reestablish normal hematopoiesis and immune system. The classification of HSCT is predicated on the source of the hematopoietic cells, namely, autologous and allogeneic transplantation. Autologous HSCT involves utilizing the patient's hematopoietic cells, whereas allogeneic HSCT relies on cells received from a human leukocyte antigen-matched donor. HSCT has shown remarkable efficacy in treating a wide array of malignant hematological conditions (Saad et al., 2020).

Despite its potential curative outcomes, allogeneic HSCT is accompanied by significant risks and complications. One of the major challenges is graft-versus-host disease (GVHD), in which the immune cells of the donor attacks the tissues of recipient (Copelan et al., 2019). Additionally, patients undergoing allogeneic HSCT face considerable risk of infections, encompassing bacterial, fungal, and viral etiologies due to their immunocompromised state. Notable infections include those caused by Pseudomonas aeruginosa, Candida, Aspergillus species, cytomegalovirus (CMV), influenza, varicella-zoster virus, adenovirus, and respiratory syncytial virus (Sahin et al., 2016). CMV infection, in particular, can manifest as pneumonia, gastrointestinal disease, and retinitis in patients undergoing allogeneic HSCT, further complicating their clinical management (Allaw et al., 2023).

Another critical concern post-allogeneic HSCT is the relapse of a hematologic disease, such as acute myeloid leukemia, which presents a significant challenge for both treatment and prevention (Kreidieh et al., 2022). Graft failure, where the transplanted hematopoietic cells fail to integrate into the bone marrow of recipient and subsequently fail to generate novel blood cells, is another serious complication that can occur, necessitating further therapeutic interventions (Ozdemir & Civriz Bozdağ, 2018). Furthermore, patients may experience long-term complications, including secondary cancers and organ-specific issues resulting from pre-transplant conditioning regimens (Majhail, 2017).

The management of these adeverse events necessitates a multifaceted approach, incorporating prophylactic and therapeutic apporaches to mitigate the risks of infections, GVHD, relapse and graft failure. Despite these challenges, the potential curative benefits of allogeneic HSCT for specific hematologic cancers underscore its importance as a cornerstone in therapeutic modality for these conditions (Saad et al., 2020).

1.2. Graft versus host disease

GVHD occurs when the recipient's tissues are targeted by an immune response initiated by the donor's cells, mainly affecting the skin, eyes, gastrointestinal system, kidneys, lungs, heart, liver, and muscle. This condition arises from the a complex interplay between transplanted donor T cells and recipient's antigens, with human leukocyte antigens being of particular importance (Ferrara et al., 2009).

Acute GVHD is characterized by the absence of diagnostic criteria specific to chronic GVHD. It encompasses both classic acute GVHD, manifesting within the first a hundred days post-transplantation, and persistent or delayed acute GVHD, characterized by acute GVHD symptoms emerging beyond the 100-day mark, often coinciding with immunosuppression reduction. Conversely, chronic GVHD is categorized into classic chronic GVHD, which lacks acute GVHD features, and an overlap condition, that exhibits characteristics of both acute and chronic GVHD (Filipovich et al., 2005).

1.3. Acute graft versus host disease

The clinical presentation of acute GVHD typically involves a maculopapular rash, jaundice due to biliary duct damage and subsequent cholestasis, gastrointestinal disturbances characterized by vomiting, anorexia, diarrhea, and abdominal pain. Diagnosis of acute GVHD is established through clinical evaluation, laboratory testing, and tissue biopsy of affected organs. Disease severity is graded based on skin, gastrointestinal tract, and liver involvement. Mild, moderate, severe, and very severe GVHD correspond to grades 1, 2, 3, and 4, respectively. Approximately 30-50% of individuals undergoing allogeneic HSCT experience acute GVHD, with 14% developing severe forms of the disease (Zeiser & Blazar, 2017a).

The acute GVHD severity is defined by the tissue involvement, based on the grading criteria (Rowlings et al., 1997):

Grade A (I): Involves a rash influencing less than 25% of the body, no significant liver abnormalities (total bilirubin < 34 μmol/L) and gastrointestinal involvement (volume of diarrhoea < 500 ml/day)
Grade B (II): Rash covers 25-50% of the body, stage 1-2 liver involvement (total bilirubin: 34–102 μmol/L), stage 1-2 gut involvement (volume of diarrhoea: 550–1500 ml/day)
Grade C (III): Rash covers more than 50% of the body, stage 3 liver involvement (total bilirubin: 103–255 μmol/L), stage 3 gut involvement (volume of diarrhoea > 1500 ml/day)
Grade D (IV): Widespread, blistering rash, severe liver damage (total bilirubin > 255 μmol/L), stage 4 gut involvement (severe pain and ileus)

Several factors influence the occurance of acute GVHD:
Age: Advanced age in both the recipient and donor is a recognized contributing factor for developing acute GVHD (Flowers et al., 2011).
Gender mismatch: Female donors paired with male recipients exhibit a higher incidence of the disease (Flowers et al., 2011).
HLA mismatch: HLA proteins, encoded by genes within the major histocompatibility complex, are crucial for presentation of antigen and subsequent activation of T cells. Given their extensive genetic variability, the range of HLA discordance between HSCT recipient and donor serves as a major predictor of the risk of developing acute GVHD (Flowers et al., 2011).
HSCT source: Peripheral blood stem cell transplantation (PBSCT) is linked to a greater risk of acute GVHD in comparison to bone marrow transplantation. This fact is attributed to the greater rate of allo-reactive T cells present in peripheral blood grafts. Conversely, umbilical cord blood transplants exhibit a lower acute GVHD risk, likely due to the predominance of immature cells within the transplant (Flowers et al., 2011).
Conditioning regimen intensity: Reduced-intensity conditioning therapies have been linked to less acute GVHD risk in comparison to myeloablative regimens in some studies, although this finding is not consistently reported across all research (Jagasia et al., 2012).

1.4. Chronic graft versus host disease

Chronic GVHD can impact not only the skin, gastrointestinal system, lungs, and liver, which are generally affected in acute GVHD, but also other organ systems, involving the oral cavity, esophagus, musculoskeletal system, joints, fascia, eyes, lymphohematopoietic system, hair, nails, and genitalia (Zeiser & Blazar, 2017b).

Chronic GVHD can manifest with a diverse range of organ-specific symptoms. These can include cutaneous manifestations such as dyspigmentation, and alopecia; oral complications like xerostomia and ulcerations; ocular involvement such as keratoconjunctivitis sicca; and musculoskeletal complications such as fasciitis and myositis. Gastrointestinal issues include loss of appetite, weight loss, and esophageal strictures, while liver involvement can cause jaundice. Pulmonary complications include bronchiolitis obliterans Other organs and tissues may also be affected, including the kidneys (nephrotic syndrome), heart (pericarditis), and bone marrow (cytopenias) (Ferrara et al., 2009).

Several factors contribute to the onset of chronic GVHD (Zeiser & Blazar, 2017b):

Age: Advanced age in either the donor or recipient is a well-established contributing factor for chronic GVHD. This correlation is likely attributable to age-related modifications in immune function, tissue regenerative capacity, cellular aging, and cumulative pathogen burden.

Gender mismatch: Chronic GVHD incidence is higher in male recipients of female donor hematopoietic stem cell transplants, primarily due to donor T cells targeting host H-Y antigens.

HLA mismatch: HLA mismatching between unrelated recipient and donor grafts is linked to a greater risk of chronic GVHD owing to heightened alloimmune reactivity.

HSCT source: PBSCT are linked to a greater rates of chronic GVHD in comparison to bone marrow transplants, largely due to the higher counts of T cells in the PBSC graft.

Conditioning regimen intensity: Intense conditioning regimens are linked to increased tissue damage, which can exacerbate the initial phase of chronic GVHD.

Prior acute GVHD: Prior acute GVHD is related to a greater risk of developing chronic GVHD. This relationship is likely attributable to the persistent immune activation and tissue damage caused by the acute phase. While evidence suggests that severe acute GVHD (grades III and IV) may be more strongly correlated with chronic GVHD development, further investigation is warranted.

Chronic myelogenous leukemia diagnosis: Early studies showed a heightened risk of chronic GVHD among individuals undergoing allogeneic HSCT for chronic myeloid leukemia.

1.5. Factors affecting survival

The survival of individuals undergoing allogeneic HSCT affected by several critical factors. The relapse of the malignancy continues to be the predominant factor of mortality,

followed closely by chronic GVHD (Bhatia et al., 2007). Another significant contributor to mortality in these patients is Cytomegalovirus (CMV) infection (Allaw et al., 2023). Moreover, reactivation of CMV, when accompanied by bloodstream infections, further reduces overall survival (Ren et al., 2024). Additionally, acute GVHD and graft failure are linked to a higher mortality risk. Moreover, the timing of platelet and neutrophil engraftment has been identified as a substantial contributing factor for overall survival (Ren et al., 2024).

Despite the challenges associated with allogeneic HSCT, there have been significant improvements in clinical practice over recent years, resulting in a substantial reduction in non-relapse mortality (from 30% to 16% over 25 years), relapse-related mortality, and overall mortality (McDonald et al., 2020). Furthermore, the volume of patients treated at a given center has been shown to be a significant determinant of survival outcomes. In the United States, centers with higher patient volumes (>40 allogeneic HSCT per year) have reported a 62% survival rate at one year, compared to a 56% survival rate in lower-volume centers (≤40 allogeneic HSCT per year) (Majhail et al., 2020).

In addition, pre-hospital physical exercise interventions have been linked to lower mortality rates in patients undergoing allogeneic HSCT. Patients who received physical exercise interventions prior to hospital admission demonstrated better survival outcomes than those who did not (Wiskemann et al., 2015).

1.6. Risk Scores for patients' outcomes

Two primary risk scores are utilized to evaluate mortality risk in allogeneic HSCT patients. The European Group for Blood and Marrow Transplantation (EBMT) risk score was initially created specifically for individuals diagnosed with chronic myeloid leukemia but has since been validated for individuals with acute leukemia and other hematological diseases (Gratwohl et al., 2009; Shouval et al., 2017). This risk score idetified factors contributing to higher scores, including older patient age (≥40 years), advanced disease stage (intermediate or late), prolonged time between diagnosis and transplantation (>12 months), the utilization of unrelated donors, and female-to-male donor-recipient sex mismatches (Gratwohl, 2012).

Higher EBMT scores are associated with poorer survival outcomes. This association is supported by findings from a cohort, where acute leukemia patients with an EBMT score >10

had substantially worse overall survival incomparison to the patients with scores <8.5 (hazard ratio: 2.79 [2.02-3.87], two-year overall survival) (Shouval et al., 2017).

The second key risk score Hematopoietic Cell Transplantation-Specific Comorbidity Index (HCT-CI) evaluates the impact of pre-transplant comorbidities on HSCT outcomes, assigning weighted scores to various conditions based on their severity. Conditions include arrhythmias, cardiac issues, psychiatric disturbances, mild hepatic impairment, obesity, infections, and rheumatologic disorders, among others, with higher scores assigned to more severe conditions such as moderate/severe renal, pulmonary, or hepatic dysfunction, prior solid malignancies, and heart valve disease (Sorror et al., 2007). Increasing HCT-CI values have been associated with elevated non-relapse mortality (Sorror et al., 2007) and reduced overall survival (Kataoka et al., 2010).

1.7. Implementation of machine learning in medicine

Recent advancements have facilitated the usage of artificial intelligence (AI) in medicine. A subset of AI, machine learning employs statistical methodologies to predict target variables such as clinical outcomes and diagnoses (Çubukçu et al., 2024; Herman et al., 2021).

The creation of a ML model begins with data collection and preparation. This process followed data cleaning and feature engineering to extract relevant information. The curated dataset is subsequently divided into training, tuning, and validation subsets. The training subset forms the basis for model learning, allowing it to identify underlying patterns in the data. While optional, a tuning subset can be utilized to optimize model performance through hyperparameter adjustment. Once the model is constructed and refined, its performance, interpretability, and generalizability are rigorously assessed. This evaluation encompasses both internal and external validation datasets to ensure model robustness if possible (Çubukçu et al., 2024).

ML has become an essential tool in hematology, offering potential to enhance clinical decision-making in diagnosis, prognosis, and treatment (Radakovich et al., 2020). Artificial intelligence (AI) applications in diagnostic hematology encompass digital morphological analysis of smears, hemoglobinopathy testing, and flow cytometry (Obstfeld, 2023). Moreover, ML models have demonstrated utility in predicting prognosis, such as overall and non-relapse mortality following allogeneic HSCT (Mussetti et al., 2024). In the realm of

treatment, ML-based approaches are being explored for optimizing therapy selection, as exemplified by its application in chronic lymphocytic leukemia (CLL) (Agius et al., 2020).

While numerous ML models have been developed, their performance can vary significantly across different patient populations. The limited availability of readily accessible models tailored to specific patient cohorts underscores the need for developing in-house solutions. Therefore, this study aims to construct ML models utilizing our patient data to predict critical post-transplant outcomes, including survival, the onset and severity of acute GVHD, and the chronic GVHD development.

2. MATERIALS AND METHODS

2.1. Study population

Approval for this research was granted by Ankara University's Human Research Ethics Committee (decision number: İ09-596-23). We employed a retrospective dataset encompassing patient medical records from 1988 to 2023 at the hematology department of Ankara University Faculty of Medicine. Patients included in the study satisfied the criteria below:

Inclusion Criteria:

- Underwent allogeneic HSCT.
- 18 years of age or older.

Exclusion Criteria:

- Underwent solid organ transplantation.
- Underwent autologous HSCT.
- Aged under 18 years.
- Possessed missing data on dependent variables (variables targeted for prediction), including:
  - ✓ Acute GVHD occurrence
  - ✓ Acute GVHD severity (grades 1 vs. 2, 3, 4)
  - ✓ Chronic GVHD occurrence
  - ✓ Survival status (living or deceased)

  Patients with missing data on dependent variables were excluded from their respective prediction models.

Following application of the inclusion and exclusion criteria, we gathered available information on the following elements, to be used as input and dependent variables for subsequent analyses:

<u>HLA conformance information:</u>

Mismatched locus status: This variable indicates which specific HLA loci are mismatched between the recipient and donor, including A, B, C, and DRB1.

DRB1 mismatch status: This focuses specifically on mismatches at the DRB1 locus, which is part of the HLA class II region.

Mismatched locus count: This variable likely represents the total count of HLA loci mismatches between recipient and donor.
HLA numeric conformance: This is a numerical representation of HLA compatibility, ranging between 0 and 1. A value of 1 indicates a fully matched conformance, while lower values represent increasing degrees of mismatch between donor and recipient HLA profiles.
Haploidentical transplant status: This binary variable likely indicates whether the transplant is haploidentical, meaning the donor is a half-match with the recipient.

Demographical information:
Recipient age: The age of the individual receiving the transplant.
Donor age: The age of the individual donating the tissue.
Recipient gender: The gender identity of the transplant recipient.
Donor gender: The gender identity of the tissue donor.
Gender mismatch: A binary variable that indicates whether the recipient and donor have mismatched genders.
Recipient blood type: Denotes the blood type of the transplant recipient.
Donor blood type: Specifies the blood type of the organ or tissue donor.
Blood type mismatch: A binary variable that indicates whether the recipient and donor have mismatched blood types.
Main diagnosis: Provides the primary medical condition for which the transplant is needed.
Diagnosis: Additional details about the recipient's medical condition, complementing the main diagnosis.
Subtype: A subtype or variant of the diagnosis.

Transplantation-related information:
Stem cell source: Specifies the origin of the cells used for transplantation, such as peripheral blood and bone marrow.
Conditioning therapy: Describes the preparatory therapy administered before transplantation.
Total Body Irradiation (TBI) Status: Indicates whether TBI was incorporated in the recipient's conditioning therapy.
Conditioning therapy type: Specifies the type of conditioning therapy administered (Reduced-Intensity Conditioning, Myeloablative, Non-Myeloablative).
GVHD Prophylaxis: Refers to the preventive medication administered to minimize the risk of GVHD.

Time (month) between diagnosis and transplantation: Indicates the duration between the recipient's diagnosis of the underlying medical condition and the transplantation procedure.

Antifungal therapy: Describes the administration of medications to prevent or treat fungal infections.

Defibrotide prophylaxis: Refers to the use of defibrotide as a preventive measure against hepatic veno-occlusive disease (VOD).

Antimicrobial prophylaxis: The administration of antimicrobial medication to prevent infections in recipients.

G-CSF given to donor: The administration of G-CSF to stimulate the white blood cell production, facilitating stem cell mobilization in donors.

G-CSF given to recipient: The administration of G-CSF to recipients to enhance engraftment in recipients.

Number of apheresis applied to donor: The total number of apheresis procedures conducted to harvest stem cells from the donor for transplantation.

Number of donor lymphocyte infusion: The number of infusions of donor lymphocytes administered post-transplant.

The amount of graft given to the recipient (milliliters): The volume of the transplanted graft received by the recipient.

Engraftment status: Indicates whether the transplanted cells have successfully engrafted within the recipient's bone marrow.

The day the neutrophil count exceeds 500: Specifies the number of days following the transplant until the neutrophil count exceed 500.

The day the neutrophil count exceeds 1000: Specifies the number of days following the transplant until the neutrophil count exceed 1000.

The day the platelet count exceeds 20000: Specifies the number of days following the transplant until the platelet count exceed 20000.

The day the platelet count exceeds 50000: Specifies the number of days following the transplant until the platelet count exceed 50000.

Total Nucleated Cells 10E8: The total count of nucleated cells in the graft.

Mononuclear cells 10E8: The number of mononuclear cells in the graft including lymphocytes and monocytes.

CD34-positive cells 10E6: The number of hematopoietic stem cells in the graft.

CD3-positive cells 10E7: The number of T cells in the graft.

CD3-16-56+ 10E7: The number of NK cells in the graft.

CD3+4+ 10E7: The number of helper T cells in the graft.

CD3+8+ 10E7: The number of cytotoxic T cells in the graft.
CD19+ 10E7: The number of B cells in the graft.

Clinical information:

Viral infection: Indicates whether the recipient has experienced a viral infection including:

- CMV: Cytomegalovirus infection.
- HBV: Hepatitis B Virus infection.
- ADV: Adenovirus infection.
- BK: BK Virus infection.
- Mumps: Mumps Virus infection.
- HSV: Herpes Simplex Virus infection.
- VZV: Varicella-Zoster Virus infection.
- No Viral Infection: Indicates the absence of any viral infection.

HSV infection status: The status of HSV infection history in the recipient.
CMV infection status: The status of CMV infection history in the recipient.
Presence of hemorrhagic cystitis: Indicates whether the recipient has developed hemorrhagic cystitis, a condition characterized by inflammation and bleeding of the bladder lining.
Hemorrhagic cystitis grade: The severity or grade of hemorrhagic cystitis.
Number of relapses: The number of disease relapses experienced by the recipient.
Acute GVHD: This variable covers the grade of acute GVHD from 1 to 4. Additionally, patients without acute GVHD represented as 0.
Acute GVHD severity (1,2 vs. 3,4): Categorizes the severity of acute GVHD into:

- Grade 1-2: Mild to moderate acute GVHD.
- Grades 3-4: Severe to very severe acute GVHD.

Variables to be predicted:

Acute GVHD presence: Indicates whether the recipient has developed acute GVHD post-transplant.
Acute GVHD severity (1 vs. 234): Categorizes the acute GVHD severity into:

- Grade 1: Mild acute GVHD.
- Grades 2-4: Moderate to very severe acute GVHD.

Chronic GVHD presence: Indicates whether the recipient has developed chronic GVHD post-transplant.
Survival status: Indicates the survival status of the recipient (survived or deceased).

The initial dataset encompassed 1313 patients. We obtained survival information for 1306 patients, data about the acute GVHD occurrence for 1132 patients, and data on the chronic GVHD occurrence for 956 patients. Notably, all patients with acute GVHD (n=536) possessed information regarding the severity of their condition.

2.2. Machine learning model development

A detailed overview of our ML development process is shown in Figure 2.1.

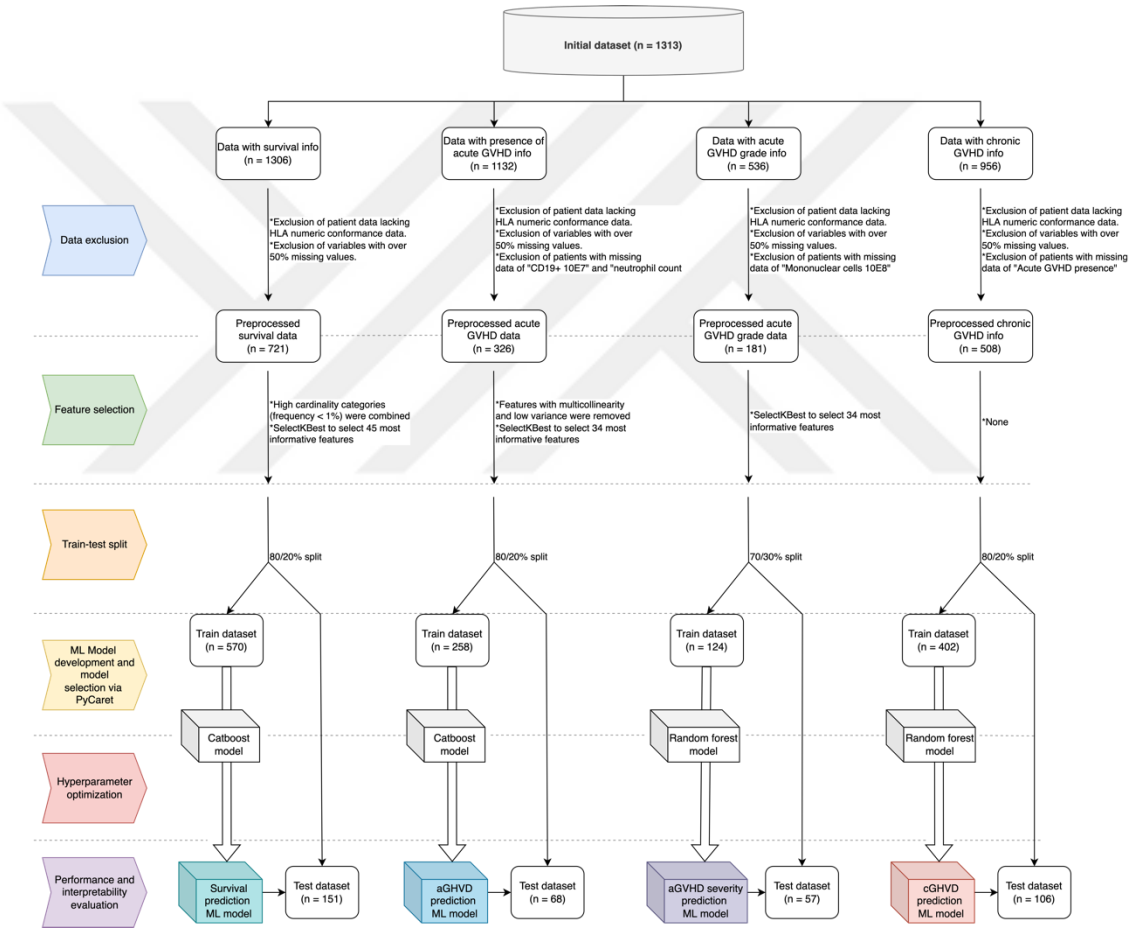

Figure 2.1: Machine learning development process.

2.2.1. Data preprocessing

Prior to model development, data cleaning and preparation were conducted using Pandas 1.3.5 (Team, 2020) and NumPy 1.21.2 (Harris et al., 2020) libraries within the Python programming language. This step ensured data quality and consistency for optimal model performance.

2.2.2. Machine learning library and algorithms considered

PyCaret version 3.0.0, a low-code ML library designed to streamline workflows, was employed for model development (Ali, 2020). This library facilitated the efficient creation, evaluation, and comparison of multiple ML algorithms, enabling the choosing the most appropriate models for our particular prediction tasks.

PyCaret offers a comprehensive suite of classification and regression algorithms. For this study, we explored a broad range of algorithms commonly used for categorical outcome prediction, including:

- Decision Tree Classifiers (Decision Tree, Random Forest)
- Gradient Boosting Classifiers (CatBoost, XGBoost, LightGBM)
- Support Vector Machines (SVM - Linear Kernel)
- Ensemble Classifiers (AdaBoost, Gradient Boosting)
- Statistical Classifiers (Logistic Regression, Naive Bayes)
- Discriminant Analysis Classifiers

2.2.3. Model development for each outcome

We developed four distinct machine learning models, each aiming to predict a different patient outcome (target variable) following allogeneic HSCT:

1. Survival Prediction: This model aimed to forecast recipient survival post-transplant.
2. Acute GVHD Prediction: This model focused on predicting the development of acute GVHD.
3. Acute GVHD Severity Prediction: This model differentiated between mild (grade I) and moderate to very severe (grades ≥ II) forms of acute GVHD.

4. Chronic GVHD Prediction: This model predicted the development of chronic GVHD after transplantation.

2.2.4. Data cleaning and exclusion criteria

To ensure data quality, we implemented a three-step exclusion process for all models:

1. Missing Target Variable: Patients with missing data in the target variable for each respective model were excluded.
2. Missing HLA Conformance Data: Any patient lacking HLA numeric conformance data was removed.
3. High Missing Value Variables: Variables with over 50% missing values were excluded from the analysis.

2.2.5. Specific preprocessing for each model

In addition to the general exclusion process, each model employed further data cleaning steps specific to the target variable and relevant features:

- Survival Prediction: Missing values were handled by applying K-Nearest Neighbors (KNN) imputation for numeric variables and using the mode (most frequently occurring value) for categorical variables.. High cardinality categories (frequency < 1%) were combined into a "rare" category. Feature selection was conducted using scikit-learn's SelectKBest package to identify the 45 most informative features, which was followed by an 80/20 training and testing split.
- Acute GVHD Prediction: Patients with missing "CD19+ 10E7" and "neutrophil count exceeding 1000" data were excluded. Additionally, the "Subtype" variable, with a high proportion (42%) of missing values, was excluded. Missing values were handled by applying KNN for numeric variables and median for categorical variables. Features with multicollinearity and low variance were removed for stability and noise reduction. The data was normalized using Z-score transformation. Feature selection identified 34 features using SelectKBest, which was followed by an 80/20 training and testing split.
- Acute GVHD Severity Prediction: Patients with missing "Mononuclear cells 10E8" data were excluded, along with the "Subtype" variable due to its high (35%) missing value rate. The target variable was "Acute GVHD severity (1 vs. 234)". Preprocessing

steps mirrored those of acute GVHD prediction. Feature selection identified 34 features using SelectKBest, followed by a 70/30 train-test split.

- Chronic GVHD Prediction: Patients with missing "Acute GVHD presence" data were excluded. The target variable was set as "Chronic GVHD presence". Preprocessing followed the same approach as Acute GVHD prediction without feature selection followed by an 80/20 train-test split.

2.2.6. Model selection and hyperparameter optimization

Following data preprocessing and feature selection for each model, the aforementioned ML algorithms' performances were evaluated using PyCaret. The algorithm with the best performance on the testing set was chosen for further training with hyperparameter optimization. This process involves systematically adjusting the model's internal configurations to maximize accuracy of prediction.

2.2.7. Performance evaluation

The performances of all developed models on test datasets were firstly assessed using a confusion matrix that provides an overview of model performance by classifying predictions into true negatives (TN), true positives (TP), false negatives (FN), and false positives (FP). Moreover, receiver operating characteristic (ROC) curves with area under the curve (AUC) values, Kappa values, and metrics such as specificity, sensitivity, positive predictive value, accuracy, and F1 score were used to evaluate performances. The following formulas were used for the calculation of the metrics:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

$$Positive\ Predictive\ Value = \frac{TP}{TP + FP}$$

$$F1\ score\ = \frac{2 \times TP}{2 \times TP + FP + FN}$$

2.2.8. Evaluation of models' interpretability

To better understand the decision-making processes of the models, we used SHapley Additive exPlanations (SHAP) as an interpretability technique. SHAP draws inspiration from game theory to explain how each feature contributes to a specific model prediction. It treats features as players, each contributing to the final prediction. By leveraging the Shapley value concept, SHAP ensures fair credit allocation to each feature by considering all possible feature combinations. It calculates feature importance based on the average influence on the model's prediction (Molnar, 2019).

We utilized SHAP beeswarm plots to visualize the results. These plots use pre-computed SHAP values to illustrate the average influence of each feature on the outcome prediction for every data point. Each data point is represented by a single dot, with the X-axis displaying the model's features. The location of a dot on the X-axis represents the SHAP value of that feature for the corresponding data point. Colors typically range from red (high values) to blue (low values). Features with a broader spread of dots generally have a more substantial influence on predictions across various data points (Lundberg et al., 2020).

2.2.9. The web-based application development

Backend design of the application is summarised below:

Model Integration: Machine learning models were saved as .pkl files during the training phase in PyCaret. The four ML models were stored in a directory (./ML_models) and loaded dynamically using the PyCaret load_model() function based on the user's choice. The application dynamically loads the appropriate model using the @st.cache_resource decorator to optimize performance.

Data Preprocessing: Input features correspond to the models' training data feature sets. Features are either directly entered by the user or derived (e.g., gender mismatch and blood type mismatch). An input dataset serves as a reference for drop-down options and value constraints.

Frontend design of the application is as follows:

Dynamic Forms: Input forms are dynamically generated based on the selected model, with input fields configured to accept numeric inputs, categorical selections, or binary options.
Sidebar: Provides a navigation menu to select models and includes developer information.
Main Page: Displays input forms and prediction results.
Disclaimer: A disclaimer appears at the beginning of the main page as the application is clearly labeled for research use only.

Application workflow is given below:
Model Selection: The sidebar features a radio button interface allowing users to select one of the four predictive models.
Data Input: Forms are generated dynamically to accept input features relevant to the selected model. Input features were predefined based on the models' training datasets, including variables about demographics, transplantation details, laboratory parameters, and clinical outcomes. All input features are validated against training data distributions using drop-downs, number inputs with defined ranges, or predefined options.
Prediction: Upon entry, input data are collected into a Python dictionary and converted to a Pandas DataFrame. The selected model processes the input using PyCaret's predict_model() function, returning the predicted outcome and the associated probability. Prediction output is displayed as an information box, formatted with the prediction outcome and its probability.

The application was built using Streamlit version 1.38.0. The application's code is publicly accessable on GitHub at https://github.com/hikmetc/hsct_outcome. Moveover, the app is available at https://hsct-outcome.streamlit.app .

3. RESULTS

3.1. Performance of survival prediction model

Following the data preprocessing steps outlined previously, the final dataset for the survival prediction model included information from 721 patients. An initial performance comparison was conducted using a subset of 570 patients' data.

The CatBoost model emerged as the best-performing model and was subsequently trained using a cross-validation (10-fold) process. This method involves dividing the training data into 10 folds, using 9 folds are utilized for model training and the remaining part to evaluate its performance. This procedure is repeated 10 times to ensure a robust evaluation across the entire dataset. Additionally, hyperparameter optimization was conducted with 100 iterations, aiming to maximize the model's accuracy in predicting patient survival.

The final performance of the optimized CatBoost model was assessed on a hold-out test set consisting of 151 patients' data. This data was not used during model training or hyperparameter optimization, providing an unbiased assessment of the model's generalizability to unseen data. The performance of the CatBoost model is presented in the confusion matrix (**Figure 3.2**), performance metrics table (**Table 3.1**), and the ROC curve (**Figure 3.3**).

The survival model attained 76.82% accuracy, 66.07% sensitivity, and 83.16% specificity (**Table 3.1**). Additionally, the AUC value of the ROC curve was 0.82 (**Figure 3.3**).

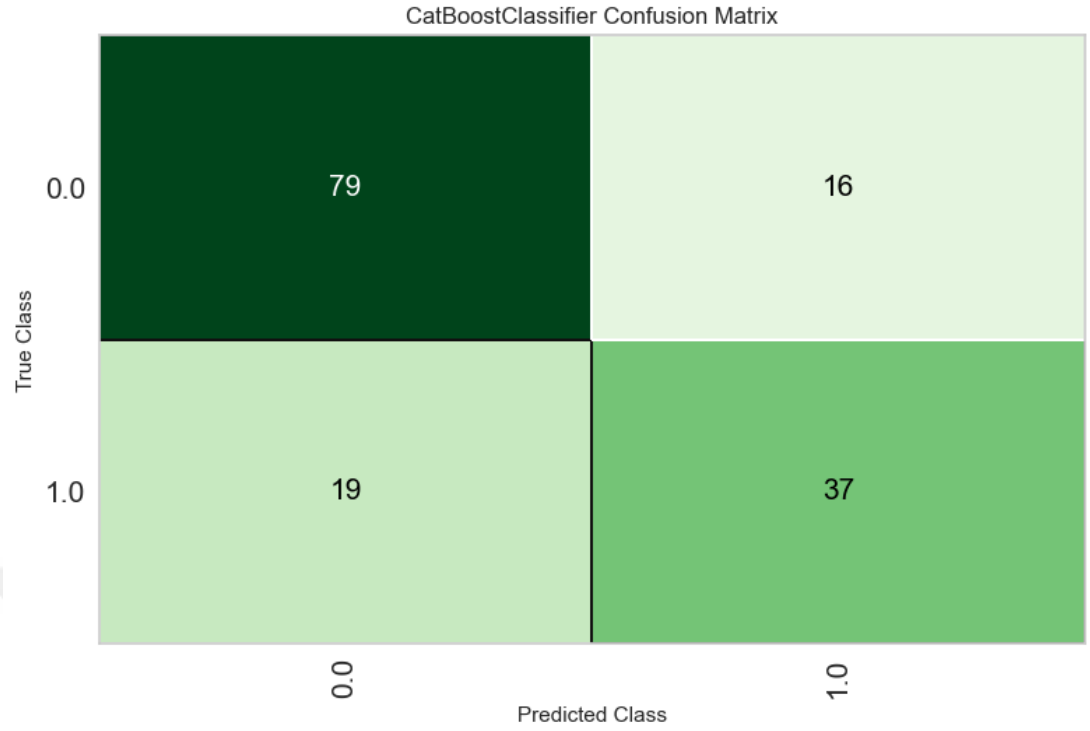

Figure 3.2: Confusion matrix of the CatBoost model for survival prediction. Class 0 and 1 represents deceased and survived individuals, respectively.

Table 3.1: The performance statistics of the CatBoost model for survival prediction. AUC: Area under the curve.

Statistic	**Value**
Sensitivity	66.07%
Specificity	83.16%
Positive Predictive Value	69.81%
Negative Predictive Value	80.61%
Accuracy	76.82%
F1 Score	67.89%
Kappa value	0.50
AUC	0.82

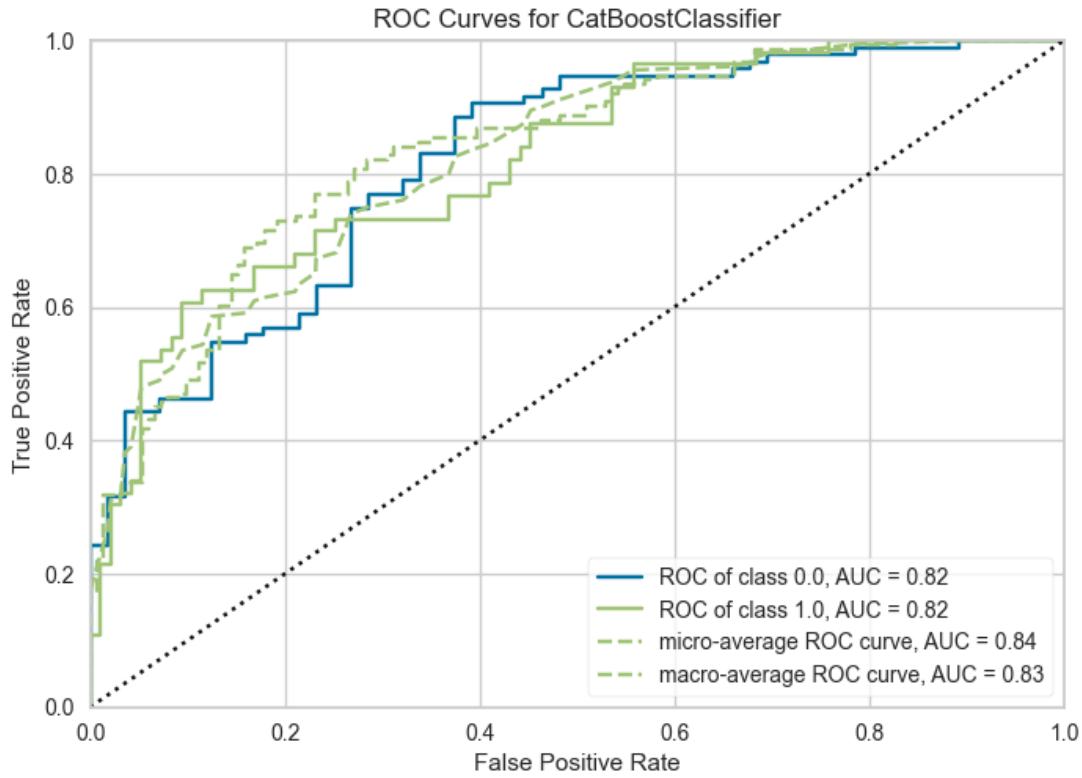

Figure 3.3: The ROC curve of the CatBoost model for survival prediction. Class 0 and 1 represents deceased and survived individuals, respectively.

Positive factors for survival in the survival prediction model (**Figure 3.4**) included positive engraftment, cyclosporine A + methotrexate GVHD prophylaxis, sibling donor and acute myeloid leukemia (AML) diagnosis. Conversely, higher relapse count, higher grades of acute GVHD, higher number of mismatched locus, male recipient gender, higher recipient age, and TBI were linked to lower survival.

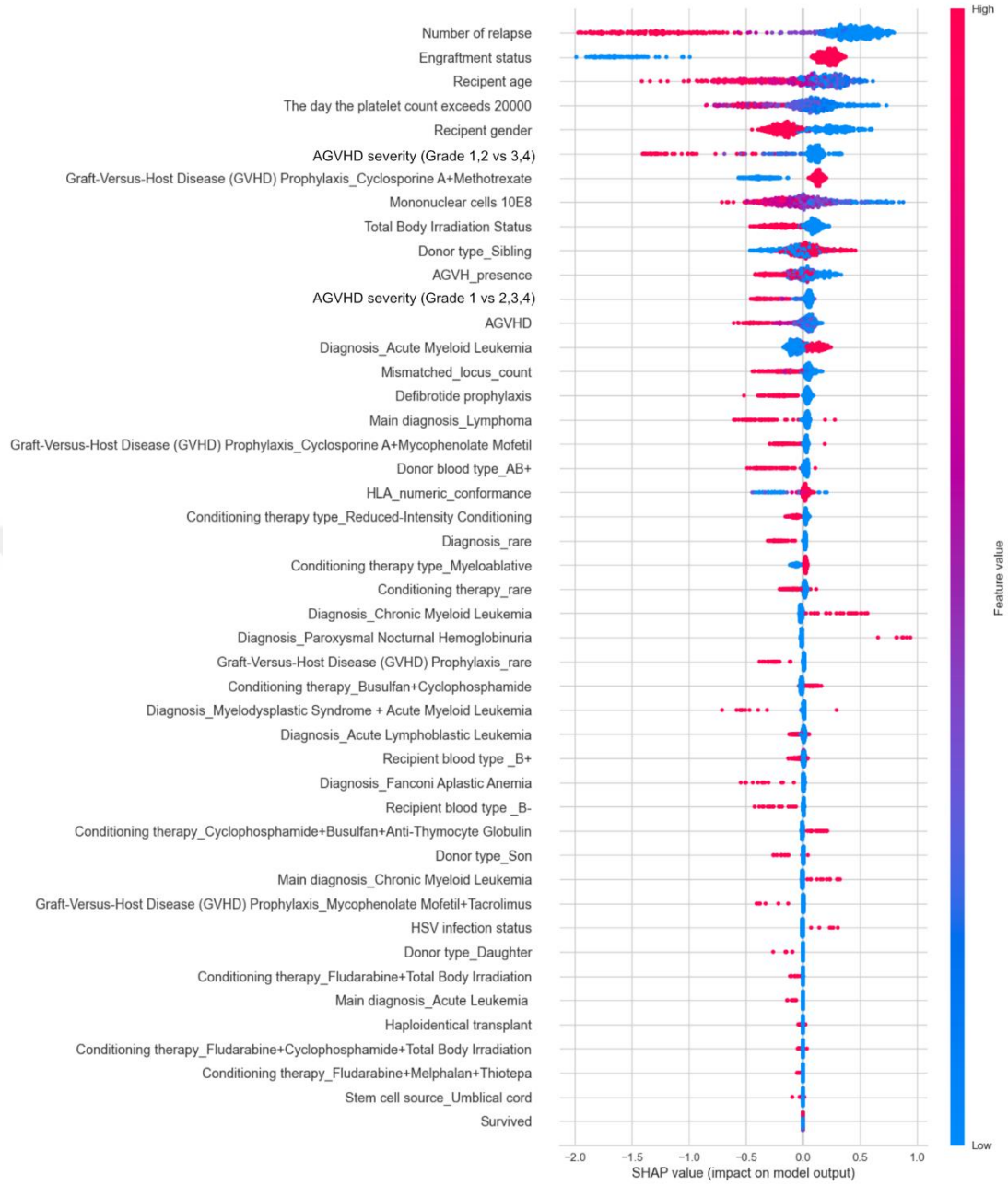

Figure 3.4: SHAP beeswarm plot of the CatBoost model for survival prediction.

3.2. Performance of acute GVHD prediction model

Following the data preprocessing steps outlined previously, the final dataset for the acute GVHD prediction model included information from 326 patients. An initial performance comparison was conducted using a subset of 258 patients' data. The CatBoost model outperformed other models in the initial performance evalutaion. Subsequently, the CatBoost model was further trained using a 10-fold cross-validation process. Hyperparameter

optimization did not result in improved performance. Thus, the initial CatBoost model was selected as the final acute GVHD prediction algorithm.

The final model's performance was assessed on a hold-out test set consisting of 68 patients' data. The performance of the acute GVHD prediction model is presented in the confusion matrix (**Figure 3.5**), performance metrics table (**Table 3.2**) and the ROC curve (**Figure 3.6**). The acute GVHD prediction model showed 64.71% accuracy, 64.71% sensitivity, and 64.71% specificity (**Table 3. 2**). Moreover, the AUC value of the ROC curve's was 0.68 (**Figure 3.6**).

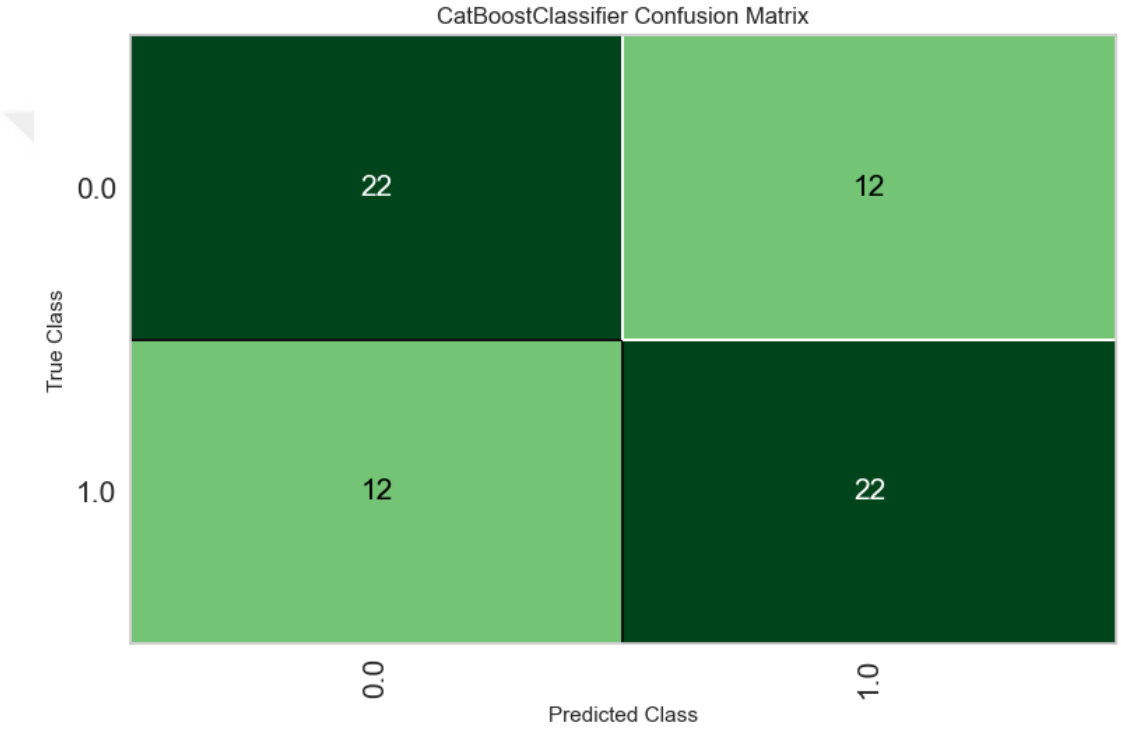

Figure 3.5: Confusion matrix of the CatBoost model for acute GVHD prediction. Class 1 represents patients with acute GVHD.

Table 3.2: The performance statistics of the CatBoost model for acute GVHD prediction. AUC: Area under the curve.

Statistic	**Value**
Sensitivity	64.71%
Specificity	64.71%
Positive Predictive Value	64.71%
Negative Predictive Value	64.71%
Accuracy	64.71%
F1 Score	64.71%
Kappa value	0.29
AUC	0.68

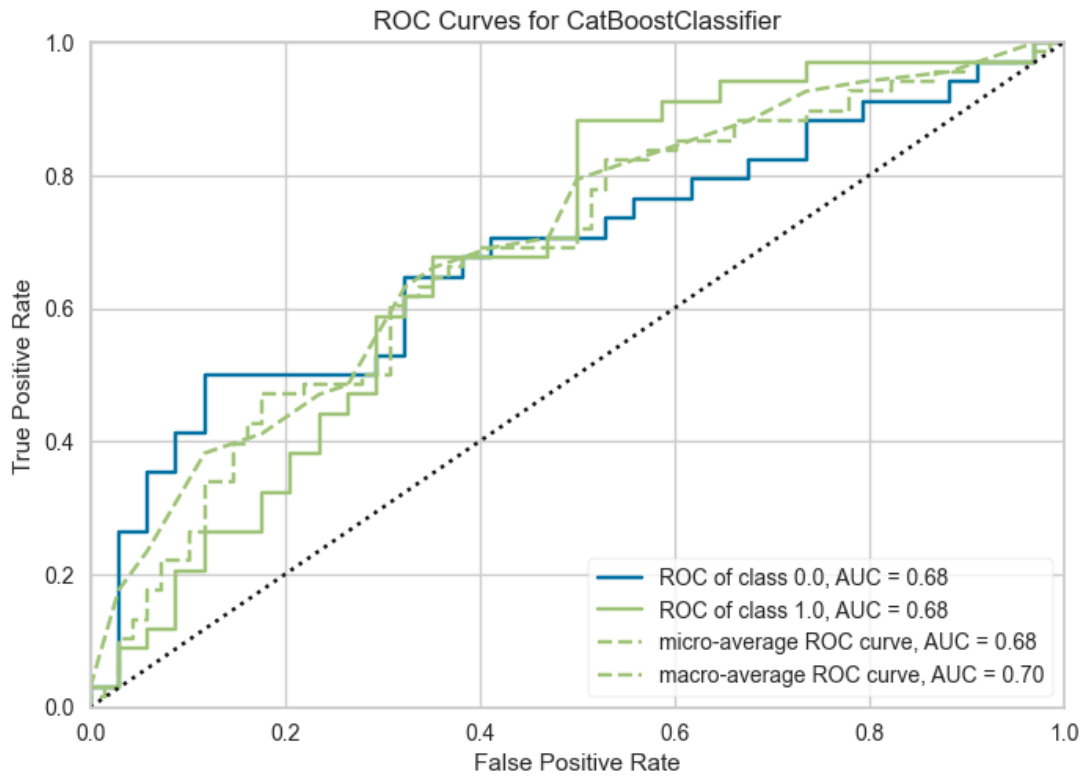

Figure 3.6: The ROC curve of the CatBoost model for acute GVHD prediction. Class 1 represents patients with acute GVHD.

Positive predictors of acute GVHD in the model (**Figure 3.7**) included CMV infection, acute lymphoblastic leukemia diagnosis, donor age, blood type mismatch, conditioning therapy with cyclophosphamide + TBI, and conditioning therapy with busulfan+ cyclophosphamide. Conversely, the amount of graft given to the recipient and the number of CD34 positive cells in the graft were linked to negative acute GVHD status. The number of total nucleated cells in the graft and the day the neutrophil count exceeds 1000 appeared complex based on SHAP analysis, potentially depending on interactions with other factors.

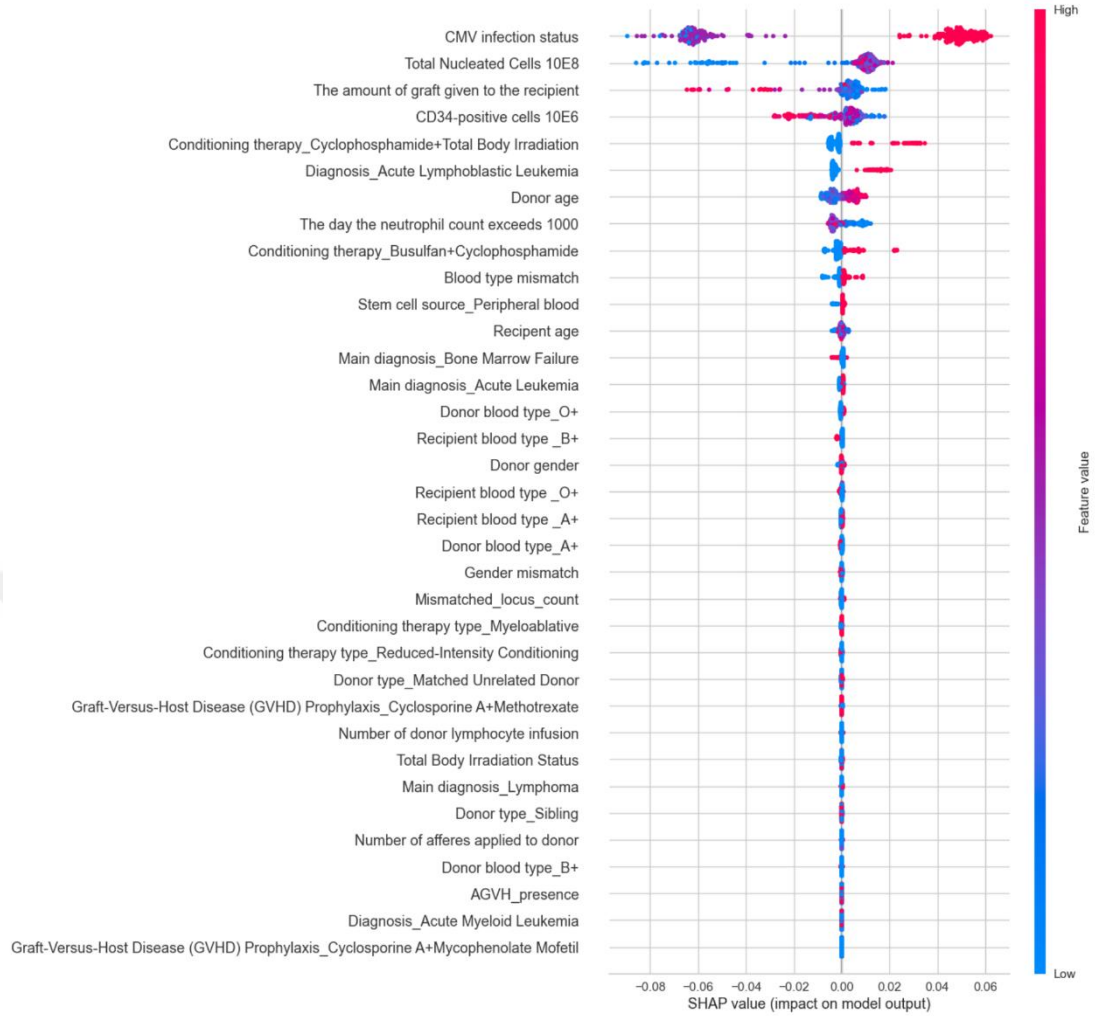

Figure 3.7: SHAP beeswarm plot of the CatBoost model for acute GVHD prediction.

3.3. Performance of acute GVHD severity prediction model

Following the data preprocessing steps outlined previously, the final dataset for the moderate to very severe acute GVHD prediction model included information from 181 patients. An initial performance comparison was conducted using a subset of 124 patients' data. During the initial performance comparison, CatBoost and Random Forest algorithms outperformed the other models. The Random Forest model performed better than the CatBoost model on the test dataset (Accuracy: 73.68% vs. 73.68%, AUC: 0.79 vs. 0.78, Sensitivity: 76.67% vs. 73.33%). Thus, the Random Forest algorithm was selected to be trained using a 10-fold cross-validation process with hyperparameter optimization over 100 iterations to achieve maximum accuracy.

The test set consisting of 57 patients' data was used to assess the final model's performance. The acute GVHD severity prediction model's performance is presented in the confusion matrix (**Figure 3.8**), performance metrics table (**Table 3.3**) and the ROC curve (**Figure 3.9**).

The acute GVHD severity prediction model showed 71.93% accuracy, 76.67% sensitivity, and 66.67% specificity (**Table 3.3**). Moreover, the AUC value of the ROC curve's was 0.78 (**Figure 3.9**).

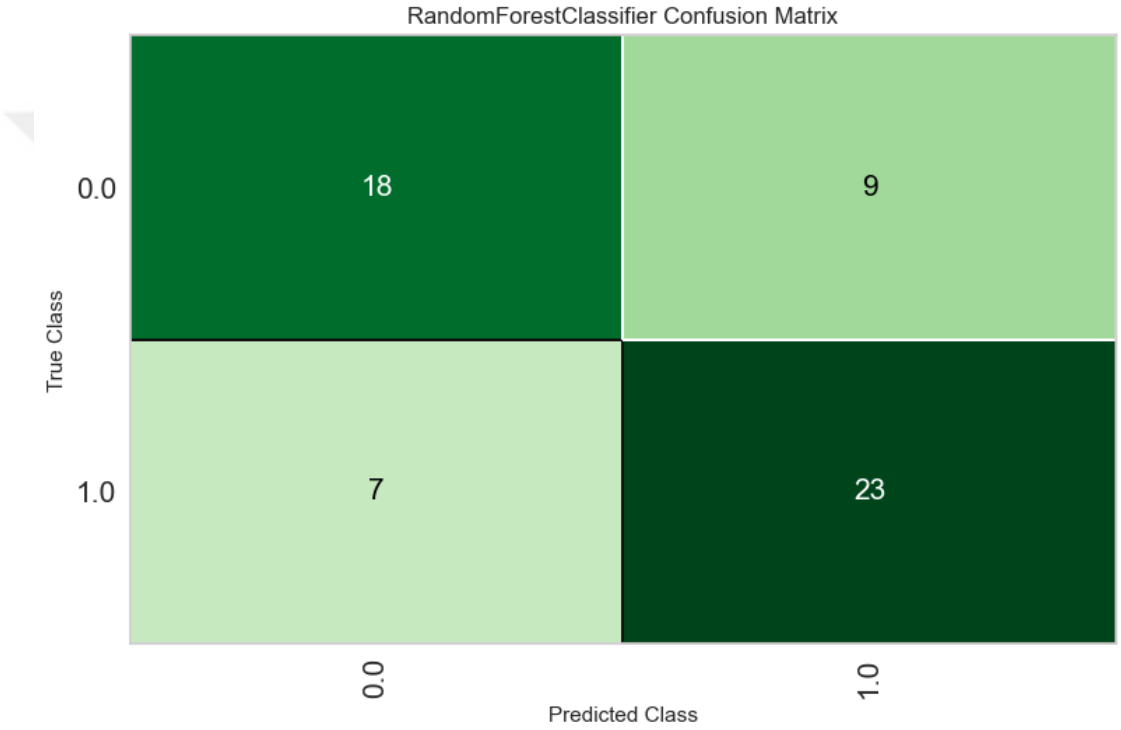

Figure 3.8: Confusion matrix of the Random forest model for acute GVHD severity prediction. Class 1 represents patients with acute GVHD severity higher than 1.

Table 3.3: The performance statistics of the Random forest model for acute GVHD severity prediction. AUC: Area under the curve.

Statistic	**Value**
Sensitivity	76.67%
Specificity	66.67%
Positive Predictive Value	71.88%
Negative Predictive Value	72.00%
Accuracy	71.93%
F1 Score	74.19%
Kappa value	0.43
AUC	0.78

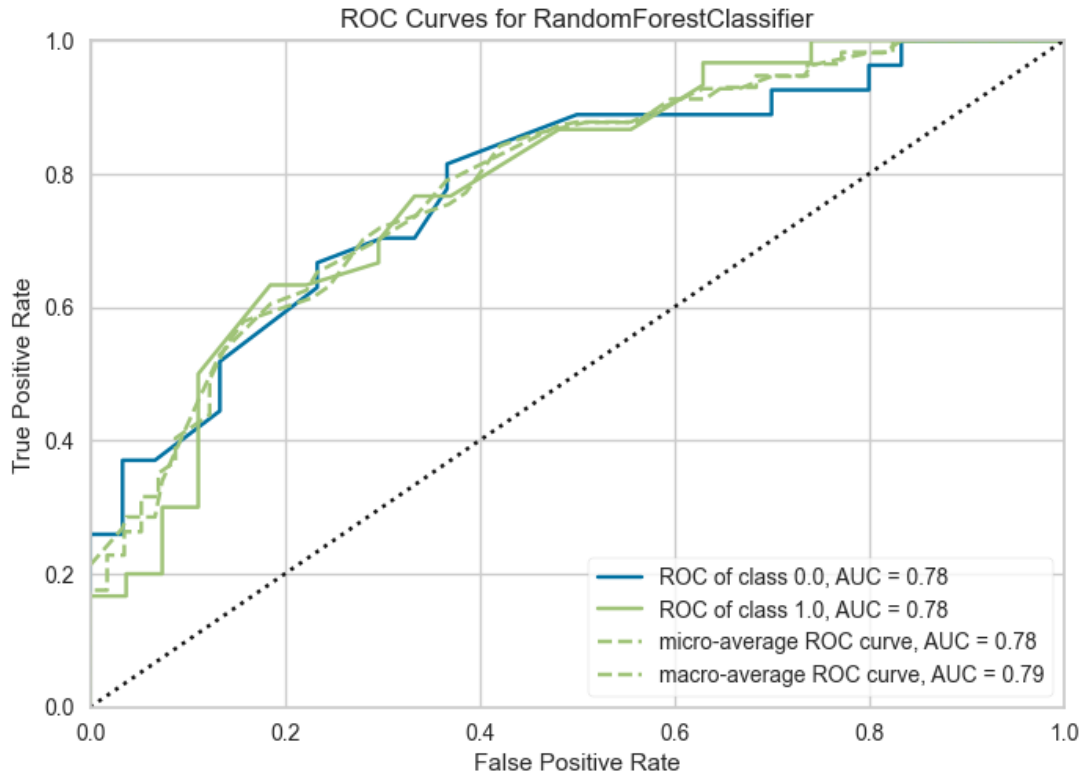

Figure 3.9: The ROC curve of the Random forest model for acute GVHD severity prediction. Class 1 represents patients with acute GVHD severity higher than 1.

Positive predictors of moderate to very severe acute GVHD in the Random Forest model (**Figure 3.10**) included TBI therapy, mismatched locus count, matched unrelated donor, and acute lymphoblastic leukemia (ALL) diagnosis. Conversely, blood type mismatch, donor blood type of O+, and sibling donor were associated with low grade (Grade 1) of acute GVHD. The day the neutrophil count exceeds 1000, mononuclear cells, CD3+ cells, CD3+4+ cells, CD3+8+ cells, total nucleated cells, CD19+ cells, CD34+ cells count in the graft, number of donor lymphocyte infusions and the amount of graft given to the recipient appeared complex based on SHAP analysis, potentially depending on interactions with other factors.

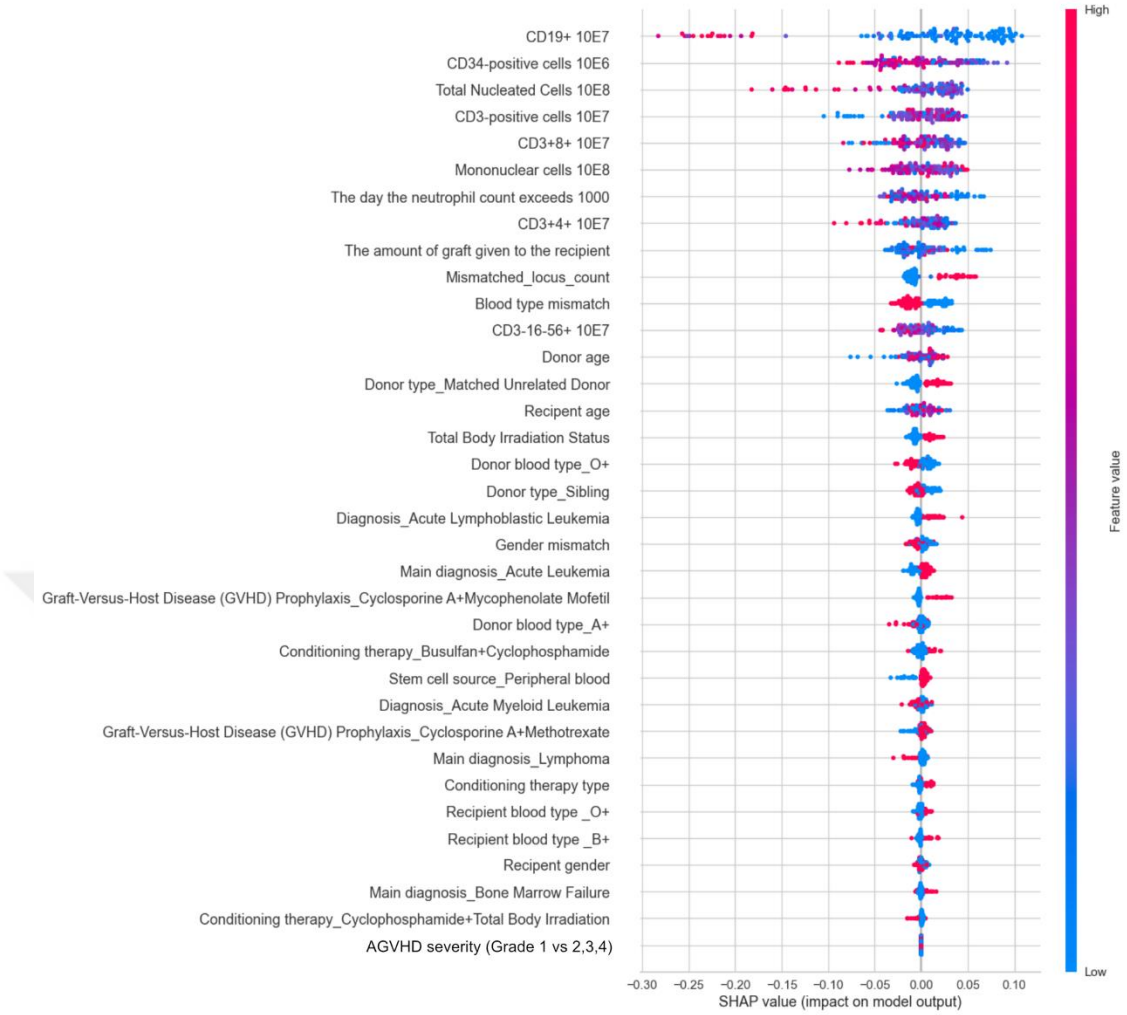

Figure 3.10: SHAP beeswarm plot of the Random forest model for moderate to very severe acute GVHD prediction.

3.4. Performance of chronic GVHD prediction model

Following the data preprocessing steps outlined previously, the final dataset for the chronic GVHD prediction model included information from 508 patients. An initial performance comparison was conducted using a subset of 402 patients' data. During the initial performance comparison, the Random Forest algorithm showed the best performance compared to the others. Therefore, the Random Forest algorithm was selected to be trained with hyperparameter optimization over 100 iterations to achieve maximum accuracy.

The final model's performance was assessed using the test set consisting of 106 patients' data. The chronic GVHD prediction model's performance is presented in the confusion matrix (**Figure 3.11**), the performance metrics table (**Table 3.4**) and the ROC curve (**Figure 3.12**).

The chronic GVHD prediction model showed 78.30% accuracy, 60.00% sensitivity, and 87.32% specificity (**Table 3.4**). Moreover, the AUC value of the ROC curve's was 0.75 (**Figure 3.12**).

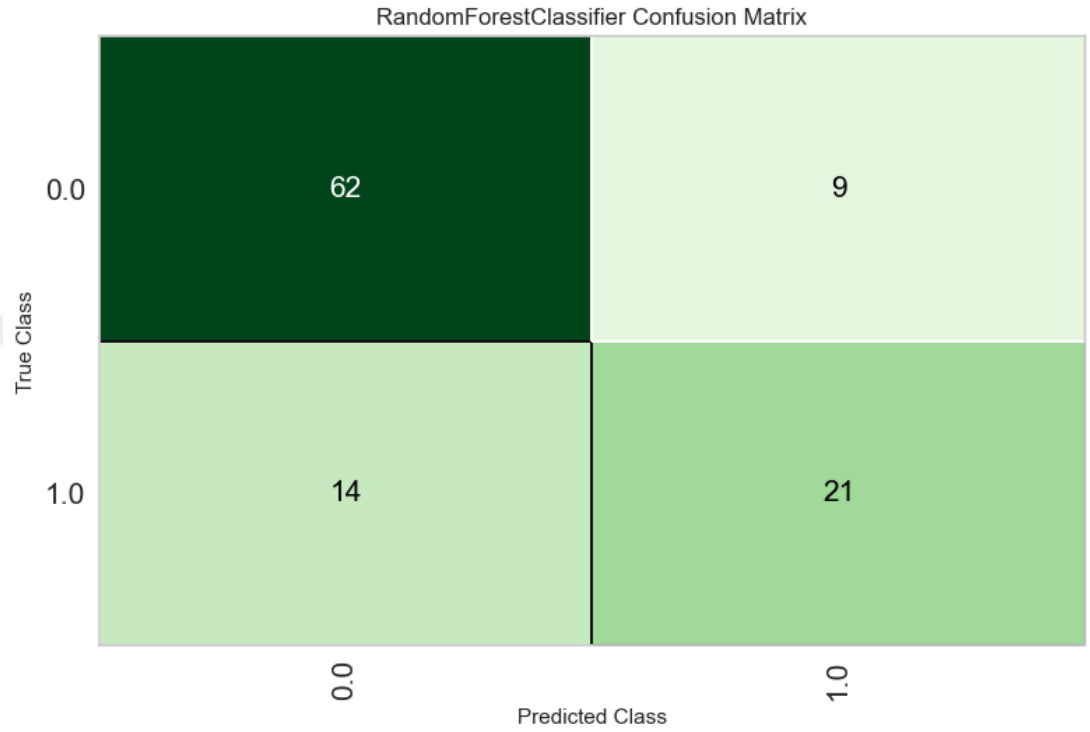

Figure 3.11: Confusion matrix of the Random forest model for chronic GVHD prediction. Class 1 represents patients with chronic GVHD.

Table 3.4: The performance statistics of the Random forest model for chronic GVHD prediction. AUC: Area under the curve.

Statistic	**Value**
Sensitivity	60.00%
Specificity	87.32%
Positive Predictive Value	70.00%
Negative Predictive Value	81.58%
Accuracy	78.30%
F1 Score	64.62%
Kappa value	0.49
AUC	0.75

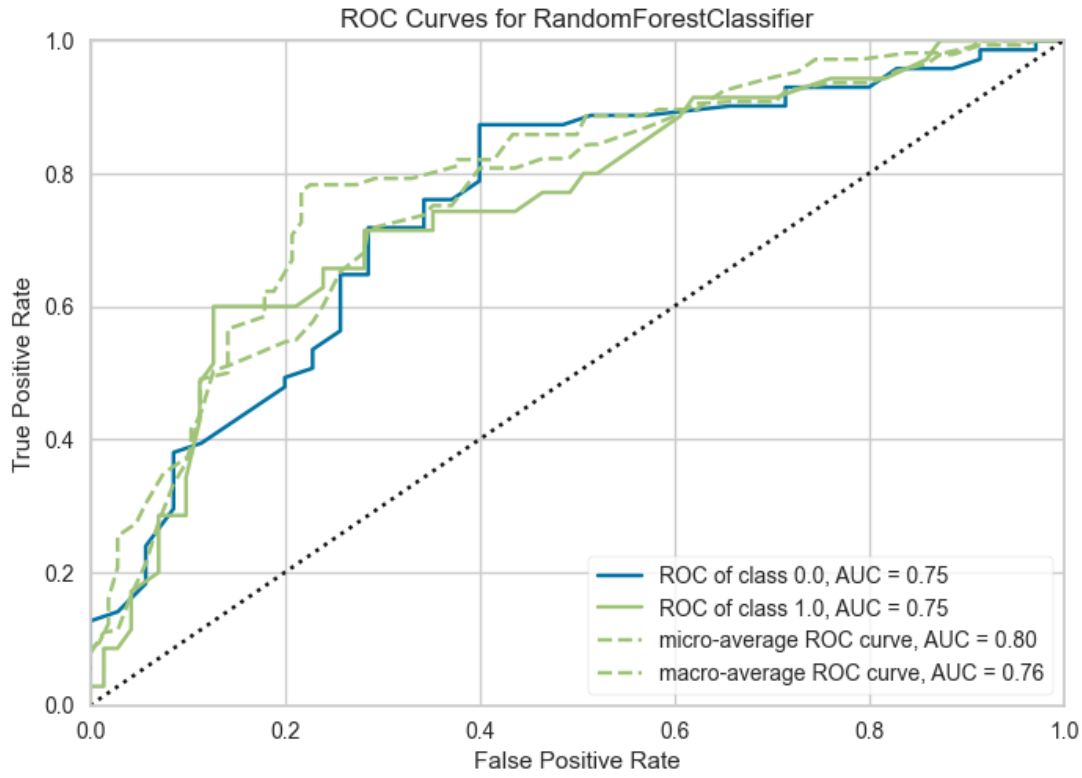

Figure 3.12: The ROC curve of the Random forest model for chronic GVHD prediction. Class 1 represents patients with chronic GVHD.

Positive predictors of chronic GVHD in the Random Forest model (**Figure 3.13**) included acute GVHD (presence and increasing grades), sibling donor, conditioning therapy with busulfan + cyclophosphamide, recipient blood type as A+, myeloablative conditioning therapy, GVHD prophylaxis with cyclosporine A + methotrexate, conditioning therapy with cyclophosphamide + TBI, and female donor gender. Conversely, matched unrelated donor, defibrotide prophylaxis, and reduced-intensity conditioning therapy were negatively linked to chronic GVHD. The amount of graft given to the recipient, the day the neutrophil count exceeds 1000, recipient and donor age appeared complex based on SHAP analysis, potentially depending on interactions with other factors.

Figure 3.13: SHAP beeswarm plot of the Random forest model for chronic GVHD prediction.

3.5. Web-based application

The user interface of the web-based application is illustrated in **Figure 3.14**. Four distinct machine learning (ML) models were integrated into the application for the following purposes: acute GVHD prediction, moderate to very severe acute GVHD prediction, chronic GVHD prediction, and survival prediction. Users can input variables and subsequently click the "predict" button to receive the predicted outcome.

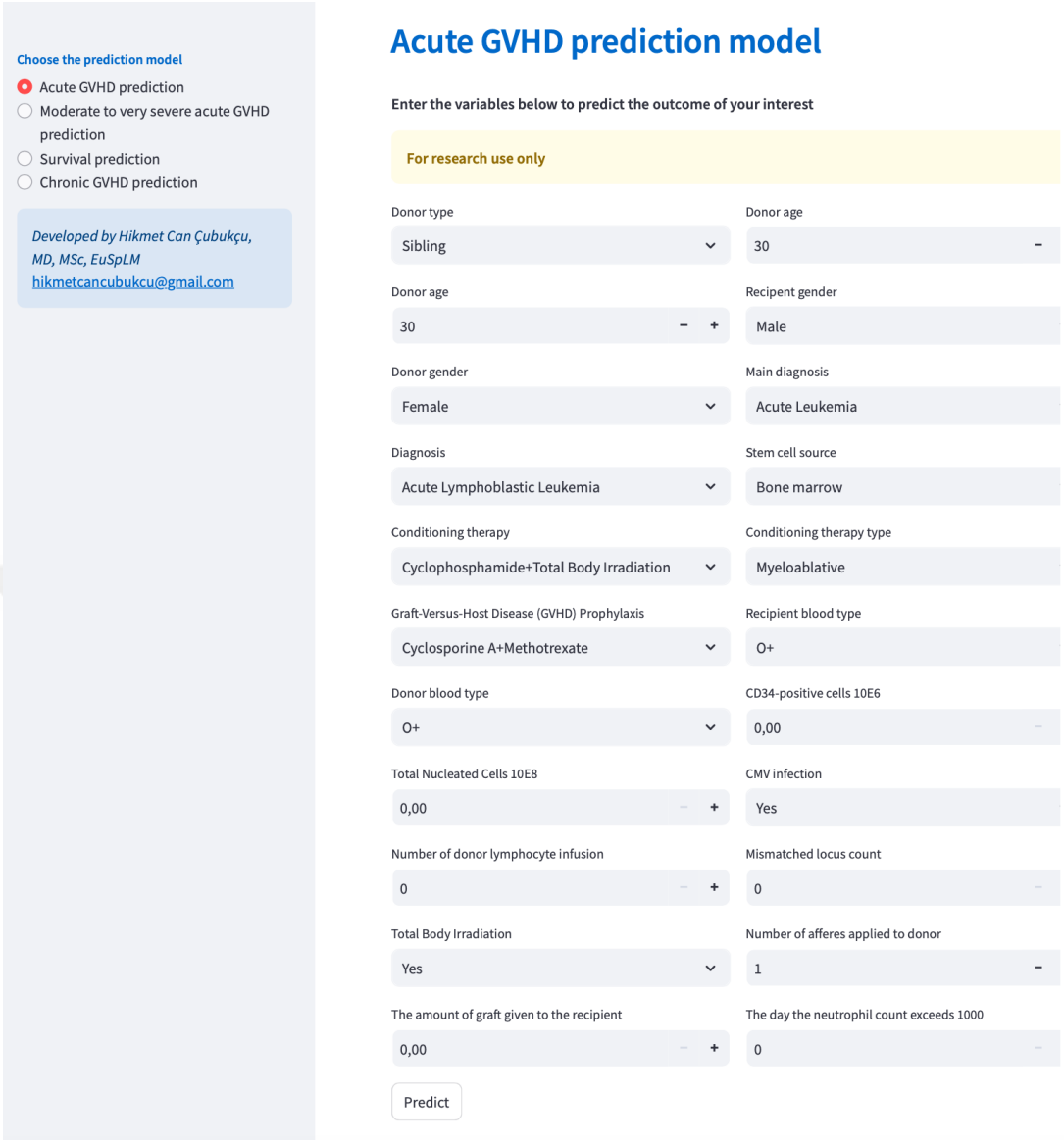

Figure 3.14: The user interface of the web-based application

4. DISCUSSION

This study employed machine learning approaches to develop four prognostic models for allogeneic HSCT outcomes, including survival status, emergence of acute GVHD, acute GVHD severity, and chronic GVHD. Three of the machine learning models (survival, severe acute GVHD, and chronic GVHD prediction models) achieved an AUC exceeding 0.70, indicating acceptable discriminative ability in predicting the respective outcomes, while the acute GVHD prediction model achieved an AUC of 0.68.

4.1. Survival prediction model

Our study's findings provided significant insights into survival factors post-allogeneic HSCT (post-allo-HSCT) using a machine learning-based survival prediction model using CatBoost appoach. This model attained an 76.82% accuracy, 66.07% sensitivity, and 83.16% specificity, with an AUC of 0.82, demonstrating its robustness in predicting survival outcomes. The performance metrics of our survival prediction CatBoost model were comparable to those reported in previous studies utilizing machine learning techniques, although the specific variables used to predict outcomes varied. For instance, Eisenberg et al. (2022) developed a gradient boosting machine (GBM) model with an AUC of 0.92 for predicting mortality post-allo-HSCT, highlighting the efficacy of machine learning in this context (Eisenberg et al., 2022). Their model identified higher C-reactive protein, glutamate oxaloacetate transaminase, blood urea nitrogen, and lower total protein levels as the most predictive features for mortality at day 21 post-transplantation, based on SHAP values (Eisenberg et al., 2022). Similarly, Choi et al. (2022) reported an AUC of 0.79 for their gradient boosting model for survival prediction, indicating comparable predictive capabilities to our survival prediction CatBoost model (Choi et al., 2022). While relapse was determined as a major factor for long-term survival, engraftment status, GVHD, and treatment-related toxicities were also found to be associated with survival. Conditioning regimen, HLA-matching status, disease status, donor and patient ages, and genetic risk were associated with disease relapse (Choi et al., 2022). Moreover, Shouval et al. (2016) employed six ML algorithms, including random forest, logistic regression, alternating decision trees, multilayer perceptron, and ADABoost, to predict transplantation-related mortality (Shouval et al., 2016). Their models achieved an AUC of 0.67, with donor type, disease stage, and conditioning therapy identified as predictive factors for non-relapse mortality (Shouval et al., 2016).

Predictors of survival prediction model:

Our study identified positive engraftment, cyclosporine A + methotrexate GVHD prophylaxis, sibling donor, and an AML diagnosis as factors associated with improved survival. Graft failure is well known complication that leads to life treatening effects (Ozdemir & Civriz Bozdağ, 2018). Similar to our study's findings, Choi et al. showed that positive engraftment linked to post-allogeneic-HSCT survival (Choi et al., 2022). Thus, positive engraftment has been consistently linked to improved outcomes in allogeneic-HSCT.

Unlike our findings, Harashima et al. (2019) observed that a prophylaxis regimen including cyclosporine A, rather than one with tacrolimus, was linked to poorer post-allogeneic-HSCT overall survival (Harashima et al., 2019). Similarly, Kamel et al. (2016) reported lower acute GVHD rates in renal transplant patients treated with tacrolimus (Kamel et al., 2016). Moreover, recent evidence has favored the use of cyclophosphamide-containing regimens over other immunosuppressive drug combinations, demonstrating improved GVHD and relapse-free survival (Bolaños-Meade et al., 2023; Ruggeri et al., 2017).

Nevertheless, as shown in Figure 3.3, our SHAP analysis did not reveal any predictive potential of GVHD prophylaxis regimens including tacrolimus or cyclophosphamide. Thus, our findings could not infer any positive effect of regimens with tacrolimus or cyclophosphamide on survival rates. While, cyclosporin and tacrolimus agents are both calcineurin inhibitors (Lee et al., 2023), tacrolimus is often preferred for patients experiencing rejection while on cyclosporine treatment (Lee et al., 2023). Patients with such a more severe clinical condition in the tacrolimus prophylaxis group may have poorer survival rates, potentially supporting our study's findings.

On the other hand, our retrospective dataset, spanning patient medical records from 1988 to 2023, encompasses a period marked by significant evolution in GVHD prophylaxis regimens. Notably, only 73 of the 570 patients in the survival prediction sub-dataset received a post-transplant cyclophosphamide regimen. This limited usage of post-transplant cyclophosphamide restricted our ability to definitively evaluate its effectiveness across the entire dataset. Consequently, its impact on patient outcomes could not be accurately assessed in this context.

Regarding the unexpected positive association between AML diagnosis and survival, it is plausible that this result may be influenced by the inclusion of other hematological malignancies with poorer prognoses in the study population.

The SHAP analysis (Figure 3.4) identified sibling donors as being associated with improved survival outcomes. This finding is consistent with current clinical practice, which prioritizes matched sibling donors over other donor types due to their superior HLA compatibility and associated higher survival rates (Lee et al., 2007).

Conversely, higher relapse count, higher grades of acute GVHD, a higher number of mismatched loci, male recipient gender, older recipient ages, and TBI were linked to reduced survival rates.

The detrimental influence of higher relapse count has been widely recognized. Choi et al. and Sayehmiri et al. emphasized relapse as a critical factor negatively influencing long-term (Choi et al., 2022) and overall survival (Sayehmiri et al., 2008), respectively. Additionally, some researchers identified disease relapse as a major factor of mortality (Wingard et al., 2011; Wu et al., 2023).

Higher grades of acute GVHD are strongly associated with reduced survival, as corroborated by Wu et al. and Ramdial et al., who identified grade III-IV acute GVHD as significant predictors of lower overall survival (Wu et al., 2023) and non-relapse mortality (Ramdial et al., 2021), respectively.

Our finding of a higher number of mismatched loci affecting survival aligns with Garcia-Horton et al. (2022), who demonstrated that HLA mismatch status significantly impacts overall survival (Garcia-Horton et al., 2022).

Additionally, the gender disparity observed, with male recipient gender being linked to lower survival, is consistent with findings from Wong et al. (2020) and Solh et al. (2018), who identified male gender as a contributing factor to mortality (Solh et al., 2018; Wong et al., 2020). Moreover, the pairing of a female donor with a male recipient is incorporated in the EBMT risk score, which is related to non-relapse mortality (Gratwohl, 2012).

Our SHAP analysis revealed that higher recipient age linked with lower survival. These findings align with the EBMT risk score, which incorporates recipient age as a risk factor (Gratwohl, 2012). Older recipient age has been consistently linked to decreased post-transplant survival (Garcia-Horton et al., 2022; Harada et al., 2021; MacMillan et al., 2012; Shaw et al., 2024; Wingard et al., 2011). This relationship is likely influenced by the fact that advanced age is often accompanied by multiple comorbidities that negatively impact survival outcomes.

We observed that TBI is linked to lower survival rates. TBI is utilized as part of the conditioning therapy. While effective in eliminating hematopoietic cells within the bone marrow, TBI is linked to significant toxicities, including pulmonary complications such as radiation-induced interstitial pneumonitis, hepatic complications, renal impairment, and VOD (Wong et al., 2018). Furthermore, it has been tied to a higher probability of acute GVHD, a major cause of complications after allogeneic HSCT (Przepiorka et al., 1999).

4.2.Acute GVHD and moderate to very severe acute GVHD prediction models

In our study, we developed two distinct models to predict acute GVHD. The first model, a CatBoost classifier, predicted the occurrence of acute GVHD with a 64.71% accuracy and an AUC of 0.68. The second model, using a Random Forest algorithm, predicted the emergence of moderate to very severe acute GVHD with 71.93% accuracy, 76.67% sensitivity, 66.67% specificity, and an AUC of 0.78. Our models outperform most of the previously published models in terms of accuracy and AUC. Compared to the ML classifier developed by Chandra et al. (2024), which achieved 58% accuracy and 0.59 AUC for acute GVHD grade prediction (Grade 1 vs 2,3,4) (Chandra et al., 2024), our moderate to very severe acute GVHD prediction model demonstrated superior performance (Accuracy: 71.93%, AUC: 0.78). Similarly, Tang et al. developed a model to predict grade ≥ II acute GVHD using the demographical, clinical and vital features. The model's AUC value was 0.659 which was lower than our model's AUC value (0.78) (Tang et al., 2020). Likewise, Arai et al. built an alternating decision tree model using clinical and demographical features (HLA mismatch status, donor age, use of irradiation, underlying disease, and time from diagnosis to HSCT) to predict grade ≥ II acute GVHD. The model achieved AUC of 0.616 which was also lower than our model's AUC value (0.78) (Arai et al., 2019). However, Salehnasab et al. (2021) built a model (a extreme gradient boositing classifier using clinical, demographical, and laboratory features) with seemingly superior performance (accuracy: 90.70%, AUC: 0.91). It's important to

consider the limitations of their study. Their model development utilized data from only 182 patients, which might negatively impact its generalizability (Salehnasab et al., 2021).

Predictors of acute GVHD and moderate to very severe acute GVHD prediction model:

Our acute GVHD prediction model identified CMV infection, ALL diagnosis, donor age, blood type mismatch, a cyclophosphamide plus TBI conditioning therapy, and conditioning therapy with busulfan+ cyclophosphamide as positive predictors of acute GVHD. Conversely, the amount of graft given to the recipient and the count of CD34 positive cells in the graft were negative predictors of acute GVHD.

In the Random Forest model, positive predictors of moderate to very severe acute GVHD were TBI therapy, a higher number of mismatched loci, matched unrelated donor, and acute lymphoblastic leukemia diagnosis. Conversely, blood type mismatch, donor blood type of O+, and sibling donor were negative predictors of moderate to very severe acute GVHD.

CMV infection has been established as a contributing factor for developing GVHD (Gale et al., 1987). Reactivation of latent CMV infection following allogeneic HSCT is linked to increased mortality (Akahoshi et al., 2022). In alignment with these findings, our study identified CMV infection history confirmed by serological tests as a significant positive predictor of acute GVHD. While CMV infection did not emerge as a negative predictor of survival in our survival prediction model, its established association with GVHD underscores the importance of CMV monitoring and prophylaxis in this patient population.

The diagnosis of ALL emerged as a positive predictor of both acute and moderate to very severe GVHD in our models. This finding is unexpected, as while ALL is associated with lower survival rates (Jagasia et al., 2012), a direct link to GVHD has not been previously established (Lee et al., 2018). Furthermore, we did not identify any relationship between an ALL diagnosis and survival in our study. This apparent discrepancy between our findings on GVHD and survival, coupled with the absence of supporting evidence in the current literature, underscores the need for further investigation to elucidate the complex interplay between ALL and GVHD.

While blood type mismatch is considered as a contributor to acute GVHD (Kimura et al., 2008; Ludajic et al., 2009), the specific role of individual blood types in this relationship

remains unclear. Recent studies have challenged this notion, reporting no significant association between blood type mismatch and GVHD occurrence (Ciftciler et al., 2020; Damodar et al., 2017). Contrary to these findings, our analysis identified blood type mismatch as a positive predictor of GVHD, while also identifying blood type mismatch and donor blood type O+ as negative predictors of moderate-to-very-severe acute GVHD. These complex and seemingly contradictory results emphasize the necessity for further investigation.

Increasing doses of TBI have been linked to an increased acute GVHD incidence (Przepiorka et al., 1999). Moreover, the timing of TBI administration within the conditioning regimen has been determined as a risk factor; patients receiving TBI prior to chemotherapy have a greater risk of grade ≥ II GVHD in comparison to those receiving TBI post-chemotherapy (Kato et al., 2014). Additionally, conditioning regimens combining TBI and cyclophosphamide have been linked to higher rates of grade ≥ II GVHD in comparison to busulfan-cyclophosphamide regimens in individuals with CML (Kröger et al., 2001). Consistent with previous findings (Giebel et al., 2023), we identified conditioning therapy with cyclophosphamide + TBI as a positive predictor of acute GVHD. Interestingly, our analysis also identified conditioning therapy with busulfan + cyclophosphamide as a positive predictor of acute GVHD. Furthermore, TBI administration was found to be a positive predictor of moderate-to-very-severe (grade 2-4) acute GVHD in our study. These results highlight the complexity of conditioning regimens in influencing GVHD risk..

Eisner and August (1995) showed that higher donor age was linked to a higher acute GVHD incidence (Eisner & August, 1995). Similarly, Di Francesco et al. (2023) identified donor age greater than 40 as a variable linked to grade ≥ II acute GVHD (Di Francesco et al., 2023). Consistent with these findings, our analysis also identified donor age as a positive predictor for the emergence of acute GVHD. However, SHAP analysis revealed a more complex relationship between higher donor age and the emergence of grade 2-4 acute GVHD, suggesting that its impact may depend on interactions with other factors (Figure 3.10). These results highlight the multifactorial nature of donor age as a contributing factor to acute GVHD.

HLA matching is a primary determinant of acute GVHD following allogeneic HSCT (Kanda, 2013). Previous research demonstrated an association between the number of mismatched loci and grade 3-4 acute GVHD (Kawase et al., 2007). Aligned with these findings, our study identified the number of mismatched loci as a positive predictor of grade 2-4 GVHD.

Our findings indicate that matched unrelated donors are linked to more severe acute GVHD compared to sibling donors, 95% of whom in our dataset are matched sibling donors. Similarly, Battipaglia et al. (2018) reported lower GVHD/relapse-free survival and a greater occurances of severe acute GVHD in AML patients transplanted with matched unrelated donors' grafts in comparison to those receiving transplants from matched sibling donors. This difference was attributed to the presence of minor histocompatibility antigens (Battipaglia et al., 2018). Furthermore, Lu et al. (2021) observed that AML patients receiving grafts from matched unrelated donors faced higher rates of grade ≥ II and grade ≥ III acute GVHD compared to those with matched sibling donors (Lu et al., 2021). Minor antigen mismatches can persist despite HLA compatibility between unrelated donors and recipients. These minor mismatches may trigger stronger immune responses. In contrast, matched siblings are more likely to possess identical minor histocompatibility antigens (Martin et al., 2017), thereby reducing the risk of T cell-mediated immune responses targeting recipient tissues.

We observed that both the total graft volume administered to the recipient and the CD34-positive cells in the transplant were negative predictors of acute GVHD. Remberger et al. showed that greater doses of CD34-positive cell transplantation were linked to reduced acute GVHD incidence. This effect was attributed to the higher count of regulatory T cells within the graft, which modulate the T-cell response (Remberger et al., 2020). In our study, the higher graft volume administered to recipients likely correlated with a higher count of CD34-positive cells, which can explain our findings in light of the results reported in the previous study.

4.3. Chronic GVHD prediction model

Our chronic GVHD prediction model achieved 78.30% accuracy, 60.00% sensitivity, 87.32% specificity, and an AUC of 0.75 (Table 3.4, Figure 3.12). Avni et al. developed a multivariate logistic regression model incorporating features such as alanine transaminase and circulating cell-free DNA (cfDNA) levels, which attained an AUC of 0.8, 86% specificity, and 89% positive predictive value for chronic GVHD prediction (Avni et al., 2024). While the latter model demonstrated superior performance, its reliance on cfDNA, which may not be widely accessible in clinical settings, limits its broader applicability.

Predictors of chronic GVHD prediction model:

Positive predictors of chronic GVHD in the Random Forest model (Figure 3.13) included sibling donor, busulfan-cyclophosphamide and cyclophosphamide-TBI conditioning therapies, the occurance and severity of acute GVHD, cyclosporine A + methotrexate GVHD prophylaxis, myeloablative conditioning therapy, female donor gender, and recipient blood type A+. Conversely, defibrotide prophylaxis, matched unrelated donor, and reduced-intensity conditioning therapy were negatively associated with chronic GVHD.

Busulfan-cyclophosphamide conditioning emerged as a positive predictor of chronic GVHD in this study (Figure 3.13). These findings align with those of Uzay et al., who observed a rise in the rates of chronic GVHD in individuals receiving busulfan-based conditioning regimens compared to those undergoing treosulfan, attributing this to a potential post-inflammatory effect of busulfan (Uzay et al., 2024).

In our study, conditioning regimen using cyclophosphamide and TBI emerged as a positive predictor of chronic GVHD. This finding contrasts with the observations made by Giebel et al., who showed comparable chronic GVHD incidence between cyclophosphamide-TBI and fludarabine-TBI conditioning therapies (Giebel et al., 2023). Notably, the fludarabine-TBI conditioning regimen did not emerge as a predictive variable for chronic GVHD in our analysis. Furthermore, a direct comparison between these two conditioning regimens was not conducted in our study, highlighting the need for further investigation to elucidate the underlying reasons for these findings.

In our study, cyclosporine A + methotrexate prophylaxis emerged as a positive predictor of chronic GVHD. While Nagler et al. reported comparable chronic GVHD incidence between post-transplantation cyclophosphamide and cyclosporine A + methotrexate prophylaxis (Nagler et al., 2022), contrasting results were presented by Shouman et al., who observed higher rates of chronic GVHD in individuals receiving cyclophosphamide + cyclosporine A prophylaxis compared to cyclosporine A + methotrexate prophylaxis (Shouman et al., 2023). However, our study did not include direct comparisons between specific prophylaxis regimens, restricting our capability to make definitive determinations about the factors contributing to these discrepancies. Further research is warranted to clarify the influence of prophylaxis regimens on chronic GVHD risk.

In our study, myeloablative conditioning therapy emerged as a positive predictor of chronic GVHD, whereas reduced-intensity conditioning therapy was identified as a negative predictor. These findings align with the results of Couriel et al. (2004), who documented a higher occurance of chronic GVHD in individuals receiving myeloablative conditioning in comparison to those undergoing non-myeloablative therapies (Couriel et al., 2004). Conversely, Afram et al. (2018) demonstrated an increased chronic GVHD risk linked to reduced-intensity conditioning therapy (Afram et al., 2018). These results also contrast with those of Pérez-Simón et al. (2005), who found no substantial difference in chronic GVHD incidence between myeloablative and reduced-intensity regimens (Pérez-Simón et al., 2005). The discrepancy between our findings and those of previous studies can be attributed to variations in patient populations or study designs, warranting further investigation.

We observed that sibling donors were a positive predictor of chronic GVHD, while matched unrelated donors were a negative predictor. This finding is intriguing, and Remberger et al. (2016) attributed a similar observation to the use of anti-thymocyte globulin in unrelated donor transplants, which depletes T cells within the graft (Remberger et al., 2016). Nevertheless, in our study, the relationship between type of donor and chronic GVHD might be influenced by other factors. In our study, sibling donor transplants were more frequently associated with myeloablative conditioning regimens (79.6% vs. 66.8%; $p < 0.001$), cyclosporine A + methotrexate prophylaxis (80.8% vs. 72.7%; $p < 0.001$), and female donors (45.5% vs. 34.3%; $p < 0.001$) compared to matched unrelated donors. All these factors were identified as positive predictors of chronic GVHD in our analysis (Figure 3.13). In contrast, matched unrelated donor transplants were more commonly associated with defibrotide prophylaxis (41.7% vs. 20.2%; $p < 0.001$) compared to sibling donors, and defibrotide prophylaxis was identified as a negative predictor of chronic GVHD (Figure 3.13). These findings highlight the complex interplay of donor type, conditioning regimens, prophylaxis strategies, and donor characteristics in influencing chronic GVHD risk, warranting further investigation.

In our study, prior acute GVHD and grade >1 acute GVHD emerged as positive predictors of chronic GVHD. These findings align with previous study demonstrating a profound connection between acute and chronic GVHD. Ozawa and colleagues emphasized that a prior acute GVHD significantly increases the chronic GVHD risk, with a greater likelihood for grade > 1 compared to grade 1 (Ozawa et al., 2007). Furthermore, Arora et al. found that

children who experienced acute GVHD were more likely to face with chronic GVHD after receiving a haploidentical HSCT (Arora et al., 2024).

Our study unexpectedly found that recipients with blood type A+ had a greater risk of developing chronic GVHD. This finding diverges from the established knowledge about the relationship between blood type and chronic GVHD, as no prior studies have reported such an association. Furthermore, researchers like Damodar et al. (2017) found no significant link between blood type mismatch and the occurrence of chronic GVHD (Damodar et al., 2017). The underlying mechanisms for this novel observation remain unclear and warrant further investigation.

Defibrotide, with its antithrombotic effects, utilized to treat VOD, a complication of allogeneic HSCT (Richardson et al., 2018). While studies in pediatric HSCT patients have shown a reduced acute GVHD incidence with defibrotide prophylaxis in comparison to controls, no such protective effect was observed for chronic GVHD (Corbacioglu et al., 2012; Squillaci et al., 2023). This aligns with the understanding that defibrotide's endothelial protective and anti-inflammatory effects primarily mitigate acute GVHD (Corbacioglu et al., 2012). Interestingly, despite the lack of prior evidence supporting its impact on chronic GVHD incidence, our study identified defibrotide prophylaxis as a negative predictor of chronic GVHD. This unexpected finding requires further investigation to explore the potential mechanisms through which defibrotide may influence chronic GVHD risk and its role in prevention strategies.

In our study, female donor gender was identified as a positive predictor of chronic GVHD. This finding aligns partially with previous research by Kanda et al., who showed increased chronic GVHD incidence in transplantations where a female donor provided grafts to male recipients, relative to male donor-female recipient pairings (Kanda et al., 2014). While our SHAP analysis highlighted female donor gender as a positive predictor, the absence of a significant impact of gender mismatch in our model prevents drawing definitive conclusions about the role of specific gender combinations in chronic GVHD development. More investigation is warranted to elucidate the complex interplay between donor and recipient gender in this context.

We additionally developed a proof-of-concept web-based application (https://hsct-outcome.streamlit.app) that allows users to input variables and obtain predicted outcomes.

Currently, this application is intended for research use only. However, following training with larger datasets and further external validation, it has the potential to become a valuable clinical decision support tool.

4.4. Limitations

The generalizability of our machine learning models is contingent upon external validation, which involves assessing model performance on an independent dataset from a different population or institution. Due to the absence of an external validation dataset in the present study, the generalizability of our models remains to be established. Future research should prioritize external validation to enhance the model's applicability.

Additionally, the presence of missing data in our allo-HSCT dataset reduced the sample size, particularly for the moderate-to-severe acute GVHD prediction model. A larger dataset would likely increase the model's predictive ability by providing more robust training and validation opportunities. Although data imputation was used to address some missing values, this approach may affect the interpretability and performance of the models. Addressing these limitations in future studies will be crucial for improving model accuracy and clinical relevance.

5. CONCLUSION AND RECOMMENDATIONS

In this study, we developed ML models to predict critical outcomes following allogeneic HSCT. While these models demonstrate promising proof-of-concept results, their generalizability and predictive performance could be further improved through external validation and access to larger, more diverse datasets. These efforts will be crucial for refining the models and ensuring their applicability across varied clinical contexts.

6. REFERENCES

Afram, G., Simón, J. A. P., Remberger, M., Caballero-Velázquez, T., Martino, R., Piñana, J. L., Ringden, O., Esquirol, A., Lopez-Corral, L., Garcia, I., López-Godino, O., Sierra, J., Caballero, D., Ljungman, P., Vazquez, L., & Hägglund, H. (2018). Reduced intensity conditioning increases risk of severe cGVHD: identification of risk factors for cGVHD in a multicenter setting. *Med Oncol*, *35*(6), 79. https://doi.org/10.1007/s12032-018-1127-2

Agius, R., Brieghel, C., Andersen, M. A., Pearson, A. T., Ledergerber, B., Cozzi-Lepri, A., Louzoun, Y., Andersen, C. L., Bergstedt, J., von Stemann, J. H., Jørgensen, M., Tang, M. E., Fontes, M., Bahlo, J., Herling, C. D., Hallek, M., Lundgren, J., MacPherson, C. R., Larsen, J., & Niemann, C. U. (2020). Machine learning can identify newly diagnosed patients with CLL at high risk of infection. *Nat Commun*, *11*(1), 363. https://doi.org/10.1038/s41467-019-14225-8

Akahoshi, Y., Kimura, S. I., Tada, Y., Matsukawa, T., Tamaki, M., Doki, N., Uchida, N., Tanaka, M., Nakamae, H., Kuriyama, T., Matsuoka, K. I., Ikeda, T., Kimura, T., Fukuda, T., Kanda, Y., Atsuta, Y., Murata, M., Terakura, S., & Nakasone, H. (2022). Cytomegalovirus gastroenteritis in patients with acute graft-versus-host disease. *Blood Adv*, *6*(2), 574-584. https://doi.org/10.1182/bloodadvances.2021005885

Ali, M. (2020). *PyCaret: An open source, low-code machine learning library in Python*. In (Version 3.0.0) https://www.pycaret.org

Allaw, F., Haddad, S. F., Zakhour, J., & Kanj, S. S. (2023). Management of cytomegalovirus infection in allogeneic hematopoietic stem cell transplants. *Int J Antimicrob Agents*, *62*(2), 106860. https://doi.org/10.1016/j.ijantimicag.2023.106860

Arai, Y., Kondo, T., Fuse, K., Shibasaki, Y., Masuko, M., Sugita, J., Teshima, T., Uchida, N., Fukuda, T., Kakihana, K., Ozawa, Y., Eto, T., Tanaka, M., Ikegame, K., Mori, T., Iwato, K., Ichinohe, T., Kanda, Y., & Atsuta, Y. (2019). Using a machine learning algorithm to predict acute graft-versus-host disease following allogeneic transplantation. *Blood Adv*, *3*(22), 3626-3634. https://doi.org/10.1182/bloodadvances.2019000934

Arora, S., Thakkar, D., Upasana, K., Yadav, A., Rastogi, N., Sharma, P. S., & Yadav, S. P. (2024). Incidence, Risk Factors, Characteristics, and Outcome of Chronic Graft Versus Host Disease in Children Undergoing Haploidentical Peripheral Blood Stem Cell Transplant With Post-transplant Cyclophosphamide. *J Pediatr Hematol Oncol*, *46*(1), e44-e50. https://doi.org/10.1097/mph.0000000000002786

Avni, B., Neiman, D., Shaked, E., Gal-Rosenberg, O., Grisariu, S., Kuzli, M., Avni, I., Fracchia, A., Stepensky, P., Zuckerman, T., Lev-Sagie, A., Fox-Fisher, I., Piyanzin, S., Moss, J., Salpeter, S. J., Glaser, B., Shemer, R., & Dor, Y. (2024). Chronic graft-versus-host disease detected by tissue-specific cell-free DNA methylation biomarkers. *J Clin Invest*, *134*(2). https://doi.org/10.1172/jci163541

Battipaglia, G., Ruggeri, A., Labopin, M., Volin, L., Blaise, D., Socié, G., Tabrizi, R., Cornelissen, J. J., Ghavamzadeh, A., Huynh, A., Wu, D., Yakoub-Agha, I., Maertens, J., Chevallier, P., Mohty, M., & Nagler, A. (2018). Refined graft-versus-host disease/relapse-free survival in transplant from HLA-identical related or unrelated donors in acute myeloid leukemia. *Bone Marrow Transplant*, *53*(10), 1295-1303. https://doi.org/10.1038/s41409-018-0169-6

Bhatia, S., Francisco, L., Carter, A., Sun, C. L., Baker, K. S., Gurney, J. G., McGlave, P. B., Nademanee, A., O'Donnell, M., Ramsay, N. K., Robison, L. L., Snyder, D., Stein, A., Forman, S. J., & Weisdorf, D. J. (2007). Late mortality after allogeneic hematopoietic cell transplantation and functional status of long-term survivors: report from the Bone Marrow Transplant Survivor Study. *Blood*, *110*(10), 3784-3792. https://doi.org/10.1182/blood-2007-03-082933

Bolaños-Meade, J., Hamadani, M., Wu, J., Al Malki Monzr, M., Martens Michael, J., Runaas, L., Elmariah, H., Rezvani Andrew, R., Gooptu, M., Larkin Karilyn, T., Shaffer Brian, C., El Jurdi, N., Loren Alison, W., Solh, M., Hall Aric, C., Alousi Amin, M., Jamy Omer, H., Perales, M.-A., Yao Janny, M., . . . Holtan Shernan, G. (2023). Post-Transplantation Cyclophosphamide-Based Graft-versus-Host Disease Prophylaxis. *New England Journal of Medicine*, *388*(25), 2338-2348. https://doi.org/10.1056/NEJMoa2215943

Chandra, G., Wang, J., Siirtola, P., & Röning, J. (2024). Leveraging machine learning for predicting acute graft-versus-host disease grades in allogeneic hematopoietic cell transplantation for T-cell prolymphocytic leukaemia. *BMC Med Res Methodol*, *24*(1), 112. https://doi.org/10.1186/s12874-024-02237-y

Choi, E. J., Jun, T. J., Park, H. S., Lee, J. H., Lee, K. H., Kim, Y. H., Lee, Y. S., Kang, Y. A., Jeon, M., Kang, H., Woo, J., & Lee, J. H. (2022). Predicting Long-term Survival After Allogeneic Hematopoietic Cell Transplantation in Patients With Hematologic Malignancies: Machine Learning-Based Model Development and Validation. *JMIR Med Inform*, *10*(3), e32313. https://doi.org/10.2196/32313

Ciftciler, R., Goker, H., Buyukasık, Y., Karaagac, T., Aksu, S., Tekin, F., & Demiroglu, H. (2020). Impact of ABO blood group incompatibility on the outcomes of allogeneic hematopoietic stem cell transplantation. *Transfus Apher Sci*, *59*(1), 102597. https://doi.org/10.1016/j.transci.2019.06.024

Copelan, E. A., Chojecki, A., Lazarus, H. M., & Avalos, B. R. (2019). Allogeneic hematopoietic cell transplantation; the current renaissance. *Blood Rev*, *34*, 34-44. https://doi.org/10.1016/j.blre.2018.11.001

Corbacioglu, S., Cesaro, S., Faraci, M., Valteau-Couanet, D., Gruhn, B., Rovelli, A., Boelens, J. J., Hewitt, A., Schrum, J., Schulz, A. S., Müller, I., Stein, J., Wynn, R., Greil, J., Sykora, K. W., Matthes-Martin, S., Führer, M., O'Meara, A., Toporski, J., . . . Dini, G. (2012). Defibrotide for prophylaxis of hepatic veno-occlusive disease in paediatric haemopoietic stem-cell transplantation: an open-label, phase 3, randomised controlled trial. *Lancet*, *379*(9823), 1301-1309. https://doi.org/10.1016/s0140-6736(11)61938-7

Couriel, D. R., Saliba, R. M., Giralt, S., Khouri, I., Andersson, B., de Lima, M., Hosing, C., Anderlini, P., Donato, M., Cleary, K., Gajewski, J., Neumann, J., Ippoliti, C., Rondon, G., Cohen, A., & Champlin, R. (2004). Acute and chronic graft-versus-host disease after ablative and nonmyeloablative conditioning for allogeneic hematopoietic transplantation. *Biol Blood Marrow Transplant*, *10*(3), 178-185. https://doi.org/10.1016/j.bbmt.2003.10.006

Çubukçu, H. C., Topcu, D., & Yenice, S. (2024). Machine learning-based clinical decision support using laboratory data. *Clin Chem Lab Med*, *62*(5), 793-823. https://doi.org/10.1515/cclm-2023-1037

Damodar, S., Shanley, R., MacMillan, M., Ustun, C., & Weisdorf, D. (2017). Donor-to-Recipient ABO Mismatch Does Not Impact Outcomes of Allogeneic Hematopoietic Cell Transplantation Regardless of Graft Source. *Biol Blood Marrow Transplant*, *23*(5), 795-804. https://doi.org/10.1016/j.bbmt.2017.02.009

Di Francesco, A., Raiola, A. M., Dominietto, A., Di Grazia, C., Gualandi, F., Van Lint, M. T., Bregante, S., Chiusolo, P., Laurenti, L., Sora, F., Giammarco, S., Metafuni, E., Fresa, A., Sica, S., Angelucci, E., & Bacigalupo, A. (2023). Acute graft versus host disease 1976-2020: reduced incidence and predictive factors. *Front Med (Lausanne)*, *10*, 1320692. https://doi.org/10.3389/fmed.2023.1320692

Eisenberg, L., Brossette, C., Rauch, J., Grandjean, A., Ottinger, H., Rissland, J., Schwarz, U., Graf, N., Beelen, D. W., Kiefer, S., Pfeifer, N., & Turki, A. T. (2022). Time-dependent prediction of mortality and cytomegalovirus reactivation after allogeneic hematopoietic cell transplantation using machine learning. *Am J Hematol*, *97*(10), 1309-1323. https://doi.org/10.1002/ajh.26671

Eisner, M. D., & August, C. S. (1995). Impact of donor and recipient characteristics on the development of acute and chronic graft-versus-host disease following pediatric bone marrow transplantation. *Bone Marrow Transplant*, *15*(5), 663-668.

Ferrara, J. L., Levine, J. E., Reddy, P., & Holler, E. (2009). Graft-versus-host disease. *Lancet*, *373*(9674), 1550-1561. https://doi.org/10.1016/s0140-6736(09)60237-3

Filipovich, A. H., Weisdorf, D., Pavletic, S., Socie, G., Wingard, J. R., Lee, S. J., Martin, P., Chien, J., Przepiorka, D., Couriel, D., Cowen, E. W., Dinndorf, P., Farrell, A., Hartzman, R., Henslee-Downey, J., Jacobsohn, D., McDonald, G., Mittleman, B., Rizzo, J. D., . . . Flowers, M. E. (2005). National Institutes of Health consensus development project on criteria for clinical trials in chronic graft-versus-host disease: I. Diagnosis and staging working group report. *Biol Blood Marrow Transplant*, *11*(12), 945-956. https://doi.org/10.1016/j.bbmt.2005.09.004

Flowers, M. E., Inamoto, Y., Carpenter, P. A., Lee, S. J., Kiem, H. P., Petersdorf, E. W., Pereira, S. E., Nash, R. A., Mielcarek, M., Fero, M. L., Warren, E. H., Sanders, J. E., Storb, R. F., Appelbaum, F. R., Storer, B. E., & Martin, P. J. (2011). Comparative analysis of risk factors for acute graft-versus-host disease and for chronic graft-versus-host disease according to National Institutes of Health consensus criteria. *Blood*, *117*(11), 3214-3219. https://doi.org/10.1182/blood-2010-08-302109

Gale, R. P., Bortin, M. M., van Bekkum, D. W., Biggs, J. C., Dicke, K. A., Gluckman, E., Good, R. A., Hoffmann, R. G., Kay, H. E., Kersey, J. H., & et al. (1987). Risk factors for acute graft-versus-host disease. *Br J Haematol*, *67*(4), 397-406. https://doi.org/10.1111/j.1365-2141.1987.tb06160.x

Garcia-Horton, A., Cyriac, S. L., Gedde-Dahl, T., Floisand, Y., Remberger, M., Mattsson, J., & Michelis, F. V. (2022). Patient age and donor HLA matching can stratify allogeneic hematopoietic cell transplantation patients into prognostic groups. *Eur J Haematol*, *109*(6), 672-679. https://doi.org/10.1111/ejh.13850

Giebel, S., Labopin, M., Schroeder, T., Swoboda, R., Maertens, J., Bourhis, J. H., Grillo, G., Salmenniemi, U., Hilgendorf, I., Kröger, N., Poiré, X., Cornelissen, J. J., Arat, M., Savani, B., Spyridonidis, A., Nagler, A., & Mohty, M. (2023). Fludarabine versus cyclophospamide in combination with myeloablative total body irradiation as conditioning for patients with acute myeloid leukemia treated with allogeneic hematopoietic cell transplantation. A study from the Acute Leukemia Working Party of the European Society for Blood and Marrow Transplantation. *Am J Hematol*, *98*(4), 580-587. https://doi.org/10.1002/ajh.26825

Gratwohl, A. (2012). The EBMT risk score. *Bone Marrow Transplant*, *47*(6), 749-756. https://doi.org/10.1038/bmt.2011.110

Gratwohl, A., Stern, M., Brand, R., Apperley, J., Baldomero, H., de Witte, T., Dini, G., Rocha, V., Passweg, J., Sureda, A., Tichelli, A., & Niederwieser, D. (2009). Risk score for outcome after

allogeneic hematopoietic stem cell transplantation: a retrospective analysis. *Cancer*, *115*(20), 4715-4726. https://doi.org/10.1002/cncr.24531

Harada, K., Kimura, S. I., Fuji, S., Najima, Y., Yakushijin, K., Uchida, N., Onizuka, M., Ikegame, K., Yano, S., Shingai, N., Matsuoka, K. I., Onishi, Y., Sawa, M., Takada, S., Kawakita, T., Fukuda, T., Kanda, J., Atsuta, Y., & Nakasone, H. (2021). Prognostic factors in salvage transplantation for graft failure following allogeneic hematopoietic stem cell transplantation. *Bone Marrow Transplant*, *56*(9), 2183-2193. https://doi.org/10.1038/s41409-021-01310-0

Harashima, S., Yoneda, R., Horie, T., Fujioka, Y., Nakamura, F., Kurokawa, M., & Yoshiuchi, K. (2019). Psychosocial Assessment of Candidates for Transplantation scale (PACT) and survival after allogeneic hematopoietic stem cell transplantation. *Bone Marrow Transplant*, *54*(7), 1013-1021. https://doi.org/10.1038/s41409-018-0371-6

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., . . . Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, *585*(7825), 357-362. https://doi.org/10.1038/s41586-020-2649-2

Herman, D. S., Rhoads, D. D., Schulz, W. L., & Durant, T. J. S. (2021). Artificial Intelligence and Mapping a New Direction in Laboratory Medicine: A Review. *Clin Chem*, *67*(11), 1466-1482. https://doi.org/10.1093/clinchem/hvab165

Jagasia, M., Arora, M., Flowers, M. E., Chao, N. J., McCarthy, P. L., Cutler, C. S., Urbano-Ispizua, A., Pavletic, S. Z., Haagenson, M. D., Zhang, M. J., Antin, J. H., Bolwell, B. J., Bredeson, C., Cahn, J. Y., Cairo, M., Gale, R. P., Gupta, V., Lee, S. J., Litzow, M., . . . Hahn, T. (2012). Risk factors for acute GVHD and survival after hematopoietic cell transplantation. *Blood*, *119*(1), 296-307. https://doi.org/10.1182/blood-2011-06-364265

Kamel, M., Kadian, M., Srinivas, T., Taber, D., & Posadas Salas, M. A. (2016). Tacrolimus confers lower acute rejection rates and better renal allograft survival compared to cyclosporine. *World J Transplant*, *6*(4), 697-702. https://doi.org/10.5500/wjt.v6.i4.697

Kanda, J. (2013). Effect of HLA mismatch on acute graft-versus-host disease. *Int J Hematol*, *98*(3), 300-308. https://doi.org/10.1007/s12185-013-1405-x

Kanda, J., Nakasone, H., Atsuta, Y., Toubai, T., Yokoyama, H., Fukuda, T., Taniguchi, S., Ohashi, K., Ogawa, H., Eto, T., Miyamura, K., Morishima, Y., Nagamura-Inoue, T., Sakamaki, H., Murata, M., & on behalf of the, G. W. G. o. t. J. S. f. H. C. T. (2014). Risk factors and organ involvement

of chronic GVHD in Japan. *Bone Marrow Transplantation*, *49*(2), 228-235. https://doi.org/10.1038/bmt.2013.151

Kataoka, K., Nannya, Y., Ueda, K., Kumano, K., Takahashi, T., & Kurokawa, M. (2010). Differential prognostic impact of pretransplant comorbidity on transplant outcomes by disease status and time from transplant: a single Japanese transplant centre study. *Bone Marrow Transplant*, *45*(3), 513-520. https://doi.org/10.1038/bmt.2009.194

Kato, M., Shiozawa, R., Koh, K., Nagatoshi, Y., Takita, J., Ida, K., Kikuchi, A., & Hanada, R. (2014). The effect of the order of total body irradiation and chemotherapy on graft-versus-host disease. *J Pediatr Hematol Oncol*, *36*(1), e9-12. https://doi.org/10.1097/MPH.0b013e318279e81c

Kawase, T., Morishima, Y., Matsuo, K., Kashiwase, K., Inoko, H., Saji, H., Kato, S., Juji, T., Kodera, Y., & Sasazuki, T. (2007). High-risk HLA allele mismatch combinations responsible for severe acute graft-versus-host disease and implication for its molecular mechanism. *Blood*, *110*(7), 2235-2241. https://doi.org/10.1182/blood-2007-02-072405

Kimura, F., Sato, K., Kobayashi, S., Ikeda, T., Sao, H., Okamoto, S., Miyamura, K., Mori, S., Akiyama, H., Hirokawa, M., Ohto, H., Ashida, H., & Motoyoshi, K. (2008). Impact of AB0-blood group incompatibility on the outcome of recipients of bone marrow transplants from unrelated donors in the Japan Marrow Donor Program. *Haematologica*, *93*(11), 1686-1693. https://doi.org/10.3324/haematol.12933

Kreidieh, F., Abou Dalle, I., Moukalled, N., El-Cheikh, J., Brissot, E., Mohty, M., & Bazarbachi, A. (2022). Relapse after allogeneic hematopoietic stem cell transplantation in acute myeloid leukemia: an overview of prevention and treatment. *Int J Hematol*, *116*(3), 330-340. https://doi.org/10.1007/s12185-022-03416-7

Kröger, N., Zabelina, T., Krüger, W., Renges, H., Stute, N., Kabisch, H., Jaburg, N., Löliger, C., Krüll, A., & Zander, A. R. (2001). Comparison of total body irradiation vs busulfan in combination with cyclophosphamide as conditioning for unrelated stem cell transplantation in CML patients. *Bone Marrow Transplant*, *27*(4), 349-354. https://doi.org/10.1038/sj.bmt.1702802

Lee, C., Haneuse, S., Wang, H. L., Rose, S., Spellman, S. R., Verneris, M., Hsu, K. C., Fleischhauer, K., Lee, S. J., & Abdi, R. (2018). Prediction of absolute risk of acute graft-versus-host disease following hematopoietic cell transplantation. *PLoS One*, *13*(1), e0190610. https://doi.org/10.1371/journal.pone.0190610

Lee, H., Myoung, H., & Kim, S. M. (2023). Review of two immunosuppressants: tacrolimus and cyclosporine. *J Korean Assoc Oral Maxillofac Surg*, *49*(6), 311-323. https://doi.org/10.5125/jkaoms.2023.49.6.311

Lee, S. J., Klein, J., Haagenson, M., Baxter-Lowe, L. A., Confer, D. L., Eapen, M., Fernandez-Vina, M., Flomenberg, N., Horowitz, M., Hurley, C. K., Noreen, H., Oudshoorn, M., Petersdorf, E., Setterholm, M., Spellman, S., Weisdorf, D., Williams, T. M., & Anasetti, C. (2007). High-resolution donor-recipient HLA matching contributes to the success of unrelated donor marrow transplantation. *Blood*, *110*(13), 4576-4583. https://doi.org/https://doi.org/10.1182/blood-2007-06-097386

Lu, Y., Zhao, Y. L., Zhang, J. P., Xiong, M., Cao, X. Y., Liu, D. Y., Sun, R. J., Wei, Z. J., Zhou, J. R., & Lu, D. P. (2021). Comparable outcomes among unmanipulated haploidentical, matched unrelated, and matched sibling donors in BU-based myeloablative hematopoietic stem cell transplantation for intermediate and adverse risk acute myeloid leukemia in complete remission: a single-center study. *Ann Hematol*, *100*(6), 1579-1591. https://doi.org/10.1007/s00277-020-04355-1

Ludajic, K., Balavarca, Y., Bickeböller, H., Rosenmayr, A., Fischer, G. F., Faé, I., Kalhs, P., Pohlreich, D., Kouba, M., Dobrovolna, M., & Greinix, H. T. (2009). Minor ABO-mismatches are risk factors for acute graft-versus-host disease in hematopoietic stem cell transplant patients. *Biol Blood Marrow Transplant*, *15*(11), 1400-1406. https://doi.org/10.1016/j.bbmt.2009.07.002

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S.-I. (2020). From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence*, *2*(1), 56-67. https://doi.org/10.1038/s42256-019-0138-9

MacMillan, M. L., DeFor, T. E., & Weisdorf, D. J. (2012). What predicts high risk acute graft-versus-host disease (GVHD) at onset?: identification of those at highest risk by a novel acute GVHD risk score. *Br J Haematol*, *157*(6), 732-741. https://doi.org/10.1111/j.1365-2141.2012.09114.x

Majhail, N. S. (2017). Long-term complications after hematopoietic cell transplantation. *Hematol Oncol Stem Cell Ther*, *10*(4), 220-227. https://doi.org/10.1016/j.hemonc.2017.05.009

Majhail, N. S., Mau, L. W., Chitphakdithai, P., Denzen, E. M., Joffe, S., Lee, S. J., LeMaistre, C. F., Loberiza, F., Parsons, S. K., Repaczki-Jones, R., Robinett, P., Douglas Rizzo, J., Murphy, E., Logan, B., & Le-Rademacher, J. (2020). Transplant center characteristics and survival after allogeneic hematopoietic cell transplantation in adults. *Bone Marrow Transplant*, *55*(5), 906-917. https://doi.org/10.1038/s41409-019-0748-1

Martin, P. J., Levine, D. M., Storer, B. E., Warren, E. H., Zheng, X., Nelson, S. C., Smith, A. G., Mortensen, B. K., & Hansen, J. A. (2017). Genome-wide minor histocompatibility matching as related to the risk of graft-versus-host disease. *Blood*, *129*(6), 791-798. https://doi.org/10.1182/blood-2016-09-737700

McDonald, G. B., Sandmaier, B. M., Mielcarek, M., Sorror, M., Pergam, S. A., Cheng, G. S., Hingorani, S., Boeckh, M., Flowers, M. D., Lee, S. J., Appelbaum, F. R., Storb, R., Martin, P. J., Deeg, H. J., Schoch, G., & Gooley, T. A. (2020). Survival, Nonrelapse Mortality, and Relapse-Related Mortality After Allogeneic Hematopoietic Cell Transplantation: Comparing 2003-2007 Versus 2013-2017 Cohorts. *Ann Intern Med*, *172*(4), 229-239. https://doi.org/10.7326/m19-2936

Molnar, C. (2019). *Interpretable Machine Learning A Guide for Making Black Box Models Explainable*. Lean Publishing.

Mussetti, A., Rius-Sansalvador, B., Moreno, V., Peczynski, C., Polge, E., Galimard, J. E., Kröger, N., Blaise, D., Peffault de Latour, R., Kulagin, A., Mousavi, A., Stelljes, M., Hamladji, R. M., Middeke, J. M., Salmenniemi, U., Sengeloev, H., Forcade, E., Platzbecker, U., Reményi, P., . . . Peric, Z. (2024). Artificial intelligence methods to estimate overall mortality and non-relapse mortality following allogeneic HCT in the modern era: an EBMT-TCWP study. *Bone Marrow Transplant*, *59*(2), 232-238. https://doi.org/10.1038/s41409-023-02147-5

Nagler, A., Labopin, M., Dholaria, B., Wu, D., Choi, G., Aljurf, M., Ciceri, F., Gedde-Dahl, T., Meijer, E., Niittyvuopio, R., Bondarenko, S., Bourhis, J. H., Cornelissen, J. J., Socié, G., Koc, Y., Canaani, J., Savani, B., Bug, G., Spyridonidis, A., . . . Mohty, M. (2022). Graft-versus-Host Disease Prophylaxis with Post-Transplantation Cyclophosphamide versus Cyclosporine A and Methotrexate in Matched Sibling Donor Transplantation. *Transplant Cell Ther*, *28*(2), 86.e81-86.e88. https://doi.org/10.1016/j.jtct.2021.11.013

Obstfeld, A. E. (2023). Hematology and Machine Learning. *J Appl Lab Med*, *8*(1), 129-144. https://doi.org/10.1093/jalm/jfac108

Ozawa, S., Nakaseko, C., Nishimura, M., Maruta, A., Cho, R., Ohwada, C., Sakamaki, H., Sao, H., Mori, S., Okamoto, S., Miyamura, K., Kato, S., Kawase, T., Morishima, Y., & Kodera, Y. (2007). Chronic graft-versus-host disease after allogeneic bone marrow transplantation from an unrelated donor: incidence, risk factors and association with relapse. A report from the Japan Marrow Donor Program. *Br J Haematol*, *137*(2), 142-151. https://doi.org/10.1111/j.1365-2141.2007.06543.x

Ozdemir, Z. N., & Civriz Bozdağ, S. (2018). Graft failure after allogeneic hematopoietic stem cell transplantation. *Transfus Apher Sci*, *57*(2), 163-167. https://doi.org/10.1016/j.transci.2018.04.014

Pérez-Simón, J. A., Díez-Campelo, M., Martino, R., Brunet, S., Urbano, A., Caballero, M. D., de León, A., Valcárcel, D., Carreras, E., del Cañizo, M. C., López-Fidalgo, J., Sierra, J., & San Miguel, J. F. (2005). Influence of the intensity of the conditioning regimen on the characteristics of

acute and chronic graft-versus-host disease after allogeneic transplantation. *Br J Haematol*, *130*(3), 394-403. https://doi.org/10.1111/j.1365-2141.2005.05614.x

Przepiorka, D., Smith, T. L., Folloder, J., Khouri, I., Ueno, N. T., Mehra, R., Körbling, M., Huh, Y. O., Giralt, S., Gajewski, J., Donato, M., Cleary, K., Claxton, D., Braunschweig, I., van Besien, K., Andersson, B. S., Anderlini, P., & Champlin, R. (1999). Risk factors for acute graft-versus-host disease after allogeneic blood stem cell transplantation. *Blood*, *94*(4), 1465-1470.

Radakovich, N., Nagy, M., & Nazha, A. (2020). Machine learning in haematological malignancies. *Lancet Haematol*, *7*(7), e541-e550. https://doi.org/10.1016/s2352-3026(20)30121-6

Ramdial, J. L., Mehta, R. S., Saliba, R. M., Alousi, A. M., Bashir, Q., Hosing, C., Kebriaei, P., Olson, A. L., Oran, B., Qazilbash, M. H., Srour, S. A., Andersson, B. S., Champlin, R. E., & Popat, U. (2021). Acute graft-versus-host disease is the foremost cause of late nonrelapse mortality. *Bone Marrow Transplant*, *56*(8), 2005-2012. https://doi.org/10.1038/s41409-021-01274-1

Remberger, M., Afram, G., Sundin, M., Uhlin, M., LeBlanc, K., Björklund, A., Mattsson, J., & Ljungman, P. (2016). High incidence of severe chronic GvHD after HSCT with sibling donors. A single center analysis. *Bone Marrow Transplantation*, *51*(11), 1518-1521. https://doi.org/10.1038/bmt.2016.159

Remberger, M., Grønvold, B., Ali, M., Mattsson, J., Egeland, T., Lundin, K. U., Myhre, A., Abrahamsen, I., Heldal, D., Dybedal, I., Tjønnfjord, G. E., Gedde-Dahl, T., & Fløisand, Y. (2020). The CD34(+) Cell Dose Matters in Hematopoietic Stem Cell Transplantation with Peripheral Blood Stem Cells from Sibling Donors. *Clin Hematol Int*, *2*(2), 74-81. https://doi.org/10.2991/chi.d.200221.001

Ren, J., Xu, J., Sun, J., Wu, X., Yang, X., Nie, C., Lan, L., Zeng, Y., Zheng, X., Li, J., Lin, Q., Hu, J., & Yang, T. (2024). Reactivation of cytomegalovirus and bloodstream infection and its impact on early survival after allogeneic haematopoietic stem cell transplantation: a multicentre retrospective study. *Front Microbiol*, *15*, 1405652. https://doi.org/10.3389/fmicb.2024.1405652

Richardson, P. G., Carreras, E., Iacobelli, M., & Nejadnik, B. (2018). The use of defibrotide in blood and marrow transplantation. *Blood Adv*, *2*(12), 1495-1509. https://doi.org/10.1182/bloodadvances.2017008375

Rowlings, P. A., Przepiorka, D., Klein, J. P., Gale, R. P., Passweg, J. R., Henslee-Downey, P. J., Cahn, J. Y., Calderwood, S., Gratwohl, A., Socié, G., Abecasis, M. M., Sobocinski, K. A., Zhang, M. J., & Horowitz, M. M. (1997). IBMTR Severity Index for grading acute graft-versus-host

disease: retrospective comparison with Glucksberg grade. *Br J Haematol*, *97*(4), 855-864. https://doi.org/10.1046/j.1365-2141.1997.1112925.x

Ruggeri, A., Sun, Y., Labopin, M., Bacigalupo, A., Lorentino, F., Arcese, W., Santarone, S., Gülbas, Z., Blaise, D., Messina, G., Ghavamzadeh, A., Malard, F., Bruno, B., Diez-Martin, J. L., Koc, Y., Ciceri, F., Mohty, M., & Nagler, A. (2017). Post-transplant cyclophosphamide versus anti-thymocyte globulin as graft- versus-host disease prophylaxis in haploidentical transplant. *Haematologica*, *102*(2), 401-410. https://doi.org/10.3324/haematol.2016.151779

Saad, A., de Lima, M., Anand, S., Bhatt, V. R., Bookout, R., Chen, G., Couriel, D., Di Stasi, A., El-Jawahri, A., Giralt, S., Gutman, J., Ho, V., Horwitz, M., Hsu, J., Juckett, M., Kharfan-Dabaja, M. A., Loren, A., Meade, J., Mielcarek, M., . . . Pluchino, L. (2020). Hematopoietic Cell Transplantation, Version 2.2020, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw*, *18*(5), 599-634. https://doi.org/10.6004/jnccn.2020.0021

Sahin, U., Toprak, S. K., Atilla, P. A., Atilla, E., & Demirer, T. (2016). An overview of infectious complications after allogeneic hematopoietic stem cell transplantation. *J Infect Chemother*, *22*(8), 505-514. https://doi.org/10.1016/j.jiac.2016.05.006

Salehnasab, C., Hajifathali, A., Asadi, F., Parkhideh, S., Kazemi, A., Roshanpoor, A., Mehdizadeh, M., Tavakoli-Ardakani, M., & Roshandel, E. (2021). An Intelligent Clinical Decision Support System for Predicting Acute Graft-versus-host Disease (aGvHD) following Allogeneic Hematopoietic Stem Cell Transplantation. *J Biomed Phys Eng*, *11*(3), 345-356. https://doi.org/10.31661/jbpe.v0i0.2012-1244

Sayehmiri, K., Eshraghian, M. R., Mohammad, K., Alimoghaddam, K., Foroushani, A. R., Zeraati, H., Golestan, B., & Ghavamzadeh, A. (2008). Prognostic factors of survival time after hematopoietic stem cell transplant in acute lymphoblastic leukemia patients: Cox proportional hazard versus accelerated failure time models. *J Exp Clin Cancer Res*, *27*(1), 74. https://doi.org/10.1186/1756-9966-27-74

Shaw, B. E., Flynn, K. E., He, N., Cusatis, R., D'Souza, A., Hamilton, B. K., Horowitz, M. M., Mattila, D., Phelan, R., Lee, S. J., & Brazauskas, R. (2024). Incorporating patient-reported outcome data into a predictive calculator for allogeneic hematopoietic cell transplantation recipients. *Cancer*, *130*(10), 1826-1835. https://doi.org/10.1002/cncr.35189

Shouman, M. T., Mansour, O. M., El Gammal, M. M., Abdel-Fattah, R. M., Samra, M. A., Elhaddad, A. M., Maher, M. A., & Mahmoud, H. K. (2023). Post-transplantation Cyclophosphamide-based Graft-versus-host-disease Prophylaxis Compared to Methotrexate-cyclosporine a in Matched-related Allogeneic Hematopoietic Stem Cell Transplantation. *Hematol Oncol Stem Cell Ther*, *16*(4), 379-387. https://doi.org/10.56875/2589-0646.1065

Shouval, R., Bonifazi, F., Fein, J., Boschini, C., Oldani, E., Labopin, M., Raimondi, R., Sacchi, N., Dabash, O., Unger, R., Mohty, M., Rambaldi, A., & Nagler, A. (2017). Validation of the acute leukemia-EBMT score for prediction of mortality following allogeneic stem cell transplantation in a multi-center GITMO cohort. *Am J Hematol*, *92*(5), 429-434. https://doi.org/10.1002/ajh.24677

Shouval, R., Labopin, M., Unger, R., Giebel, S., Ciceri, F., Schmid, C., Esteve, J., Baron, F., Gorin, N. C., Savani, B., Shimoni, A., Mohty, M., & Nagler, A. (2016). Prediction of Hematopoietic Stem Cell Transplantation Related Mortality- Lessons Learned from the In-Silico Approach: A European Society for Blood and Marrow Transplantation Acute Leukemia Working Party Data Mining Study. *PLoS One*, *11*(3), e0150637. https://doi.org/10.1371/journal.pone.0150637

Solh, M. M., Bashey, A., Solomon, S. R., Morris, L. E., Zhang, X., Brown, S., & Holland, H. K. (2018). Long term survival among patients who are disease free at 1-year post allogeneic hematopoietic cell transplantation: a single center analysis of 389 consecutive patients. *Bone Marrow Transplant*, *53*(5), 576-583. https://doi.org/10.1038/s41409-017-0076-2

Sorror, M. L., Giralt, S., Sandmaier, B. M., De Lima, M., Shahjahan, M., Maloney, D. G., Deeg, H. J., Appelbaum, F. R., Storer, B., & Storb, R. (2007). Hematopoietic cell transplantation specific comorbidity index as an outcome predictor for patients with acute myeloid leukemia in first remission: combined FHCRC and MDACC experiences. *Blood*, *110*(13), 4606-4613. https://doi.org/10.1182/blood-2007-06-096966

Squillaci, D., Marcuzzi, A., Rimondi, E., Riccio, G., Barbi, E., Zanon, D., & Maximova, N. (2023). Defibrotide impact on the acute GVHD disease incidence in pediatric hematopoietic stem cell transplant recipients. *Life Sci Alliance*, *6*(5). https://doi.org/10.26508/lsa.202201786

Tang, S., Chappell, G. T., Mazzoli, A., Tewari, M., Choi, S. W., & Wiens, J. (2020). Predicting Acute Graft-Versus-Host Disease Using Machine Learning and Longitudinal Vital Sign Data From Electronic Health Records. *JCO Clin Cancer Inform*, *4*, 128-135. https://doi.org/10.1200/cci.19.00105

Team, T. P. D. (2020). *pandas-dev/pandas: Pandas*. In (Version latest) Zenodo. https://doi.org/10.5281/zenodo.3509134

http://dx.doi.org/10.5281/zenodo.3509134

Uzay, A., Gündoğdu, Y., Koşan, B., Yetiş, T., & Kartı, S. S. (2024). Treosulfan is a safe and effective alternative to busulfan for conditioning in adult allogeneic HSCT patients: Data from a single center. *Cancer Med*, *13*(10), e7292. https://doi.org/10.1002/cam4.7292

Wingard, J. R., Majhail, N. S., Brazauskas, R., Wang, Z., Sobocinski, K. A., Jacobsohn, D., Sorror, M. L., Horowitz, M. M., Bolwell, B., Rizzo, J. D., & Socié, G. (2011). Long-term survival and late deaths after allogeneic hematopoietic cell transplantation. *J Clin Oncol*, *29*(16), 2230-2239. https://doi.org/10.1200/jco.2010.33.7212

Wiskemann, J., Kleindienst, N., Kuehl, R., Dreger, P., Schwerdtfeger, R., & Bohus, M. (2015). Effects of physical exercise on survival after allogeneic stem cell transplantation. *Int J Cancer*, *137*(11), 2749-2756. https://doi.org/10.1002/ijc.29633

Wong, F. L., Teh, J. B., Atencio, L., Stiller, T., Kim, H., Chanson, D., Forman, S. J., Nakamura, R., & Armenian, S. H. (2020). Conditional Survival, Cause-Specific Mortality, and Risk Factors of Late Mortality After Allogeneic Hematopoietic Cell Transplantation. *J Natl Cancer Inst*, *112*(11), 1153-1161. https://doi.org/10.1093/jnci/djaa022

Wong, J. Y. C., Filippi, A. R., Dabaja, B. S., Yahalom, J., & Specht, L. (2018). Total Body Irradiation: Guidelines from the International Lymphoma Radiation Oncology Group (ILROG). *International Journal of Radiation Oncology, Biology, Physics*, *101*(3), 521-529. https://doi.org/10.1016/j.ijrobp.2018.04.071

Wu, L., Wu, Y., Shi, J., Lai, X., Zhao, Y., Liu, L., Yu, J., Yang, L., Zhu, P., Zheng, W., Hu, Y., Wu, W., Zhu, Y., Cai, Z., Huang, H., & Luo, Y. (2023). Survival and late mortality among patients who survived disease-free for 2 years after stem cell transplantation. *Br J Haematol*, *202*(3), 608-622. https://doi.org/10.1111/bjh.18905

Zeiser, R., & Blazar, B. R. (2017a). Acute Graft-versus-Host Disease - Biologic Process, Prevention, and Therapy. *N Engl J Med*, *377*(22), 2167-2179. https://doi.org/10.1056/NEJMra1609337

Zeiser, R., & Blazar, B. R. (2017b). Pathophysiology of Chronic Graft-versus-Host Disease and Therapeutic Targets. *N Engl J Med*, *377*(26), 2565-2579. https://doi.org/10.1056/NEJMra1703472

7. APPENDIX

İNSAN ARAŞTIRMALARI ETİK KURULU KARAR FORMU

ETİK KURULUN ADI	ANKARA ÜNİVERSİTESİ TIP FAKÜLTESİ İNSAN ARAŞTIRMALARI ETİK KURULU
AÇIK ADRES	Ankara Üniversitesi Tıp Fakültesi Morfoloji Binası 06100 Sıhhıye/ANKARA
TELEFON	
FAKS	
E-POSTA	

BAŞVURU BİLGİLERİ	ARAŞTIRMANIN AÇIK ADI	Başvuru No: 2023000542(2023/542) Makine Öğrenmesi Kullanılarak Allojenik Hematopoetik Kök Hücre Nakli Yapılan Hastalarda Sağkalım ve Graft Versus Host Hastalığı Tahmini	
	KOORDİNATÖR/SORUMLU ARAŞTIRMACI ÜNVANI/ADI/SOYADI		
	PROJE YÜRÜTÜCÜSÜ UNVANI/ADI/SOYADI (TÜBİTAK vb. gibi kaynaklardan destek alanlar için)		
	KOORDİNATÖR/SORUMLU ARAŞTIRMACININ UZMANLIK ALANI	Hematoloji	
	KOORDİNATÖR/SORUMLU ARAŞTIRMACININ BULUNDUĞU MERKEZ	Ankara Üniversitesi Tıp Fakültesi İç Hastalıkları Anabilim Dalı, Hematoloji Bilim Dalı	
	ARAŞTIRMAYA KATILAN MERKEZLER	TEK MERKEZ ☒	ÇOK MERKEZLİ ☐

KARAR BİLGİLERİ	**Karar No:İ09-596-23**	**Tarih: 12 Ekim 2023**
	Yukarıda bilgileri verilen başvuru dosyası ile ilgili belgeler araştırmanın/çalışmanın gerekçe, amaç, yaklaşım ve yöntemleri dikkate alınarak incelenmiş ve uygun bulunmuş olup araştırmanın/çalışmanın başvuru dosyasında belirtilen merkezde gerçekleştirilmesinde etik ve bilimsel sakınca bulunmadığına toplantıya katılan etik kurul üye tam sayısının salt çoğunluğu ile karar verilmiştir.	

İNSAN ARAŞTIRMALARI ETİK KURULU

ÇALIŞMA ESASI	İyi Klinik Uygulamaları Kılavuzu
BAŞKANIN UNVANI / ADI / SOYADI:	Prof.Dr.Hakan ERGÜN

Unvanı/Adı/Soyadı	Uzmanlık Alanı	Kurumu	Araştırma ile ilişki		İmza
Prof.Dr.Hakan ERGÜN	Tıbbi Farmakoloji	A.Ü. Tıp Fakültesi	E ☐	H ☒	
Prof.Dr.Berna ARDA	Tıp Tarihi ve Etik	A.Ü. Tıp Fakültesi	E ☐	H ☒	
Prof.Dr.Hatice ILGIN RUHİ	Tıbbi Genetik	A.Ü. Tıp Fakültesi	E ☐	H ☒	
Prof.Dr.Sevim AYDIN	Histoloji ve Embriyoloji	A.Ü. Tıp Fakültesi	E ☐	H ☒	
Prof.Dr.Berna SAVAŞ	Tıbbi Patoloji	A.Ü. Tıp Fakültesi	E ☐	H ☒	
Prof.Dr.Yüksel ÜRÜN	Tıbbi Onkoloji	A.Ü.Tıp Fakültesi	E ☐	H ☒	
Prof.Dr.Cihangir AKYOL	Genel Cerrahi	A.Ü. Tıp Fakültesi	E ☐	H ☒	
Prof.Dr.Başak Ceyda MEÇO	Anesteziyoloji ve Reanimasyon	A.Ü. Tıp Fakültesi	E ☐	H ☒	
Doç.Dr.Emel OKULU	Çocuk Sağlığı ve Hastalıkları	A.Ü. Tıp Fakültesi	E ☐	H ☒	
Doç.Dr.Zahide Çiler BÜYÜKATALAY YALDIZ	Kulak, Burun ve Boğaz Hastalıkları	A.Ü. Tıp Fakültesi	E ☐	H ☒	
Doç.Dr.Rezzak YILMAZ	Nöroloji	A.Ü. Tıp Fakültesi	E ☐	H ☒	
Doç.Dr.Serkan AKBULUT	Cerrahi Onkoloji	A.Ü. Tıp Fakültesi	E ☐	H ☒	
Doç.Dr.Miraç YILDIRIM	Çocuk Sağlığı ve Hastalıkları	A.Ü. Tıp Fakültesi	E ☐	H ☒	
Dr.İrem KAR	Biyoistatistik	A.Ü. Tıp Fakültesi	E ☐	H ☒	

Funda BAYKAL KILIÇ
A.Ü.T.F. İnsan Araştırmaları
Etik Kurulu
Aslı Gibidir