

令和 6 年 6 月 2 1 日現在

機関番号：14301

研究種目：研究活動スタート支援

研究期間：2022～2023

課題番号：22K21082

研究課題名（和文）アンサンブル学習及び公開データベースを用いた造血幹細胞移植予後予測モデルの開発

研究課題名（英文）Establishment of a Stacked Ensemble Model for Risk Stratification after
Allogeneic HSCT Utilizing Big Data

研究代表者

岩崎 惇（Iwasaki, Makoto）

京都大学・医学研究科・医員

研究者番号：40967527

交付決定額（研究期間全体）：（直接経費） 2,200,000 円

研究成果の概要（和文）：本研究では、急性GVHD発症が移植後の中長期的な予後に与える可能性を示唆し、それらの因子をアンサンブル学習を用いた中長期予後予測モデルに組み込んで開発する事に成功した。開発する中で、移植後の急性GVHDおよびその治療を考慮に入れる事で、アンサンブル学習モデルの精度が時間経過に伴い徐々に向上する事を明らかにした。更に、予後に影響する因子を明らかにするため、SHapley Additive Explanationsを用いると、重症急性GVHDは、性別一致、ドナーソースや移植前疾患状態を含む既知の予後因子と同様に、非再発死亡率、全生存を予測する因子として重要である事が明らかとなった。

研究成果の学術的意義や社会的意義

造血幹細胞移植は現代においても、再発・移植後合併症で約半数は病気の治癒が難しく、予後を予測して適応を慎重に判断していく必要がある。特に、移植後1年を過ぎた中長期の予後の予測は、移植後の合併症や治療の影響も受けることから現代においても難しい課題の一つである。本研究において、機械学習・アンサンブル学習を用いて、移植後の因子も取り入れた予後の予測を行う事で、従来よりも優れたモデルが開発できる事を示しており、将来の造血幹細胞移植の適応判断や治療法選択に生きるものと考ええる。

研究成果の概要（英文）：We successfully developed a mid- to long-term prognosis prediction model incorporating these factors using stacked ensemble model. By taking into account post-transplant acute GVHD and its treatment, we demonstrated that the accuracy of the stacked ensemble model gradually improves over time, which underscores the significance of acute GVHD events in mid-to-long-term post-transplantation outcomes. Furthermore, to identify factors influencing prognosis, we used SHapley Additive Explanations, which revealed that severe acute GVHD is an important predictor of non-relapse mortality and overall survival, comparable to known prognostic factors such as sex matching, donor source, and pre-transplant disease status.

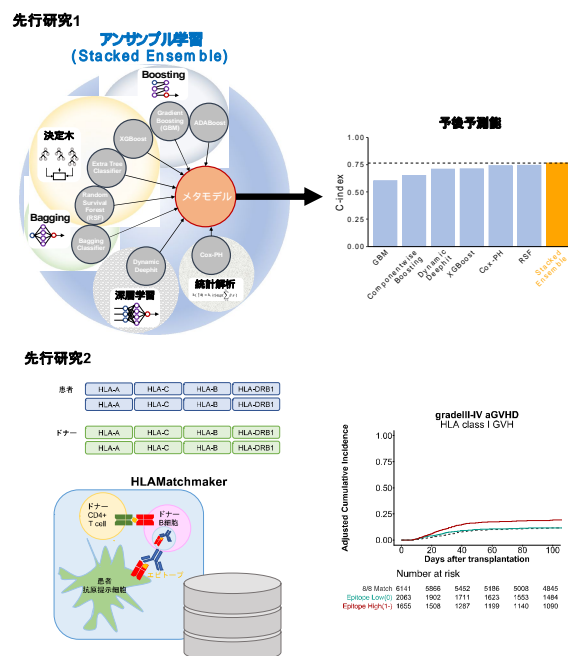
研究分野：血液学

キーワード：造血幹細胞移植 移植片対宿主病 機械学習 アンサンブル学習 生存時間解析 競合リスク

1. 研究開始当初の背景

他者の造血・免疫細胞を移植する同種造血幹細胞移植は、分子標的薬や細胞治療を含む新規治療法が日々開発されている現在においても、多くの血液腫瘍において唯一の根治を目指せる治療方法である。造血幹細胞移植においては、強力な化学放射線療法（前処置）を実施し、ドナーから採取した細胞を移植する。移植を実施後にも、生着の担保、及び移植後に特異的な移植片対宿主病（GVHD）と呼ばれる免疫学的合併症の予防のために免疫抑制剤を使用する治療であり、移植後には再発、GVHD や感染症等、様々な事象を想定して進めて行かなければならない治療である。そのため、複数の要因が重なり合って移植治療に影響を与えるため、多くの移植方法の選択肢に対してその後の予後を予測し、最適だと考えられる治療を選択する事が実臨床でも非常に重要である。疾患及び移植前病期、移植前の併存症の状態といったそれぞれの因子に基づいて移植後の予後を予測する手法はこれまで開発され、実際に臨床における移植の是非や移植方法の判断に役立てられている。疾患及び移植前病期といった疾患状態から移植のリスクを考える上では、Disease Risk Index（DRI）が用いられている(Armand et al., 2014)。心臓、腎臓、肝臓といった各臓器における既往症や感染症の状態等、併存症に基づき移植後の予後を予測するスコアとしては、Hematopoietic Cell Transplantation Comorbidity Index (HCT-CI)が用いられている(Sorrer et al., 2005)。また、患者年齢、疾患病期、診断から移植までの期間、ドナーの種類、患者とドナーの性別に基づいた項目をそれぞれスコア化し、複合的に移植後予後を予測するための指標として、European Bone Marrow Transplantation (EBMT)が提唱したリスクスコアや、DRI と HCT-CI を組合わせた HCT-CR などが報告されている (Passweg et al., 2004; Kongtim et al., 2019)。また、造血幹細胞移植後の予後を予測する上では、再発、GVHD 以外にも感染症や移植特異的な臓器の合併症を予測する事が必要であり、それぞれには特有のリスク因子が存在し、それらを複合的に解釈する事であらゆるエンドポイントの予測が可能になる。2015 年には、造血幹細胞移植後の死亡、再発、重症 GVHD の発症に関する複合エンドポイントとして GVHD, Relapse-Free Survival (GRFS)が提唱され、移植後に重篤な疾患状態のない生存を測る指標として用いられている(Holtan et al., 2015)。このように、多くの因子が影響する造血幹細胞移植において、移植前因子の中で移植後予後に影響を与える因子を抽出する事、その上でそれらの因子から複合的に解釈の必要なエンドポイントを予測する事が必要であり、既存の手法では十分な予後予測が行えているとは言い難い現状がある。

先行研究において、古典的統計学解析手法と複数の機械学習モデルを、スタッキングと呼ばれるアンサンブル学習手法を用いる事で統合して新たなメタモデルを作成する事で、全生存、再発、非再発死亡、GVHD 発症といった様々なエンドポイントにおいて従来の統計学的解析や機械学習モデルと比較して優れた予後予測能を示すアルゴリズムを開発した(Iwasaki et al., 2021, 先行研究 1)。また、移植前の因子として、特に造血幹細胞移植特異的な合併症である GVHD や、逆に腫瘍に対して免疫反応が働き抑制に繋がる移植片対腫瘍効果 (GVL) を考える上では、Human Leukocyte Antigen (HLA) と呼ばれる自己と他者の細胞を判別する際に働く分子の患者・ドナー間での一致度が重要であるが、非血縁者間移植において、HLAMatchmaker と呼ばれる免疫グロブリン抗体が



認識するエピトープが、移植後の重症急性 GVHD 発症、非再発死亡のリスクとなり、全生存に影響する事を示した(Iwasaki et al., 2022, 先行研究 2)。正確な造血幹細胞移植後予後予測手法の開発は、大きな課題として現在も盛んに研究が進められている分野であり、上記知見に基づき新たな予後予測モデルが開発できるかは、本研究における重要な問いと考えた。

2．研究の目的

本研究では、機械学習手法を用いて、造血幹細胞移植後予後をより正確に予測する事ができるアルゴリズムを開発する事を目的としており、アンサンブル学習を用いた新規アルゴリズムを解析に用いている点で独自性がある。アウトカムとしては、移植後の合併症・治療の影響から予測が困難とされてきた慢性 GVHD を始めとする中長期的予後を対象とし、より正確な予後を予測できる事で、個々の症例に即した移植方法の選択に寄与する事を目指す。

3．研究の方法

アンサンブル学習の基礎に用いる機械学習アルゴリズムの選定や、用いるアルゴリズムの最適化を行っていきたいと考えている。基礎となるモデル個々の hyperparameter の調整や、組み合わせの最適化に関しては、更に検討の余地がある課題と考えており、その事でモデルの精度向上を目指す。更に、移植前後の予後予測因子を最適化する事で、更なる予測精度の向上に取り組む。データは、the Transplant Registry Unified Management Program (TRUMP)に登録されている 2011 年から 2020 年までに実施された造血幹細胞移植データ約 30000 症例を対象とする。予測精度に関しては、C-index 及び Calibration を基に、先行研究におけるアルゴリズムと比較する。

4．研究成果

アンサンブルモデル学習モデルを使用して、造血幹細胞移植後中長期的予後を予測するモデルを開発した。症例は、Transplant Registry Unified Management Program (TRUMP) を使用し、2010 年から 2018 年の間に最初の allo-HSCT を受けた 21,840 人の血液悪性腫瘍患者を対象とした。移植後の急性 GVHD およびその治療を考慮に入れ、移植後 30 日目、60 日目、100 日目それぞれの地点で、中等症以上の aGVHD 発症、コルチコステロイド、抗胸腺細胞グロブリンへの治療反応性を含めた移植後の項目を含めた予後因子を含む事で、アンサンブル学習モデルの cGVHD、NRM および OS において、精度の指標である C-index が時間が経つにつれて徐々に向上する事を明らかにした(慢性 GVHD/NRM/OS の C-index: 30 日目-0.589/0.643/0.635、60 日目-0.598/0.655/0.640、100 日目-0.598/0.688/0.640)。移植後 100 日時点のモデルを用いる事で、4 つのリスクグループに層別化できる事を確認した。予後に影響する因子を明らかにするため、SHapley Additive Explanations を用いると、重症急性 GVHD は、性別一致、ドナーソースや移植前疾患状態を含む既知の予後因子と同様に、非再発死亡率、全生存を予測する因子として重要である事が明らかとなった。また、慢性 GVHD には、中等症以上の急性 GVHD 発症及びその反応性が影響を与える事が明らかとなった。本研究では、急性 GVHD 発症が移植後の中長期的な予後に与える可能性を示唆し、それらの因子をアンサンブル学習を用いた中長期予後予測モデルに組み込んで開発する事に成功した。

5 . 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計2件（うち招待講演 0件 / うち国際学会 1件）

1 . 発表者名 Makoto Iwasaki
2 . 発表標題 THE EFFECT OF HUMAN LEUKOCYTE ANTIGEN EPITOPE MATCHING ON OUTCOMES AFTER HAPLOIDENTICAL TRANSPLANTATION DEPENDS ON GRAFT-VERSUS-HOST DISEASE PROPHYLAXIS AND PRETRANSPLANT DISEASE STATUS
3 . 学会等名 EBMT2023（国際学会）
4 . 発表年 2023年

1 . 発表者名 岩崎 惇
2 . 発表標題 血縁者間ハプロ半合致末梢血幹細胞移植におけるT細胞HLAエピトープの意義
3 . 学会等名 第31回日本組織適合性学会大会
4 . 発表年 2023年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6 . 研究組織

	氏名 （ローマ字氏名） （研究者番号）	所属研究機関・部局・職 （機関番号）	備考
--	---------------------------	-----------------------	----

7 . 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------