Available online at www.blood-genomics.com

Blood and Genomics

2025, 9(2): 66−82

B & G

Review Article

Machine learning in graft-versus-host disease: current applications and future frontiers

Yibo Wang[1,2,3,△], Xinwen Zhang[1,2,3,△], Ruihao Huang[1,2,3], Rui Ji[1,2,3], Xianjing Cheng[1,2,3], Xiaoli Liang[1,2,3], Xiaoqi Wang[1,2,3,*], Xi Zhang[1,2,3,*]

[1] *Medical Center of Hematology, Institute of Science Innovation for Blood Ecology and Intelligent Cells, Xinqiao Hospital of Army Medical University, Chongqing 400037, China;*

[2] *Chongqing Key Laboratory of Hematology and Microenvironment, Chongqing 400037, China;*

[3] *State Key Laboratory of Trauma and Chemical Poisoning, Army Medical University, Chongqing 400037, China.*

ABSTRACT

Allogeneic hematopoietic stem cell transplantation (allo-HSCT) may be curative; however, after which graft-versus-host disease (GVHD) remains a leading cause of mortality. For the sake of high heterogeneity and the absence of widely used clinical biomarkers, clinicians are faced with obstacles when diagnosing and prognosticating GVHD. Current diagnostic approaches primarily rely on typical clinical manifestations, supplemented by laboratory indices such as inflammatory cytokine levels, often resulting in suboptimal efficacy. The emergence of machine learning (ML), however, offers a promising alternative, leveraging its unique advantages and powerful algorithms beyond traditional statistical tools. Thus, it is able to integrate existing findings on GVHD to build models for biomarker identification, diagnosis, and even prediction. In this review, we conclude present applications of ML for GVHD, exhibiting its distinctive qualitative and quantitive analytical capabilities in mining associated factors and bridging them with GVHD clinical outcomes. Additionally, we discuss future directions in this interdisciplinary field, combining the latest GVHD research frontiers with advanced bioinformatics methods, with the aim of providing guidance and insights for future progress.

Keywords: graft-versus-host disease, machine learning, artificial intelligence, allogeneic hematopoietic stem cell transplantation

INTRODUCTION

Allogeneic hematopoietic stem cell transplantation (allo-HSCT) has been practised for decades as a crucial treatment modality for various hematologic malignant and non-malignant diseases[1–2]. However, for allo-HSCT survivors, graft-versus-host disease (GVHD) remains a significant complication that affects patient mortality. This requires that clinicians be capable of early diagnosis and timely intervention to halt disease progression. Although several well-accepted approaches to GVHD prophylaxis, diagnosis and treatment currently exist[3], we still encounter challenges in addressing nonclassical manifestations

*Corresponding to: Xi Zhang and Xiaoqi Wang, Institute of Science Innovation for Blood Ecology and Intelligent Cells, Xinqiao Hospital of Army Medical University, No. 183 Xinqiao Main Street, Shapingba District, Chongqing 400037, China. E-mails: zhangxi@tmmu.edu.cn and xiaoqiwang27@gmail.com.

△These authors contributed equally to this work.

Received 10 September 2025, Revised 16 October 2025, Accepted 1 December 2025

https://doi.org/10.46701/BG.20250201160

and disease prevention due to its complex pathogenesis. Furthermore, most clinicians diagnose GVHD based on experience, which lacks accuracy and sensitivity; consequently, widely used clinical biomarkers or laboratory diagnostic methods have yet to be established. However, based on conventional research approaches and traditional statistical tools, it is challenging to achieve significant short-term progress in GVHD, a common bottleneck faced by nearly all hematologic disorders.

With the rapid development of artificial intelligence (AI), a growing number of AI techniques, especially machine learning (ML), are being applied to medical research and clinical assistance. As an interdisciplinary field integrating mathematics and computer science, ML excels in data analysis and model construction through data mining, which therefore has been widely applied in medical research for data description and predictive model building, and shows promising potential due to its robust analytical capabilities and high accuracy. Given the high heterogeneity and complexity of GVHD, ML is expected to be applied in this field through data analysis and model development for diagnosis and prediction, which may contribute to the identification of latent correlations among the intrinsic characteristics of patients with or at high risk of GVHD.

In this review, we introduce how ML can explore new dimensions for addressing the heterogeneity of GVHD and summarize its existing contributions to this field. Besides, considering the current gap between the development of ML models and their clinical application, there remains significant potential to optimize research patterns and maturing this interdisciplinary intersection. Future investigations are expected to integrate more diverse data types, such as electronic medical records, clinical examinations, and various bioinformatics, to realize multimodal exploration in GVHD diagnosis, prognosis, and prediction. Therefore, we aim to summarize the current development status and future prospects of this field through the review.

GVHD

Pathogenesis of GVHD

The pathogenesis of GVHD fundamentally adheres to Billingham's triad: immunocompetent donor cells within the graft, histocompatibility antigen disparities between donor and recipient, and the recipient's immunoincompetence that prevents donor cell elimination[4]. Acute GVHD (aGVHD) evolves through three sequential phases: conditioning-induced tissue injury and inflammatory cytokine release (*e.g.*, TNF-α, IL-1), donor T-cell activation *via* host antigen-presenting cells, and effector mechanisms involving direct cytotoxicity and cytokine storms. Chronic GVHD (cGVHD), conversely, manifests as dysregulated T-cell and B-cell interactions that drive fibrotic pathways, although its precise molecular triggers remain incompletely understood despite recent mechanistic frameworks[5].

Diagnostic challenges

Current GVHD diagnosis relies predominantly on clinical phenotyping. AGVHD presents as a triad of maculopapular rash, secretory diarrhea, and cholestatic liver injury, while cGVHD exhibits heterogeneous manifestations, such as poikiloderma, mucosal lichenoid changes, or bronchiolitis obliterans. This symptom-dependent approach faces critical limitations, including the absence of validated biomarkers, organ-specific phenotypic variability, and overlapping features between acute and chronic forms that complicate classification. Consequently, diagnostic accuracy remains suboptimal, especially in cases of early or atypical presentations, underscoring the urgent need for objective molecular signatures to complement clinical criteria[4].

Therapeutic landscape

First-line GVHD management centers on glucocorticoids; however, 40%–50% of aGVHD patients develop steroid-refractory disease requiring secondary agents such as calcineurin inhibitors or JAK inhibitors. This therapeutic hierarchy reveals significant constraints: steroid non-responsiveness correlates with higher non-relapse mortality (NRM), while intensified immunosuppression increases infectious risks without ensuring treatment efficacy. Moreover, treatment responses vary among patients, with no reliable predictors available to guide personalized regimen selection. The stark reality remains that steroid-refractory aGVHD (SR-aGVHD) carries a two-year mortality rate exceeding 50%, highlighting the systemic limitations of current therapeutic paradigms[6].

Unmet needs and future directions

Critical gaps persist across both GVHD research and clinical management. Mechanistically, the drivers of cGVHD fibrosis and the resistance pathways in SR-aGVHD require elucidation to facilitate targeted interventions. Clinically, there is an urgent need for validated biomarkers (*e.g.*, ST2, CXCL9) for early diagnosis and therapy-response prediction, as well as

risk-stratification tools to optimize immunosuppression dosing. Technologically, transitioning from reactive, symptom-driven management toward preemptive, precision-based approaches necessitates the integration of multi-omics data through advanced computational frameworks—a paradigm shift where ML holds the potential to bridge mechanistic insights with actionable clinical decision support.

Given the current situation, increasing efforts are being directed toward research on the pathogenesis, diagnosis, and prognostication of GVHD. With numerous observed risk factors including gender, age, HLA matching, immune status, diseases history, and other potential influencing variables, the complexity and challenges of studying GVHD are readily apparent[7–8]. Amid the substantial uncertainties in basic research on GVHD, clinicians also face limitations in available strategies to effectively manage the disease. Thus, there is a growing need for more powerful tools carrying strong computational and analytical capabilities, which are expected to be applied in investigating such a complicated disease.

MECHANISMS AND PROPERTIES OF ML

ML, an emerging discipline originating from computer science, is being increasingly applied in medical research[9–13]. A contemporary definition of ML is that a computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E[14]. Conventional statistical techniques are model (hypothesis)-driven, which begin with a predefined model and subsequently assess whether the data conform to it, thereby requiring certain data properties, such as normal distribution and variable independence to satisfy the assumptions of the chosen method. Furthermore, conventional statistical tools often struggle with very large datasets or datasets containing numerous variables. Thus, several potential relationships or patterns underlying the data may be overlooked during traditional statistical processing. In contrast, the paradigm underlying ML does not start with a predefined model; instead, it creates the model by detecting underlying patterns in the provided data[15]. ML is typically categorized into two types: supervised learning and unsupervised learning[16].

Supervised learning

Supervised learning aims to predict an unknown output or target by training on datasets with determined ones[16]. Similar to a recruit gaining proficiency through repeated practice, a computer analyzes labeled feature datasets and hence develops the ability to predict labels for new, unlabeled datasets (***Fig. 1***). Common supervised learning includes classification and regression. The former determines the specific type of inputs based on previously classified examples. One typical example is that our phones identify spam messages by recognizing keywords or expression patterns from pre-labeled junk messages. Regression, on the other hand, realizes prediction by estimating the relationship between the target variable and one or more feature variables, as exemplified by sales forecasting.

In the medical field, supervised learning has been utilized in a range of specialized tasks. It helps analyze basic research data, such as BIDCell, a self-supervised deep learning framework developed by Fu X *et al*. that is proficient in uncovering relationships between spatially resolved gene expression and cell morphology, thereby significantly enhancing single-cell spatial expression analysis[17]. In clinical applications, supervised learning has also demonstrated strong performance. For instance, Klauschen F *et al*. accomplished a shift from visual estimation to computational image analysis for tumor-infiltrating lymphocyte (TIL) scoring, enabling more standardized and efficient TIL quantification and facilitating clinical decision-making[18].

Unsupervised learning

Instead of output prediction, unsupervised learning identifies naturally occurring patterns or groupings within the data and performs functions such as cluster analysis, association rule learning, and dimensionality reduction based on these findings (***Fig. 1***)[16]. A classic example is customer segmentation: companies usually divide potential consumers into different groups based on specific features and unsupervised learning algorithms have the capability to generate various grouping scenarios. The value of these groupings can be evaluated by subsequent supervised learning, enabling further adjustment and optimization to achieve desired outcomes.

Unsupervised learning is most commonly applied in medicine within the field of radiomics to discover previously unknown signals and patterns related to disease evolution, progression, and treatment response underlying sufficient and mineable datasets[19]. Although not as widely adopted as supervised learning, unsupervised learning possesses unique applications and advantages, particularly when combined with supervised methods. A notable

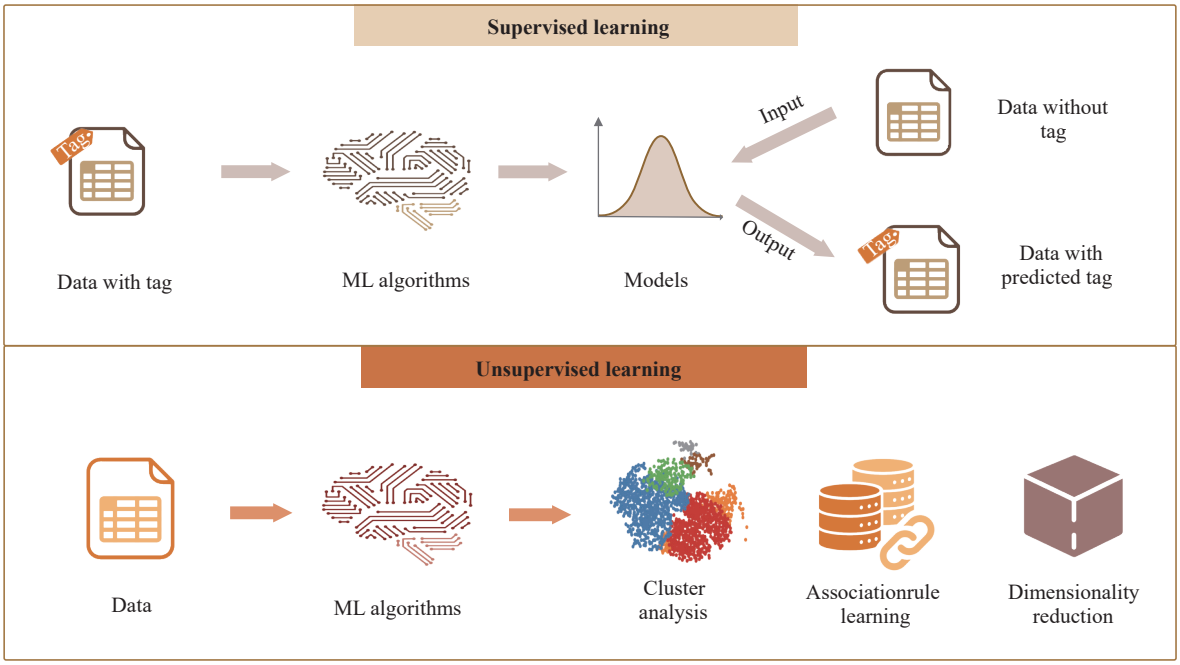

***Fig. 1* Operating mode of machine learning (ML).** Supervised learning discovers relationships between inputs (data) and outputs (tag) to build models for outcome prediction. Unsupervised learning analyzes potential patterns within input data to realize goals such as cluster analysis, association rule learning, and dimensionality reduction.

example is the work by Silburt J *et al*., who developed MORPHIOUS to identify clusters of microglial and astrocytic activation by integrating a one-class support vector machine with the density-based spatial clustering of applications with noise (DBSCAN) algorithm[20].

Model selection and adjustment

When applying ML in research, initial challenge is often selecting the most suitable ML model for the desired outcome. With multiple models available, the study's objective serves as the primary principle. The most appropriate method for selecting one from our alternatives is parallel comparison, although it can be highly time-consuming. A more practical approach for algorithm selection is to consider the volume of data and the features involved: for instance, using logistic regression and regularization when data and features are limited; tree-based models when features are relatively abundant; and neural network when the data is abundant enough. It is worth noting that support vector machines (SVMs) demonstrate considerable robustness across various situations, as evidenced by several studies[16,21]. Besides, the selected model requires substantial optimization, including hyper-parameter adjustment and feature selection, to improve performance and avoid overfitting. Several methods such as cross-validation should be considered to produce the most appropriate model.

Current situation and development of ML in medical fields

ML is revolutionizing medical research by shifting the paradigm from hypothesis-driven inquiry to data-driven discovery. Unlike classical statistical methods that rely on predefined models, ML algorithms excel at identifying complex, non-linear patterns and interactions in high-dimensional datasets without a priori assumptions about underlying distributions. This capability is particularly powerful in medicine, where multifactorial diseases involve intricate interplays among genetic, clinical, and environmental variables. By autonomously screening large-scale biomedical data, ML can uncover novel risk factors and biomarkers that might remain obscured under traditional analytical frameworks. The applications of ML in medicine are rapidly expanding across several frontiers. In medical imaging, deep learning models, particularly convolutional neural networks (CNNs), now achieve expert-level accuracy in detecting pathologies from radiographs, CT scans, and magnetic resonance imaging (MRI) scans, and can sometimes enhance low-resolution images to diagnostic quality. For instance, recent breakthroughs include AI models capable of predicting genetic mutations directly from routine histopathological slides, an approach known as "read-the-slide-to-know-the-gene". This capability may potentially reduce reliance on costly and time-consuming genetic testing in certain scenarios. In drug

discovery, ML dramatically accelerates the identification of candidate compounds and the prediction of their efficacy and toxicity, thereby streamlining the development pipeline. In the realm of personalized medicine, ML integrates multimodal data, including genomics, proteomics, and electronic health records, to forecast individual disease risk and optimize therapeutic strategies, advancing healthcare toward a more precise and proactive paradigm. Despite its transformative potential, the integration of ML into clinical practice faces significant challenges. The "black box" problem remains a major concern. The opaqueness of complex models, especially deep neural networks, can undermine clinical trust and impede adoption, as clinicians and regulators require interpretable reasoning for critical decisions. In response, research into explainable AI (XAI) is intensifying to enhance the transparency of model outputs. Additionally, the performance of ML is heavily constrained by data requirements. Model robustness demands not only large volumes of high-quality, curated data but also meticulous standardization to mitigate biases. This is compounded by issues of data privacy, security, and the logistical complexities of aggregating diverse datasets from disparate healthcare systems. Moreover, the computational resources required for training state-of-the-art models can be substantial, necessitating significant infrastructure investment. Looking ahead, the trajectory of ML in medicine points toward several promising directions. The integration of multimodal data will enable a more holistic perspective on patient health. Federated learning is emerging as a key framework for training models across multiple institutions without sharing raw patient data, thus preserving privacy while leveraging larger datasets. There is also growing emphasis on developing real-time, adaptive systems for continuous patient monitoring and predictive intervention, particularly in critical care and chronic disease management. Ultimately, the future lies not in the replacement of clinicians but in the development of a collaborative intelligence, where ML serves as a augmentative tool that enhances human expertise with data-driven insights. In the context of a complex syndrome like GVHD, which involves dynamic interactions among immune cells, tissues, and inflammatory mediators, ML offers a uniquely powerful toolkit. Its ability to deconvolute such complexity can accelerate the identification of predictive signatures, elucidate novel pathophysiological mechanisms, and ultimately contribute to more personalized risk stratification and treatment strategies.

APPLICATIONS OF ML IN GVHD

ML has been applied in hematological malignancies for decades, aiding in biomarker selection, diagnosis, and prediction[22]. As a significant complication following allo-HSCT, GVHD is also incorporated within this interdisciplinary field. In the current information era, from basic research to clinical application, ML provides new insights into GVHD with advances in the holistic medical model before, during, and after disease onset (***Fig. 2***). The following section summarizes existing research on GVHD that has introduced ML as an analytical tool and has achieved notable results (***Table 1***).

ML for selecting GVHD-associated biomarkers

Although several studies have identified GVHD-associated biomarkers such as USP17L RNA, ST2, and Elafin as potential diagnostic indicators, these discoveries primarily originate from single hypotheses based on artificially selected targets from pre-study designs or prior research, thereby lacking high-throughput screening capability and failing to adequately account for the interrelationships among these factors[23–24]. In contrast, ML, by virtue of its distinct and more powerful analytical framework, enables the data-driven discovery of a broader spectrum of potential risk factors. For instance, McCurdy SR *et al*. employed classification and regression tree (CART) analysis to construct models that elucidate relationships between biological variables and clinical outcomes[25]. Using ML, they identified risk factors, including CXCL9 and effector $CD4^+$ conventional T cells (Tconv), and further quantified their contributions to the occurrence of aGVHD following post-transplant cyclophosphamide-based myeloablative allogeneic bone marrow transplantation. The cumulative incidence curves based on the levels of CXCL9 and Tconv demonstrated the predictive efficacy of these factors for aGVHD. Notably, the study also provided profound insights into underlying mechanisms, including the CXCR3-CXCL9 axis and immune reconstitution. Similarly, Martens MJ *et al*. successfully identified five proteins (CXCL9, CXCL10, MMP3, DKK3, and CD163) associated with a high risk of cGVHD and innovatively discovered that the effects of certain biomarkers varied depending on the graft type[26]. In their work, most ML models incorporating biomarkers yielded higher time-varying area under the curve (AUCt) values for predicting cGVHD compared to models using clinical factors alone, suggesting that plasma

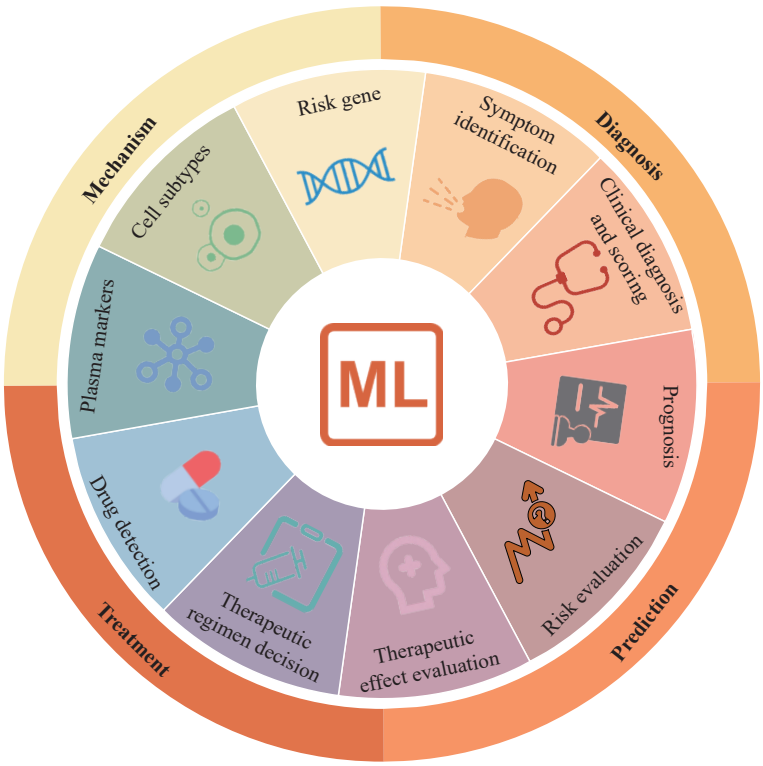

***Fig. 2* Present application situation of machine learning (ML) for graft-versus-host disease (GVHD).** From basic research like biomarker detection to diagnosis, scoring, prognosis, and even prediction and treatment, ML has been utilized in different stages of GVHD progression. At each stage, ML shows a broad application potential, encompassing various modalities.

proteins can provide early indications of underlying cGVHD biology before clinical manifestation.

The study of complex human diseases remains challenging due to multifactorial etiologies and intricate molecular mechanisms underlying genetic, genomic, and proteomic levels[27]. Given the complex pathogenesis and heterogeneous clinical manifestations of GVHD, identifying biomarkers associated with its onset and progression is crucial for improving prevention and diagnosis. ML offers a powerful approach for uncovering more reliable and clinically significant biomarkers. Furthermore, it provides novel insights into GVHD pathophysiology by revealing potential checkpoints for intervention. Existing studies further highlight ML's unique capability to quantify the contribution of specific markers to GVHD occurrence or severity, a task that is difficult to accomplish with traditional statistical methods. However, the reliability and accuracy of such quantifications require further validation[28–30].

ML for clinical diagnosis assistance and improvement

ML is increasingly being integrated into clinical decision-support systems for GVHD, demonstrating substantial potential to enhance diagnostic accuracy and facilitate early intervention. The application of ML in medical imaging has reached a relatively mature stage, with previous studies validating its effectiveness in identifying GVHD manifestations, including the differentiation of cutaneous and pulmonary complications[31–33]. Beyond distinguishing GVHD phenotypes, ML can elucidate intrinsic causal relationships between disease occurrence and associated risk factors, enabling the development of more comprehensive diagnostic models. Given the heterogeneous nature of cGVHD, Cuvelier GDE *et al*. developed a supervised support vector machine-based classifier that integrates multiple cellular and plasma biomarkers with clinical factors to diagnose pediatric cGVHD, assisting clinicians in differentiating cGVHD at its onset from other non-cGVHD manifestations[34]. The combination of selected cellular and plasma markers with clinical factors achieved an average AUC of 0.89 (± 0.07). Notably, the authors incorporated practical considerations such as time and cost during the final model optimization phase, underscoring the importance of feasibility in translating ML tools into clinical practice.

In the realm of differential diagnosis, a recent study by Elhalaby R *et al*. devised a dual-channel CNN model to automate eosinophil quantification in colonic biopsies, achieving performance comparable to pathologist assessments (F1 score: 90.7% *vs*.

Table 1 **Representative applications of ML in GVHD**

	Sample size (time window)	Data source	Prediction endpoint	Validation approach	Primary performance metrics	Key clinical factors identified	Innovation/clinical relevance
McCurdy SR *et al*. (2022)	145 patients (2019–2021)	NCT00796562/ NCT00809276 trials	aGVHD development (post-PTCy)	OOB error estimation	HR = 2.57 (high *vs*. low risk); AUC = 0.616 (Tconv+CXCL9)	Tconv cells, CXCL9, NK cells (OS protection)	First immune dynamic profiling under PTCy; linking metabolic signatures to GVHD
Cuvelier GDE *et al*. (2023)	234 pediatric patients (2013–2017)	ABLE/PBMTC 1202 prospective cohort	cGVHD diagnosis	10-fold cross-validation	AUC = 0.89 (classifier); PPV = 82%, NPV = 80%	NKregs, CXCL9/CXCL10, sCD13	Multi-site (gut/oral/nasal) biomarker integration; pediatric-specific diagnostic classifier
Arai Y *et al*. (2019)	26 695 patients (1992–2016)	Japanese Transplant Registry Unified Management Program	aGVHD (grades II–IV/III–IV)	30% independent validation cohort	C-index = 0.616 (ADTree); HR = 2.57 (high *vs*. low risk)	HLA disparity, donor age, disease risk	Largest cohort; clinically interpretable ADTree algorithm
Ingham AC *et al*. (2021)	29 children (1-year longitudinal follow-up)	36-center pediatric cohort	aGVHD severity	Leave-one-out cross-validation	Predictive taxa identified (gut/oral/nasal)	Pre-HSCT *Parabacteroides distasonis* (gut), *Actinomyces* (oral), *Rothia* (nasal)	First multi-site microbiota predictor; machine learning-based dosing protocol
Iwasaki M *et al*. (2022)	2207 patients (1996–2016)	Kyoto Stem Cell Transplantation Group	GRFS	5-fold cross-validation	C-index = 0.670 (ensemble); ΔOS = 28% (high *vs*. low risk)	Donor source, conditioning regimen, disease stage	Stacked ensemble model outperformed conventional Cox-PH (ΔC-index = 0.02)
Mo XD *et al*. (2022)	940 patients (2015–2018)	BRIGHT study (36 Chinese centers)	Steroid-refractory aGVHD response	Propensity score matching	ORR = 79.4% (D28); 3-yr OS = 64.3%	Combined therapy, grade III–IV aGVHD, Minnesota Risk Score	Real-world dosing protocol (4 doses for grade II; 5 for grades III-IV); infection risk modeling

aGVHD: acute graft-versus-host disease; cGVHD: chronic graft-versus-host disease; GVHD: graft-versus-host disease; PTCy: post-transplantation cyclophosphamide; HSCT: hematopoietic stem cell transplantation; GRFS: graft-versus-host disease-free, relapse-free survival; OOB: out-of-bag error; AUC: area under the curve; HR: hazard ratio; C-index: concordance index; PPV: positive predictive value; NPV: negative predictive value; ORR: overall response rate; OS: overall survival; ADTree: alternating decision tree; Cox-PH: Cox proportional hazards; Tconv cells: conventional T cells; NK cells: natural killer cells; NKregs: regulatory natural killer cells; CXCL9/CXCL10: C-X-C motif chemokine ligand 9/10; sCD13: soluble CD13. ΔOS (difference in OS) refers to the absolute difference in survival rates (28%) between the high-risk and low-risk groups of patients as predicted by the ML model. ΔC-index (difference in concordance index) indicates the improvement in predictive accuracy (0.02) of the novel stacked ensemble model over the conventional Cox proportional hazards (Cox-PH) model.

94.4%)[35]. The AI tool identified that the presence of neuroendocrine cell aggregates is a histological marker specific to GVHD, which was consistently absent in mycophenolate mofetil (MMF)-induced colitis ($P < 0.001$). This finding was complemented by the establishment of an eosinophil threshold (> 7.7/HPF, AUC = 0.84) indicative of MMF-induced colitis. By integrating these two distinct features, the model resolved diagnostic uncertainties in a subset of transplant patients, thereby guiding appropriate therapeutic decisions (*e.g.*, MMF discontinuation instead of unnecessary steroid escalation). This work underscores how AI can standardize the histopathologic differentiation between GVHD and drug-induced injury, directly influencing clinical management. In a similar vein, another research group employed a comparable ML methodology to develop a day-100 risk assignment algorithm for predicting subsequent cGVHD development following hematopoietic cell transplantation (HCT), successfully identifying at least two distinct cGVHD biological patterns and one associated with the establishment of immune tolerance post-HCT[36–38].

In another significant effort, Serrano-López J *et al*. developed a diagnostic (or predictive) model using a curated set of T-lymphocyte genes (including CDKN2A, SERPINB9, LYPLA, and CKTM1A/B) in cGVHD patients, which exhibited an AUC exceeding 0.75 for each defined group[39]. These findings concurrently suggested potential immune checkpoint inhibitors for subsequent cGVHD therapies, some of which have recently been reported as second-line options. Furthermore, logistic regression and two ML algorithms —LASSO regression and random

forest—were employed to analyze the distribution of five highly discriminative biomechanical parameters in sclerotic cGVHD[40]. All three analytical tools consistently demonstrated that combining frequency and relaxation time measurements yielded the highest diagnostic value. The LASSO-optimized model incorporating these variables achieved an overfit-corrected AUC of 0.87 for averages across all anatomical sites and 0.78 for averages across supine sites. However, further investigation is required to determine whether these parameters should be used individually or in combination, as bivariable analysis did not outperform univariable analyses, potentially indicating redundant information.

It is important to note that nearly all definitive cases referenced in the aforementioned studies are based on NIH consensus criteria. While these criteria assess cGVHD severity in individual organs and provide an overall severity rating, they do not integrate organ involvement to identify phenotypic clinical subgroups, potentially missing clinically relevant patterns of cGVHD manifestation. To address this limitation, Gandelman JS *et al*. applied unsupervised learning to categorize patients into seven unique clusters, leading to more significant stratification of overall survival (OS) between groups compared to stratification by NIH severity alone[41]. Subsequently, the authors developed a decision tree to classify patients into these seven clusters, further validating the clinical applicability of this ML-derived taxonomy. This demonstrates how ML augments existing methodologies allowing clinicians to attain refined diagnostic and prognostic outcomes.

In clinical practice, GVHD diagnosis primarily relies on NIH standardization. However, by the time patients definitively meet these criteria, some may have already developed irreversible organ damage, highlighting the critical importance of timely and precise detection. Previous studies have underscored the potential of AI, including ML techniques, to enhance the diagnosis of various diseases by leveraging phenotypic, genotypic, and medical imaging data[42]. For a highly heterogeneous condition such as GVHD, more comprehensive strategies are required to enable earlier and more accurate recognition. ML represents a promising approach toward this goal, with the potential to achieve superior accuracy and sensitivity. Moreover, following the initial development of such models for clinical application, rigorous evaluation of their discriminative performance and reliability remains a crucial subsequent step[43].

ML for GVHD prediction and risk assessment

Leveraging its exceptional analytical capabilities, ML offers a powerful approach for forecasting the occurrence and progression of GVHD by identifying complex, non-linear correlations within historical patient data. A seminal study by Arai Y, *et al*. employed an Alternating Decision Tree (ADTree) algorithm —selected after comparative evaluation against four other ML algorithms and one classical statistical model—to construct a predictive model for aGVHD risk using pre-transplant parameters. The model achieved AUC values of 0.616 for predicting grade II–IV aGVHD and 0.622 for grade III–IV aGVHD[44]. This ADTree-based model facilitated risk assessment *via* a cumulative prediction score, demonstrating performance superior to the traditional EBMT risk score. Notably, as one of the early applications of ML for aGVHD risk prediction, this study provided evidence for the potential superiority of ML over traditional statistical approaches and underscored the importance of algorithm selection in optimizing predictive performance.

A significant limitation of conventional predictive models is their restriction to a limited number of variables, a constraint imposed by the practical challenges of manual data entry and scoring[45]. In contrast, ML enables the integration of extensive sets of relevant variables, which is essential for effectively improving GVHD prognosis. This capability is particularly valuable given the complex clinical presentation of GVHD and the current absence of universally established biomarkers. Predictive models developed using ML algorithms, which excel at integrating validated risk factors with clinical manifestations, show promise for outperforming assessments based solely on clinical experience. Reflecting this direction, Ulschmid C *et al*. developed a more comprehensive aGVHD prediction model incorporating a wide spectrum of patient-related, disease-related, and transplant-related variables, thereby establishing a potential future standard for data-inclusive research paradigms[46]. Beyond the analysis of static variables, ML is adept at mining longitudinal data in the modern electronic health record era. For instance, Tang S *et al*. proposed an aGVHD predictive model using longitudinal vital sign data, achieving an AUC of 0.659, which represents a significant improvement over the baseline model's performance of 0.512[47]. The potential of this approach is further supported by computational analyses in preclinical models, exemplified by the study of He K *et al*., which utilized continuous body

temperature monitoring in mice for early aGVHD prediction[48]. Moreover, ML has opened new avenues for risk assessment in orphan diseases, where large datasets are inherently scarce[49].

ML's analytical capabilities also facilitates the exploration of non-traditional factors associated with GVHD. The human microbiome, for example, engages in dynamic and reciprocal interactions with local and systemic immunity, modulating host immune responses and influencing inflammatory processes[50]. Recognizing the close correlation between the microbiome and the immune system, Ingham AC *et al*. monitored microbiota dynamics in the gut, oral, and nasal cavities of pediatric allo-HSCT patients over one year. Using an ML approach, they successfully predicted aGVHD severity from pre-transplant microbiota profiles, suggesting a potential strategy for early preventive interventions[51]. This innovative study demonstrates how ML can leverage microbial community patterns to forecast GVHD, indicating future applications that integrate diverse biological data types. Complementing this, Rashidi A *et al*. identified associations between specific antibiotic regimens during the early transplant period and aGVHD incidence, suggesting considerations for antibiotic stewardship programs[52]. Building upon static and longitudinal prediction frameworks, recent research has pioneered the use of sophisticated survival analysis models integrated with multi-site, longitudinal microbiome data to achieve refined time-to-event predictions. In a landmark 2025 study, Shtossel O *et al*. applied a novel ML model named RATIO (suRvival Analysis lefT barrler IOss) to a substantial longitudinal dataset comprising over 1800 stool and saliva samples from 204 allogeneic HSCT recipients[53]. This approach uniquely addressed the challenge posed by competing risks (*e.g*., relapse, NRM) in survival analysis. Their model not only predicted the timing of seven distinct post-HSCT outcomes (including aGVHD, cGVHD, and OS) with high accuracy (median concordance index > 0.65) but also identified critical predictive time windows: early samples (weeks 1–2) were most informative for short-term complications like aGVHD, while later samples (weeks 36–70) better predicted long-term outcomes such as OS. Further reinforcing this paradigm, Artacho A *et al*. demonstrated that multimodal microbiome alterations precede GVHD development[54].

As the average age of patients undergoing allo-HSCT increases—reflecting an associated rise in risk factors and complication complexity —accurate assessment of patient risk, with GVHD being a key indicator, becomes increasingly critical[55-57]. To address the need for predicting complex composite endpoints, Iwasaki M *et al*. developed a predictive model for GVHD-free, relapse-free survival using ensemble learning. This approach initially employed various algorithms to generate multiple predictive scores, which were then integrated to produce a final, robust prediction[58]. Similarly, Reschke M *et al*. developed MatchGraft.AI, an ML-based algorithm that predicts the individual, absolute risk of aGVHD for each donor-recipient pair using pre-conditioning data, aiming to optimize donor selection[59]. Such tools have the potential to expand the pool of suitable donors and improve selection outcomes.

In summary, aGVHD, compared to cGVHD, typically involves a more defined set of risk factors, within which ML has demonstrated prominent performance in predicting occurrence and enabling precise severity stratification with high accuracy[60–62]. However, it is important to note that models incorporating a large number of variables are susceptible to overfitting, which may compromise their accuracy in practical clinical settings. To enhance the performance and generalizability of such models, experts recommend employing robust cross-validation techniques, which are anticipated to be widely adopted in future research[63]. Moreover, the integration of ML in GVHD prediction is evolving from static risk scoring to dynamic, multi-omics-informed models capable of generating personalized, time-sensitive forecasts.

ML for treatment optimization and therapeutic monitoring

Beyond its applications in disease investigation and model development, ML also demonstrates significant potential for optimizing GVHD treatment strategies and monitoring therapeutic efficacy. For instance, in a real-world analysis evaluating basiliximab for SR-aGVHD, a two-stage ML approach was implemented: first to predict the optimal drug dosage and subsequently to predict the overall response rate, infection risk, and NRM based on the estimated dosage[64]. In this study, the distributions of predicted and actual basiliximab doses were notably similar, with a mean absolute error slightly exceeding 1 mg, indicating the reasonable prediction accuracy. Furthermore, the AUCs for the three classification tasks all exceeded 0.5, demonstrating the model's effectiveness. This integrated approach, from dosage prediction to clinical efficacy assessment, provides a pathway for ML to inform optimal treatment protocols.

Beyond specific treatment decisions, another study

highlighted the positive impact of ML-derived efficacy and recovery assessment models in formulating and adapting therapeutic strategies through outcome prediction. Akahoshi Y *et al*. utilized a recursive partitioning algorithm to develop a Day-14 Mount Sinai model for predicting treatment response to aGVHD, achieving a high AUC of 0.78[65–66]. Subsequently, they reported a novel aGVHD grading system that integrated serum GVHD biomarker scores with clinical risk factors using a CART algorithm. This system improved the prediction of NRM with an AUC of 0.76, significantly outperforming models based solely on clinical symptoms at treatment initiation, such as the Minnesota risk classification. The authors concluded that such models could guide both clinical decision-making and trial design by enabling more accurate prediction of therapeutic response and long-term outcomes. ML's capacity for accurate and rapid computation facilitates a real-time feedback loop in clinical practice, allowing continuous monitoring and evaluation of ongoing treatments to guide subsequent therapeutic decisions or trial modifications.

Moreover, the integration of ML enables the refinement of existing decision-making frameworks[67]. Unlike conventional approaches that primarily consider clinical factors and patient preferences, ML incorporates additional dimensions such as biomarker dynamics and pharmacological data into a comprehensive ensemble system. This multifaceted approach is expected to yield more holistic and practically applicable treatment decisions.

PROSPECTIVES

The seminal review by Mushtaq AH *et al*. effectively established a foundation by cataloging the early applications and overarching challenges of ML in GVHD[68]. However, the field is evolving rapidly. Our review builds upon this foundation not only by updating the evidence with cutting-edge studies but also by delving deeper into the mechanistic integration and translational pathways of ML. While previous studies excelled in mapping the landscape, we aim to provide a critical appraisal of the frontier, focusing on several pivotal dimensions where recent breakthroughs are reshaping GVHD management.

First of all, it is worth mentioning that current ML applications show superior practicality and model fit in aGVHD. The primary reason we propose is that aGVHD exhibits more specific diagnostic index and relatively defined influencing factors compared to the highly heterogeneous cGVHD. Patients with aGVHD often present with lesions on the skin, gastrointestinal tract, and liver, whereas cGVHD manifestations vary across affected individuals[3]. However, this also implies a greater potential for ML to explore the mechanisms of cGVHD, which possesses enhanced analytical power and capacity to identify correlations compared to traditional methods.

Multi-omics combination and precision medicine

The increasing availability of high-throughput technologies continues to generate large volumes of omics data, capturing information across diverse yet complementary biological layers. However, most current biomarker discoveries for GVHD still derive from single-omics measurements[69–70]. While ML has been successfully applied to single-omics data (*e.g*., transcriptomics) in GVHD research, its integration with multi-omics analyses represents a logical progression to significantly enhance the sensitivity and accuracy of diagnostic and predictive models[71–72]. Unlike traditional approaches that typically focus on individual data types or manually curated correlation, ML provides powerful and novel techniques to integrate and analyze heterogeneous data types simultaneously, facilitating the discovery of previously inaccessible insights[73]. Demonstrating this potential, Sammut SJ *et al*. established a multi-omics ML predictor for breast cancer therapy response, serving as a compelling precedent for similar applications in other complex diseases like GVHD[74]. This case is not isolated, as a growing number of studies are employing ML techniques to harmoniously analyze multiple types of biological data.

Unlike traditional ML models, the integration of multi-omics data necessitates capturing complex interconnections across different biological dimensions. Consequently, researchers are increasingly adopting deep learning approaches, which are particularly adept at processing such high-dimensional inputs. Models such as graph neural networks (GNNs), transformers, and multimodal learning are particularly suited for discerning intricate relationships among diverse data types once the data have been processed and encoded into a computational format. Furthermore, scientists have developed various multimodal frameworks that successfully integrate multi-omics data across scales ranging from the individual down to the single-cell level, thereby enabling more granular bioinformatic insights[75]. This capability can be leveraged in GVHD to improve prognosis prediction and classify distinct disease endotypes based on multi-omics signal detection, moving beyond the limitations of sub-

optimal single-omics analyses[76–77]. Moreover, multi-omics integration promises to elucidate key GVHD pathogenic pathways, shedding light on disease complexity and ultimately advancing clinical practice[78–80].

Beyond refining disease classification, ML-powered multi-omics paves the way for truly precise and individualized treatment strategies, offering the potential to effectively reduce GVHD-related morbidity and mortality. Precision medicine is an emerging paradigm that focuses on understanding and treating diseases by integrating multifaceted data to inform patient-tailored clinical decisions. ML approaches applied to rich individual datasets show considerable promise for advancing the objectives of personalized medicine[81–83]. The widespread adoption of a one-size-fits-all GVHD prevention or treatment strategy may be partially aggressive or suboptimal for certain patients, potentially increasing the risk of GVHD or other complications. Given the inherent interindividual variability, tailoring initial prophylactic or therapeutic approaches based on personalized multi-omic profiles could help optimize outcomes[84]. However, integrating a patient's complete information to formulate a comprehensive clinical decision within a limited timeframe is immensely challenging for clinicians. In contemporary medicine, advanced technologies provide access to an unprecedented volume of biologic data. Co-evolving with this data explosion, computational sciences such as ML offer highly effective pathways to process and interpret the collected information. While clinicians may naturally focus on a limited set of factors during a consultation, ML systems can integrate multi-omics data, radiomics, detailed disease history, clinical manifestations, and patient-reported outcomes to support more comprehensive and appropriate diagnostic and therapeutic decisions, effectively functioning as a powerful augmentation tool similar in scope to a deeply knowledgeable family physician[85–86].

Furthermore, the paradigm of precision medicine in GVHD is rapidly evolving from static, cross-sectional assessments toward dynamic risk modeling. This emerging approach leverages longitudinal data streams from electronic health records (EHRs), such as continuously monitored vital signs, serial laboratory values, and medication administrations, to build predictive models that update in near real-time[47]. Unlike traditional static models that provide a single risk estimate at a fixed point, these dynamic systems aim to capture the evolving physiological state of the patient following HSCT. By analyzing temporal patterns and trends, ML algorithms can provide individualized, real-time early warnings of impending GVHD onset or flare-ups, potentially before overt clinical symptoms manifest. Similarly, this methodology can be applied to continuously evaluate therapeutic response, allowing for timely intervention and dosage adjustments. This capability directly addresses the clinical imperative for dynamic, individualized prediction, moving beyond a one-size-fits-all approach to truly personalized preemptive management and treatment of GVHD[87].

Big-data analysis and data mining

In the era of big data, the increasing availability of digital medical information, coupled with powerful analytical methods like ML, holds significant promise for advancing personalized medicine and improving population health management[88]. The application of ML in big data analysis provides considerable advantages in assimilating and interpreting large volumes of complex clinical data. A key strength of ML lies in its real-time updatability and autonomous learning capability, enabling the design of single- or multi-center big data models that can continuously integrate incoming patient information to support scientific research and clinical diagnostics[89]. Compared to traditional biostatistical methods, ML exhibits superior flexibility and scalability, making it well-suited for diverse tasks such as risk stratification, diagnosis, classification, and survival prediction. Its ability to analyze heterogeneous data types represents an additional major advantage.

Furthermore, ML model performance can be progressively enhanced through the analysis and integration of expanding datasets, which contributes to the development of more robust and mature systems. However, the most substantial obstacle in this domain remains the scarcity of high-quality, annotated data sources. This issue is particularly true for GVHD, where dedicated, large-scale databases are notably lacking, presenting a significant challenge for comprehensive data mining. As previously explained, most research models are based on supervised learning, which requires data with specialized labels, such as confirmed GVHD diagnoses and precise severity scores based on NIH criteria. The manual curation of such labeled data is resource-intensive, demanding expert involvement and substantial time. Consequently, a more "self-driven" data processing paradigm is highly desirable. Previous studies have demonstrated the potential of self-supervised learning, where models are first pre-trained on large, general-purpose datasets (*e.g.*, ImageNet) before being fine-tuned for specific medical tasks —an approach that

can even outperform models trained from scratch[90]. This enables a designed analytical platform to undergo continuous adjustment and optimization. More recently, the rise of semi-supervised learning techniques has provided a promising way to leverage large datasets where only a subset of entries possess target labels. Both of the aforementioned approaches could facilitate GVHD research by utilizing large-scale, general-purpose biomedical databases such as the UK Biobank (UKB), China Kadoorie Biobank (CKB), and FinnGen. Therefore, the strategic implementation of these ML algorithms in GVHD research is expected to not only expand available data sources but also significantly improve the quality and reliability of analytical outputs[91].

Clinical assistance and decision support

ML-derived clinical support tools, including diagnostic and predictive models, hold significant potential to complement clinical experience and established guidelines, thereby contributing to optimized clinical management of GVHD. By integrating these tools, clinicians can manage patients with confirmed or suspected GVHD with greater accuracy and precision than when relying solely on clinical experience. Previous studies have demonstrated the positive impact of computerized clinical decision support systems on enhancing practitioner performance[92]. A prominent example is the growing adoption of computer-aided image analysis, which alleviates the workload of radiologists amid the increasing image modalities, enhanced resolution, and growing volume of images generated by modern scanners[93]. Furthermore, the development of GVHD prediction models plays a critical role in the context of allo-HSCT. Previous studies focused on optimizing donor selection using ML incorporated GVHD risk as a key consideration[94–95]. Consequently, the integration of clinical diagnosis and prediction with ML algorithms is expected to facilitate optimal clinical practice and potentially achieve breakthroughs in the precise detection of GVHD.

Particularly for cGVHD, ML offers promise for improving the current diagnostic and prognostic landscape. Given the absence of a biomarker widely adopted in clinical practice, integrating identified markers and typical manifestations *via* ML algorithms could enhance the efficacy of cGVHD detection. For instance, a recent study proposed a tissue-specific cGVHD detection method using circulating cell-free DNA (cfDNA) and its distinct methylation patterns[96]. Given the suboptimal efficacy of relying on a single marker like cfDNA, we envisage the further introduction of ML to combine a variety of relevant markers and signals —both directly and indirectly related to cGVHD —to achieve more accurate and effective tissue localization. This would be highly valuable for clinicians in implementing timely, symptomatic treatment[97–98]. Moreover, other researchers have suggested that specific biological patterns identified by ML may be uniquely associated with organ-specific manifestations of cGVHD, potentially enabling more targeted, organ-specific therapy. Rather than depending on invasive tissue biopsies, which are currently necessary for precise diagnosis, we anticipate that ML could localize affected tissues by analyzing easily accessible biomarkers. While numerous markers, such as CXCL9 and BAFF, have been shown to correlate with cGVHD[99], there is currently no effective method to integrate existing research findings to improve cGVHD detection. The emergence and growing adoption of ML may shed new light on this research dilemma. To achieve these goals, models that integrate laboratory test indicators and clinical manifestations can be developed to assist physicians in the diagnosis and prognosis of GVHD.

In the realm of clinical treatment, ML can provide comprehensive reference insights by integrating pharmacological data, pathological information, patient-specific factors, and other relevant considerations. Similar to the real-time model developed for predicting short-term mortality in critically ill patients to support clinician decision-making[87], ML can further assist physicians in adjusting and refining strategies in a timely manner through its prognostic prediction and evaluation functions, enabled by real-time interactive capabilities[100–101]. This aligns perfectly with the vision expressed at the 2021 Transplant Complications Working Party Educational Meeting of the EBMT: "Future predictive models should address dynamic, individual predictions over the course of transplant".

Furthermore, ML holds potential for applications in pathogenic pathway detection, novel drug discovery, healthcare optimization, and beyond[102]. The technology possesses immense potential and is expected to permeate all aspects of the GVHD field (***Fig. 3***). However, it is important to note that the application of AI in clinical trials remains an area requiring careful consideration, due to the current lack of comprehensive regulations and guidelines. Thus far, AI has played a supportive role in clinical practice, with clinicians remaining the central decision-makers.

In summary, ML demonstrates great promise in both medical research and clinical practice, powered

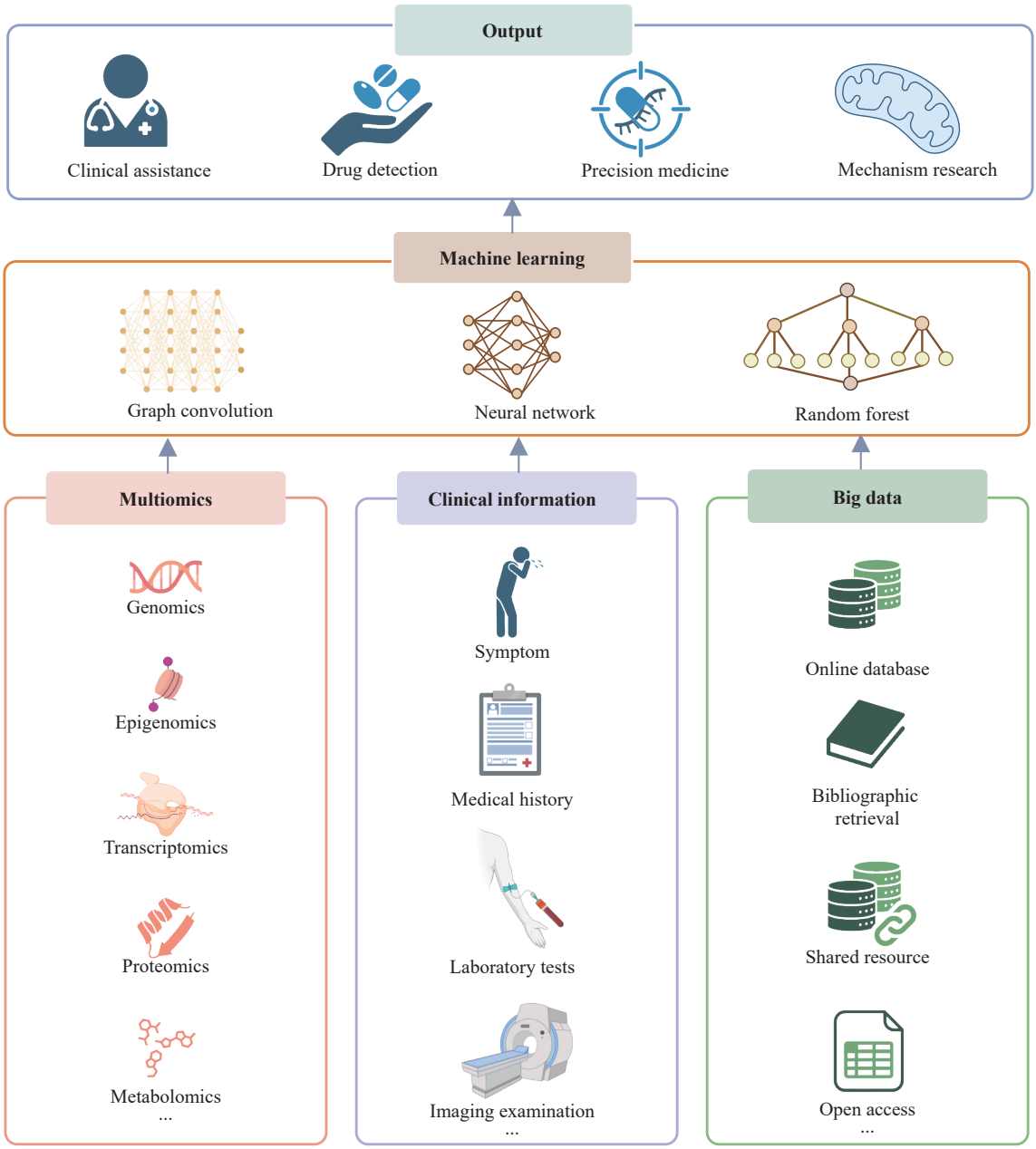

***Fig. 3* Prospects of machine learning (ML) for graft-versus-host disease (GVHD).** Through the integration and interaction of multiple aspects related to GVHD, machine learning is capable of fulfilling a range of functions and purposes, blazing a trail in the GVHD field.

by its robust and efficient algorithms. This review delineated the current landscape of ML applications within the GVHD field. Beyond its existing utility in specific tasks such as biomarker detection, disease diagnosis, and prognostication, ML is anticipated to be applied to a broader range of scenarios. Given the current challenges in managing GVHD, we have proposed several potential research directions for its diagnosis, prediction, and treatment, aiming to guide future basic and clinical investigations. Concurrently, factors critical for clinical translation and development, including the establishment of comprehensive GVHD databases, enhancement of model interpretability, and improvement of clinical acceptance, must be taken into account. As an interdisciplinary field, the integration of GVHD research with ML applications can bring new hope to patients with GVHD and individuals at high risk.

Author contributions

Writing – Original Draft Preparation, Yibo Wang and Xinwen Zhang; Visualization and Writing – Review & Editing, Ruihao Huang, Rui J, Xianjing

Cheng, and Xiaoli Liang; Supervision and Writing – Review & Editing, Xiaoqi Wang and Xi Zhang.

Ethics statement

Not applicable.

Informed consent statement

Not applicable.

Funding

This work was supported by the Noncommunicable Chronic Diseases-National Science and Technology Major Project (Grant Nos. 2025ZD0545700 and 2025ZD0545703), the Special Project for Talent Construction Program in Xinqiao Hospital (Grant Nos. 2024XKRC009 and 2022XKRC001), and the Scientific Research Cultivation Project for Undergraduates of Army Medical University (Grant No. 202490031011).

Declaration of competing interest

The authors declared no conflict of interests.

References

[1] Copelan EA. Hematopoietic stem-cell transplantation[J]. *N Engl J Med*, 2006, 354(17): 1813−1826.

[2] Wang Y, Chang YJ, Chen J, et al. Consensus on the monitoring, treatment, and prevention of leukaemia relapse after allogeneic haematopoietic stem cell transplantation in China: 2024 update[J]. *Cancer Lett*, 2024, 605: 217264.

[3] Malard F, Holler E, Sandmaier BM, et al. Acute graft-versus-host disease[J]. *Nat Rev Dis Primers*, 2023, 9: 27.

[4] Ichiki Y, Bowlus CL, Shimoda S, et al. T cell immunity and graft-versus-host disease (GVHD)[J]. *Autoimmun Rev*, 2006, 5(1): 1−9.

[5] Jagasia MH, Greinix HT, Arora M, et al. National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: I. The 2014 Diagnosis and Staging Working Group report[J]. *Biol Blood Marrow Transplant*, 2015, 21(3): 389–401. e381.

[6] Ferrara JL, Levine JE, Reddy P, et al. Graft-versus-host disease[J]. *Lancet*, 2009, 373(9674): 1550−1561.

[7] Saliba RM, de Lima M, Giralt S, et al. Hyperacute GVHD: risk factors, outcomes, and clinical implications[J]. *Blood*, 2007, 109(7): 2751−2758.

[8] Pavletic SZ, Smith LM, Bishop MR, et al. Prognostic factors of chronic graft-versus-host disease after allogeneic blood stem-cell transplantation[J]. *Am J Hematol*, 2005, 78(4): 265−274.

[9] Zhang C, Xu J, Tang R, et al. Novel research and future prospects of artificial intelligence in cancer diagnosis and treatment[J]. *J Hematol Oncol*, 2023, 16(1): 114.

[10] Wang R, Tang LV, Hu Y. Genetic factors, risk prediction and AI application of thrombotic diseases[J]. *Exp Hematol Oncol*, 2024, 13(1): 89.

[11] Zha J, Chen Q, Zhang W, et al. A machine learning-based model to predict POD24 in follicular lymphoma: a study by the Chinese workshop on follicular lymphoma[J]. *Biomark Res*, 2025, 13(1): 2.

[12] Hamamoto R, Koyama T, Kouno N, et al. Introducing AI to the molecular tumor board: one direction toward the establishment of precision medicine using large-scale cancer clinical and biological information[J]. *Exp Hematol Oncol*, 2022, 11(1): 82.

[13] Bonstingl L, Zinnegger M, Sallinger K, et al. Advanced single-cell and spatial analysis with high-multiplex characterization of circulating tumor cells and tumor tissue in prostate cancer: unveiling resistance mechanisms with the CoDuCo *in situ* assay[J]. *Biomark Res*, 2024, 12(1): 140.

[14] Mitchell TM. Machine Learning[M]. New York: McGraw-Hill, 2003.

[15] Breiman L. Statistical modeling: the two cultures (with comments and a rejoinder by the author)[J]. *Stat Sci*, 2001, 16(3): 199−215.

[16] Deo RC. Machine learning in medicine[J]. *Circulation*, 2015, 132(20): 1920−1930.

[17] Fu X, Lin Y, Lin DM, et al. BIDCell: biologically-informed self-supervised learning for segmentation of subcellular spatial transcriptomics data[J]. *Nat Commun*, 2024, 15(1): 509.

[18] Klauschen F, Müller KR, Binder A, et al. Scoring of tumor-infiltrating lymphocytes: from visual estimation to machine learning[J]. *Semin Cancer Biol*, 2018, 52(Pt 2): 151–157.

[19] Mayerhoefer ME, Materka A, Langs G, et al. Introduction to radiomics[J]. *J Nucl Med*, 2020, 61(4): 488−495.

[20] Silburt J, Aubert I. MORPHIOUS: an unsupervised machine learning workflow to detect the activation of microglia and astrocytes[J]. *J Neuroinflammation*, 2022, 19(1): 24.

[21] Greener JG, Kandathil SM, Moffat L, et al. A guide to machine learning for biologists[J]. *Nat Rev Mol Cell Biol*, 2022, 23(1): 40−55.

[22] Radakovich N, Nagy M, Nazha A. Machine learning in haematological malignancies[J]. *Lancet Haematol*, 2020, 7(7): e541−e550.

[23] Koyama M. A potential tissue-based biomarker in gut GVHD[J]. *Blood*, 2023, 142(21): 1768−1769.

[24] Paczesny S, Braun TM, Levine JE, et al. Elafin is a biomarker of graft-versus-host disease of the skin[J]. *Sci Transl Med*, 2010, 2(13): 13ra12.

[25] McCurdy SR, Radojcic V, Tsai HL, et al. Signatures of GVHD and relapse after posttransplant cyclophosphamide revealed by immune profiling and machine learning[J]. *Blood*, 2022, 139(4): 608−623.

[26] Martens MJ, Kou J, Logan BR, et al. Machine learning validates risk biomarkers of chronic graft-versus-host disease in 936 patients from BMT CTN 0201 & 1202 cohorts[J]. *Blood*, 2023, 142(Suppl 1): 479.

[27] Mi X, Zou B, Zou F, et al. Permutation-based identification of important biomarkers for complex diseases *via* machine learning models[J]. *Nat Commun*, 2021, 12(1): 3008.

[28] Sindoni MM, Toso A, Limido F, et al. Impact of ATLG on $CD4^+$ T-cell reconstitution after HSCT in children: a detailed immune profiling study[J]. *Cytotherapy*, 2025, 27(10): 1208−1218.

[29] Ma Y, Chen J, Fan Z, et al. Dynamic forecasting module for chronic graft-versus-host disease progression based on a disease-associated subpopulation of B cells: a multicenter prospective study[J]. *EBioMedicine*, 2025, 113: 105587.

[30] Loy C, Agun G, Cheng MP, et al. Plasma cell-free RNA captures immune dynamics and predicts GVHD after hematopoietic stem cell transplantation[J]. *medRxiv*, 2025.

[31] McNeil A, Parks K, Liu X, et al. Artificial intelligence recognition of cutaneous chronic graft-versus-host disease by a deep learning neural network[J]. *Br J Haematol*, 2022, 197(6): e69−e72.

[32] Sharifi H, Lai YK, Guo H, et al. Machine learning algorithms to differentiate among pulmonary complications after hematopoietic cell transplant[J]. *Chest*, 2020, 158(3): 1090−1103.

[33] Sharifi H, Bertini CD, Alkhunaizi M, et al. CT strain metrics allow for earlier diagnosis of bronchiolitis obliterans syndrome after hematopoietic cell transplant[J]. *Blood Adv*, 2024, 8(19): 5156−5165.

[34] Cuvelier GDE, Ng B, Abdossamadi S, et al. A diagnostic classifier for pediatric chronic graft-versus-host disease: results of the ABLE/PBMTC 1202 study[J]. *Blood Adv*, 2023, 7(14): 3612−3623.

[35] Elhalaby R, Sivasubramaniam P, Van Treeck BJ, et al. Mycophenolate mofetil-induced colitis versus colonic graft-versus-host disease: a comparative histologic study with artificial intelligence model development[J]. *Histopathology*, 2025, 87(4): 584−597.

[36] Rozmus J, Levine JE, Schultz KR. Planning GvHD preemptive therapy: risk factors, biomarkers, and prognostic scores[J]. *Hematology Am Soc Hematol Educ Program*, 2023, 2023(1): 149−154.

[37] Ng B, Harris AC, Abdossamadi S, et al. Distinct biological subtypes of chronic GVHD after pediatric hematopoietic cell transplantation[J]. *Blood*, 2025.

[38] Ng B, Harris AC, Abdossamadi S, et al. Identification of distinct subtypes in immune tolerance after hematopoietic cell transplantation using the prospective ABLE1.0 pediatric study cohort[J]. *Transplant Cell Ther*, 2025.

[39] Serrano-López J, Fernández JL, Lumbreras E, et al. Machine learning applied to gene expression analysis of T-lymphocytes in patients with cGVHD[J]. *Bone Marrow Transplant*, 2020, 55(9): 1668−1670.

[40] Baker LX, Chen F, Cronin A, et al. Optimal biomechanical parameters for measuring sclerotic chronic graft-versus-host disease[J]. *JID Innov*, 2021, 1(3): 100037.

[41] Gandelman JS, Byrne MT, Mistry AM, et al. Machine learning reveals chronic graft-versus-host disease phenotypes and stratifies survival after stem cell transplant for hematologic malignancies[J]. *Haematologica*, 2019, 104(1): 189−196.

[42] Lynch CJ, Liston C. New machine-learning technologies for computer-aided diagnosis[J]. *Nat Med*, 2018, 24(9): 1304−1305.

[43] Webster AP, Ecker S, Moghul I, et al. Donor whole blood DNA methylation is not a strong predictor of acute graft versus host disease in unrelated donor allogeneic haematopoietic cell transplantation[J]. *Front Genet*, 2024, 15: 1242636.

[44] Arai Y, Kondo T, Fuse K, et al. Using a machine learning algorithm to predict acute graft-versus-host disease following allogeneic transplantation[J]. *Blood Adv*, 2019, 3(22): 3626−3634.

[45] Obermeyer Z, Emanuel EJ. Predicting the future - big data, machine learning, and clinical medicine[J]. *N Engl J Med*, 2016, 375(13): 1216−1219.

[46] Ulschmid C, Li X, Wang T, et al. Development of a validated clinical risk score to predict the incidence of acute graft versus host disease after allogeneic hematopoietic cell transplantation: a CIBMTR analysis[J]. *Blood*, 2022, 140(Suppl 1): 279−281.

[47] Tang S, Chappell GT, Mazzoli A, et al. Predicting acute graft-versus-host disease using machine learning and longitudinal vital sign data from electronic health records[J]. *JCO Clin Cancer Inform*, 2020, 4: 128−135.

[48] He K, Wu Z, Fujiwara H, et al. Computational analysis of continuous body temperature provides early discrimination of graft-versus-host disease in mice[J]. *Blood Adv*, 2019, 3(23): 3977−3981.

[49] Chandra G, Wang J, Siirtola P, et al. Leveraging machine learning for predicting acute graft-versus-host disease grades in allogeneic hematopoietic cell transplantation for T-cell prolymphocytic leukaemia[J]. *BMC Med Res Methodol*, 2024, 24(1): 112.

[50] Hou K, Wu ZX, Chen XY, et al. Microbiota in health and diseases[J]. *Signal Transduct Target Ther*, 2022, 7(1): 135.

[51] Ingham AC, Kielsen K, Mordhorst H, et al. Microbiota long-term dynamics and prediction of acute graft-versus-host disease in pediatric allogeneic stem cell transplantation[J]. *Microbiome*, 2021, 9(1): 148.

[52] Rashidi A, Gao F, Fredricks DN, et al. Analysis of antibiotic exposure and development of acute graft-vs-

host disease following allogeneic hematopoietic cell transplantation[J]. *JAMA Netw Open*, 2023, 6(6): e2317188.

[53] Shtossel O, Eshel A, Fried S, et al. Microbiome-based prediction of allogeneic hematopoietic stem cell transplantation outcome[J]. *Genome Med*, 2025, 17(1): 80.

[54] Artacho A, González-Torres C, Gómez-Cebrián N, et al. Multimodal analysis identifies microbiome changes linked to stem cell transplantation-associated diseases[J]. *Microbiome*, 2024, 12(1): 229.

[55] Asteris PG, Armaghani DJ, Gandomi AH, et al. Survival prediction in allogeneic haematopoietic stem cell transplant recipients using pre- and post-transplant factors and computational intelligence[J]. *J Cell Mol Med*, 2025, 29(16): e70672.

[56] Garuffo L, Leoni A, Gatta R, et al. The applications of machine learning in the management of patients undergoing stem cell transplantation: are we ready?[J]. *Cancers (Basel)*, 2025, 17(3): 395.

[57] Rouzbahani M, Mousavi SA, Hajianfar G, et al. Predictive modeling of outcomes in acute leukemia patients undergoing allogeneic hematopoietic stem cell transplantation using machine learning techniques[J]. *Leuk Res*, 2025, 148: 107619.

[58] Iwasaki M, Kanda J, Arai Y, et al. Establishment of a predictive model for GVHD-free, relapse-free survival after allogeneic HSCT using ensemble learning[J]. *Blood Adv*, 2022, 6(8): 2618−2627.

[59] Reschke M, Gross JP, Penack O, et al. Artificial intelligence (AI) based, machine learning (ML) predicting the individual absolute risk of acute graft versus host disease (aGvHD) in a retrospective international cohort[J]. *Blood*, 2023, 142(Suppl 1): 2196.

[60] Choe H, Yuhasz N, Miller GE, et al. Advanced machine learning techniques to predict GVHD occurrence and severity with high accuracy[J]. *Blood*, 2022, 140(Suppl 1): 7658−7659.

[61] Turki AT, Peczynski C, Penack O, et al. Validating a machine learning grading system for acute GVHD. a study on behalf of the EBMT transplant complications working party[J]. *Blood*, 2021, 138(Suppl 1): 1809.

[62] Rowley SD, Gunning TS, Pelliccia M, et al. Using targeted transcriptome and machine learning of pre- and post-transplant bone marrow samples to predict acute graft-versus-host disease and overall survival after allogeneic stem cell transplantation[J]. *Cancers (Basel)*, 2024, 16(7): 1357.

[63] Charilaou P, Battat R. Machine learning models and over-fitting considerations[J]. *World J Gastroenterol*, 2022, 28(5): 605−607.

[64] Mo XD, Hong SD, Zhao YL, et al. Basiliximab for steroid-refractory acute graft-versus-host disease: a real-world analysis[J]. *Am J Hematol*, 2022, 97(4): 458−469.

[65] Spyrou N, Akahoshi Y, Kowalyk S, et al. A day 14 endpoint for acute GVHD clinical trials[J]. *Transplant Cell Ther*, 2024, 30(4): 421−432.

[66] Akahoshi Y, Spyrou N, Weber D, et al. Novel MAGIC composite scores using both clinical symptoms and biomarkers best predict treatment outcomes of acute GVHD[J]. *Blood*, 2024, 144(9): 1010−1021.

[67] Coates JT, de Koning C. Machine learning-driven critical care decision making[J]. *J R Soc Med*, 2022, 115(6): 236−238.

[68] Mushtaq AH, Shafqat A, Salah HT, et al. Machine learning applications and challenges in graft-versus-host disease: a scoping review[J]. *Curr Opin Oncol*, 2023, 35(6): 594−600.

[69] Picard M, Scott-Boyer MP, Bodein A, et al. Integration strategies of multi-omics data for machine learning analysis[J]. *Comput Struct Biotechnol J*, 2021, 19: 3735−3746.

[70] Liu X, Peng T, Xu M, et al. Spatial multi-omics: deciphering technological landscape of integration of multi-omics and its applications[J]. *J Hematol Oncol*, 2024, 17(1): 72.

[71] Albitar M, Zhang H, Pecora AL, et al. Using targeted transcriptome and machine learning of pre- and post-transplant bone marrow samples to predict acute Gvhd (aGVHD) and overall survival after allogeneic HSC transplantation[J]. *Blood*, 2022, 140(Suppl 1): 4757−4758.

[72] Albitar M, Zhang H, Pecora AL, et al. Bone marrow-based biomarkers for predicting aGVHD using targeted RNA next generation sequencing and machine learning[J]. *Blood*, 2021, 138(Suppl 1): 2892.

[73] Reel PS, Reel S, Pearson E, et al. Using machine learning approaches for multi-omics data analysis: a review[J]. *Biotechnol Adv*, 2021, 49: 107739.

[74] Sammut SJ, Crispin-Ortuzar M, Chin SF, et al. Multi-omic machine learning predictor of breast cancer therapy response[J]. *Nature*, 2022, 601(7894): 623−629.

[75] Athaya T, Ripan RC, Li X, et al. Multimodal deep learning approaches for single-cell multi-omics data integration[J]. *Brief Bioinform*, 2023, 24(5): bbad313.

[76] Poirion OB, Jing Z, Chaudhary K, et al. DeepProg: an ensemble of deep-learning and machine-learning models for prognosis prediction using multi-omics data[J]. *Genome Med*, 2021, 13(1): 112.

[77] Reel PS, Reel S, van Kralingen JC, et al. Machine learning for classification of hypertension subtypes using multi-omics: a multi-centre, retrospective, data-driven study[J]. *EBioMedicine*, 2022, 84: 104276.

[78] Liu X, Shi J, Jiao Y, et al. Integrated multi-omics with machine learning to uncover the intricacies of kidney disease[J]. *Brief Bioinform*, 2024, 25(5): bbae364.

[79] Migliozzi S, Oh YT, Hasanain M, et al. Integrative multi-omics networks identify PKCδ and DNA-PK as master kinases of glioblastoma subtypes and guide targeted cancer therapy[J]. *Nat Cancer*, 2023, 4(2): 181−202.

[80] Villiers W, Kelly A, He X, et al. Multi-omics and machine learning reveal context-specific gene regulatory activities of PML: : RARA in acute promyelocytic leukemia[J]. *Nat Commun*, 2023, 14(1): 724.

[81] MacEachern SJ, Forkert ND. Machine learning for precision medicine[J]. *Genome*, 2021, 64(4): 416−425.

[82] Nayarisseri A, Khandelwal R, Tanwar P, et al. Artificial intelligence, big data and machine learning approaches in precision medicine & drug discovery[J]. *Curr Drug Targets*, 2021, 22(6): 631−655.

[83] Plant D, Barton A. Machine learning in precision medicine: lessons to learn[J]. *Nat Rev Rheumatol*, 2021, 17(1): 5−6.

[84] Pidala JA, Gooley TA, Luznik L, et al. Chronic graft-versus-host disease: unresolved complication or ancient history?[J]. *Blood*, 2024, 144(13): 1363−1373.

[85] Infante T, Francone M, De Rimini ML, et al. Machine learning and network medicine: a novel approach for precision medicine and personalized therapy in cardiomyopathies[J]. *J Cardiovasc Med (Hagerstown)*, 2021, 22(6): 429−440.

[86] Hamilton DE, Albright J, Seth M, et al. Merging machine learning and patient preference: a novel tool for risk prediction of percutaneous coronary interventions[J]. *Eur Heart J*, 2024, 45(8): 601−609.

[87] Lim L, Gim U, Cho K, et al. Real-time machine learning model to predict short-term mortality in critically ill patients: development and international validation[J]. *Crit Care*, 2024, 28(1): 76.

[88] Weintraub WS, Fahed AC, Rumsfeld JS. Translational medicine in the era of big data and machine learning[J]. *Circ Res*, 2018, 123(11): 1202−1204.

[89] Ngiam KY, Khor IW. Big data and machine learning algorithms for health-care delivery[J]. *Lancet Oncol*, 2019, 20(5): e262−e273.

[90] Krishnan R, Rajpurkar P, Topol EJ. Self-supervised learning in medicine and healthcare[J]. *Nat Biomed Eng*, 2022, 6(12): 1346−1352.

[91] Bucci S, D'Innocente A, Liao Y, et al. Self-supervised learning across domains[J]. *IEEE Trans Pattern Anal Mach Intell*, 2022, 44(9): 5516−5528.

[92] Garg AX, Adhikari NK, McDonald H, et al. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review[J]. *JAMA*, 2005, 293(10): 1223−1238.

[93] Handelman GS, Kok HK, Chandra RV, et al. eDoctor: machine learning and the future of medicine[J]. *J Intern Med*, 2018, 284(6): 603−619.

[94] Sparapani RA, Maiers MJ, Spellman SR, et al. Is the youngest donor always the best choice to optimize outcomes for matched unrelated allogeneic transplant? Improving precision using novel statistical methodology[J]. *Blood*, 2023, 142(Suppl 1): 3552.

[95] Spellman SR, Sparapani R, Maiers M, et al. Novel machine learning technique further clarifies unrelated donor selection to optimize transplantation outcomes[J]. *Blood Adv*, 2024, 8(23): 6082−6087.

[96] Avni B, Neiman D, Shaked E, et al. Chronic graft-versus-host disease detected by tissue-specific cell-free DNA methylation biomarkers[J]. *J Clin Invest*, 2024, 134(2): e163541.

[97] Moser T, Kühberger S, Lazzeri I, et al. Bridging biological cfDNA features and machine learning approaches[J]. *Trends Genet*, 2023, 39(4): 285−307.

[98] Maimaiti A, Turhon M, Abulaiti A, et al. DNA methylation regulator-mediated modification patterns and risk of intracranial aneurysm: a multi-omics and epigenome-wide association study integrating machine learning, Mendelian randomization, eQTL and mQTL data[J]. *J Transl Med*, 2023, 21(1): 660.

[99] Sarantopoulos S, Stevenson KE, Kim HT, et al. High levels of B-cell activating factor in patients with active chronic graft-versus-host disease[J]. *Clin Cancer Res*, 2007, 13(20): 6107−6114.

[100] Zhang L, Huang T, Xu F, et al. Prediction of prognosis in elderly patients with sepsis based on machine learning (random survival forest)[J]. *BMC Emerg Med*, 2022, 22(1): 26.

[101] Li J, Qiao H, Wu F, et al. A novel hypoxia- and lactate metabolism-related signature to predict prognosis and immunotherapy responses for breast cancer by integrating machine learning and bioinformatic analyses[J]. *Front Immunol*, 2022, 13: 998140.

[102] Machine learning in drug discovery[J]. *Nat Biotechnol*, 2023, 41(7): 907.