Machine Learning Classification Algorithms to Predict aGvHD following Allo-HSCT: A Systematic Review

Cirruse Salehnasab[1] Abbas Hajifathali[2] Farkhondeh Asadi[1] Elham Roshandel[2] Alireza Kazemi[1] Arash Roshanpoor[3]

[1] Department of Health Information Technology and Management, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[2] Hematopoietic Stem Cell Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[3] Department of Computer Science, Sama Technical and Vocational Training College, Tehran Branch (Tehran), Islamic Azad University (IAU), Tehran, Iran

Address for correspondence Farkhondeh Asadi, PhD, Department of Health Information Technology and Management, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Darband St, Ghods Square, Tehran 1971653313, Iran (e-mail: asadifar@sbmu.ac.ir).

Methods Inf Med 2019;58:205–212.

Abstract

Background The acute graft-versus-host disease (aGvHD) is the most important cause of mortality in patients receiving allogeneic hematopoietic stem cell transplantation. Given that it occurs at the stage of severe tissue damage, its diagnosis is late. With the advancement of machine learning (ML), promising real-time models to predict aGvHD have emerged.

Objective This article aims to synthesize the literature on ML classification algorithms for predicting aGvHD, highlighting algorithms and important predictor variables used.

Methods A systemic review of ML classification algorithms used to predict aGvHD was performed using a search of the PubMed, Embase, Web of Science, Scopus, Springer, and IEEE Xplore databases undertaken up to April 2019 based on Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statements. The studies with a focus on using the ML classification algorithms in the process of predicting of aGvHD were considered.

Results After applying the inclusion and exclusion criteria, 14 studies were selected for evaluation. The results of the current analysis showed that the algorithms used were Artificial Neural Network (79%), Support Vector Machine (50%), Naive Bayes (43%), k-Nearest Neighbors (29%), Regression (29%), and Decision Trees (14%), respectively. Also, many predictor variables have been used in these studies so that we have divided them into more abstract categories, including biomarkers, demographics, infections, clinical, genes, transplants, drugs, and other variables.

Conclusion Each of these ML algorithms has a particular characteristic and different proposed predictors. Therefore, it seems these ML algorithms have a high potential for predicting aGvHD if the process of modeling is performed correctly.

Keywords
- machine learning
- prediction
- classification
- Allo-HSCT
- aGvHD

received
January 6, 2020
accepted after revision
February 15, 2020

DOI https://doi.org/10.1055/s-0040-1709150.
ISSN 0026-1270.

Introduction

Allogeneic Hematopoietic Stem Cell Transplantation (Allo-HSCT) has long been considered as a therapeutic approach for patients with hematologic disorders and malignancies.[1] The most important cause of mortality after transplantation is the acute graft-versus-host disease (aGvHD), which occurs in more than 35% of recipients.[2–5] aGvHD is a systemic disorder caused by allogeneic graft donor T cells and involves various organs such as skin (81%), gastrointestinal tract (54%), and liver (50%).[6–9]

The diagnosis of aGvHD is based on the patient's clinical symptoms and appears in several days to weeks after the onset of GVH when the immunologic reactions result in tissue damage in the target organs. However, many patients may develop other symptoms that interfere with the diagnosis of aGvHD, so that a combination of these factors may cause the late-stage diagnosis of aGvHD.[10]

On the other hand, after diagnosis, the severity of GvHD is determined using clinical grading systems. However, grading systems cannot predict the response to steroids.[11] Systemic steroids are the standard first-line treatment for patients with grade II or higher aGvHD. Many patients who may respond to lower doses are subjected to severe immunosuppression induced by immune-suppressive regimens and will be faced with various complications such as infection, while patients who are likely to fail treatment, receive regular doses, experience adverse effects of GvHD. Therefore, to control the aGvHD, it is necessary to identify and predict the process before the severe damage phase and the evaluation of response to treatment.[12]

In this regard, predictor variables associated with aGvHD[2] and the use of machine learning (ML) classification algorithms to evaluate the set of these predictor variables as advanced technology have been taken into consideration for prediction and response evaluation in recent years. Some researches on the prediction of aGvHD have only employed demographic and clinical variables, such as recipient gender, donor gender, recipient age, disease stage, and GvHD prophylaxis due to ease of access,[1,13,14] However, some studies have used biomarkers such as Caspase 1 (CASP1), Early Growth Response 2 (EGR2), Inducible T-cell costimulator (ICOS), Interleukin 10 (IL-10), Selectin P (SELP), Secretory Leukocyte Peptidase Inhibitor (SLPI), Signal Transducer And Activator of Transcription 6 (STAT6), B-cell lymphoma 2 Related Protein A1 (BCL2A1), Cluster of Differentiation 52 (CD52), Epidermal Growth Factor Receptor (EGFR), Interleukin 6 (IL-6), Interleukin 9 (IL-9), and Nicotinamide Phosphoribosyltransferase (NAMPT).[15–23] Also, some other studies have utilized both groups of biomarkers and demographic variables.[24,25]

Studies [1,13–25] have used ML algorithms and statistical models together or just ML algorithms to predict aGvHD.

ML is the science and art of programming computers so they can learn from past data and experiences.[26,27] ML refers to any process in which an algorithm is iteratively improved or "trained" in performing a task, usually a classification or identification task, by repeated exposure to many examples, known as the training data or training set. The trained algorithm can then be tested by measuring its performance in classifying new unseen data (the test set).[27–31]

There are four major categories: supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning.[28]

In supervised learning, the training data we feed to the algorithm include the desired solutions called labels. In unsupervised learning, the training data are unlabeled so that the system tries to learn without a supervisor.[28]

Some algorithms can deal with partially labeled training data, usually a lot of unlabeled data and a little bit of labeled data that is called semi-supervised learning.[28]

Reinforcement learning is a very different beast. The learning system called an agent in this context, can observe the environment, select and perform actions, and get rewards in return.[28]

In supervised ML, classification algorithms are applied to the prediction process. Here are some of the most crucial supervised learning algorithms including Artificial Neural Network (ANN), Support Vector Machine (SVM), Logistic Regression (LR), Naive Bayes (NB), and k-Nearest Neighbors (KNN).[28–32]

ANNs are inspired by the structure of the brain, in particular, the human visual system, making them highly useful in automated image analysis. ANN consists of many simple simulated processing units ("neurons") connected in one or more layers. Neurons receive input from preceding layers, combine the inputs according to simple summation rules, and generate an output which is fed forward to the next layer. The lowest layer of the network represents the input (e.g., variables values), while the final layer represents the output or classification. Inputs from one neuron to another are "weighted," with values analogous to synaptic weights in neural connections. As the algorithm is trained, these weights are updated according to simple feedback rules to improve the accuracy of the classification.[32,33]

SVMs apply a multidimensional transform to the input data (variables values). The algorithm then attempts to identify the hyperplane in this higher-dimensional space that best separates the training data into the desired categories (e.g., aGvHD and no aGvHD).[32,34]

KNN is an algorithm that works on the principle that data with similar characteristics will lie close to each other. For a new piece of data, the algorithm determines how close it lies with another piece of predesignated data and will then make an assumption on whether the new data has a positive or negative value.[32,35]

Considering the potential of ML in aGvHD prediction, the importance of patient safety and multiple predictors, we set out to perform a systematic review of published, real-time ML classification algorithms that predict aGvHD, highlighting algorithms and important predictor variables used.

Methods

This systematic review was conducted in 2019 based on Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statements.[36]

Eligibility Criteria

1. Inclusion criteria

This study examined a variety of studies that were eligible for inclusion in this research. These inclusion criteria include:

- Studies in which ML classification algorithms have been applied to predict aGvHD.
- Studies in which ML classification algorithms have been applied to determine the important predictors in the aGvHD prediction process.
- Studies that have applied clinical decision support systems (CDSS) based on ML classification algorithms to predict aGvHD.

2. Exclusion criteria

Studies with one or more of the following exclusion criteria were excluded from the analysis. These criteria include:

- Studies that used classical statistical methods to predict aGvHD.
- Studies that utilized ML classification algorithms to predict the effect of immune-suppressive drugs on aGvHD treatment.
- Studies that employed ML algorithms and statistical methods to predict other outcomes of Allo-HSCT or survival of patients with aGvHD.
- Studies whose full text was inaccessible.

In the cases, the abstract of papers contained the primary information, including ML classification algorithm, their evaluation indices as well as the significant predictors for the aGvHD processes. Papers that abstracts contained information according to the study objectives were not excluded from this study.

Information Sources and Search Strategy

PubMed, Embase, Web of Science, Scopus, Springer, and IEEE Xplore scientific databases before April 6, 2019 (without time limit) were explored. A combination of keywords and mesh terms associated with Allo-HSCT, aGvHD, ML, and clinical decision support system has been implemented by focusing on human studies published in English for the searching process.

The complete search strategy is presented in Supplementary Material A (online only). Besides, after evaluating full-text articles, references of the included studies were searched manually to find other suitable articles.

Study Selection and Evaluation

Three authors (FA, CS, and ER) independently examined the quality of all studies. The disagreement on whether the studies met the inclusion criteria was resolved through discussion. The reviewers also agreed on the results of all studies (Supplementary Material B [online only]).

Data items

After identifying the selected papers, using two appropriate forms designed in an Excel spreadsheet, the necessary information (such as article's primary information, applied algorithms, evaluation indices of algorithms, and essential reported predictors) was extracted from each study and recorded in these forms (Supplementary Material B (online only), ►Tables 1 and 2).

Results

Study Selection

As a result of the search strategy in selected databases, 332 articles were retrieved. Twenty-one articles were eliminated, and 266 were put aside based on reviewing the title and abstract. After examining the full-texts of 45 remaining studies, and five articles from bibliographic search based on eligibility criteria, fourteen papers were chosen for evaluation (►Fig. 1).

Study Characteristics

Exploring the extracted results illustrated that in recent years, little attention was paid to using the ML algorithm in aGvHD prediction after Allo-HSCT. Most studies have been conducted in the European Union (Italy and Spain), while the least studies have been performed in Asia (Japan).

Following the aim of the study, the results have been divided into two main categories, including the investigation of ML classification algorithms and essential predictor variables.

ML Classification Algorithms

The first step in using the algorithms when there is high dimensional data is deploying the feature extraction for dimension reduction. In some selected studies (21%), CFS-PCA methods[16,17,21] and some others (14%) multivariate analysis were applied[24,25] (►Table 1).

Selecting the validation method as the second step in applying ML algorithms is of high importance. However, the applied method is not reported in 21% of the studies.[23–25]

In these studies, the target variable was given, and the data was divided (in the 3rd step) into two sets of training and test data, applied for modeling and validation process, respectively. The training dataset was used for training purposes, and validation and the test dataset were employed for the final evaluation of the model's performance based on classified algorithms of the ML.[37] A review of the selected studies indicated that the ratio of training and test datasets was not mentioned in 21% of them.[1,22,23] In 29% of the studies,[1,13,14,23–25] the selected test data was 30% or lower (►Table 1).

The fourth step, evaluation of these classification algorithms of ML shows that 79% (11 studies) applied ANN algorithm (21%)[21,24,25] or a combination with other methods (58%).[1,15–20,23]

SVM or its combination with other methods (50%) is the second method[15,16,18–20,22,23] and then, NB algorithms and its combination with other methods (43%),[15–20] KNN and its combinations 29%,[18–20,22] different types of Regression 29%,[1,24,25] and Decision Trees 14%[13,14] were the main focus of scientists (►Table 1).

In the last step, to evaluate the performance of the applied algorithms, various evaluation indices such as Accuracy, Sensitivity, Specificity, and AUC, Positive Predictive Value, and

Table 1 Extraction of articles information based on machine learning classification algorithms and other details

Study ID	First author, year of publication, reference	Machine learning algorithms	No. of patient's	Size of training set	Size of (test set / validation set)	ML classification algorithms evaluation index
1	Fiasche et al, 2009[ax21]	ANN	59	29	30	CA training = 96%, CA test = 97%
2	Fiasche et al, 2009[ax17]	CFS-ANN and Wrapper-NB	59	29	30	CFS-ANN:CA Test = 97% Wrapper-NB:CA test = 97%
3	Fiasche et al, 2009[ax16]	CFS-ANN, PCA-ANN, Wrapper-NB, and Wrapper-SVM	59	29	30	CFS-ANN:CA test = 97%
4	La Nasa et al, 2009[cy25]	ANN and LR	78	68	10	For ANN (CS = 83.3% and CSP = 90.1%) For LR (CS = 21.7% and CSP = 80.5%)
5	Caocci et al, 2010[cy24]	ANN and LR	78	68	10	For ANN (CS = 83.3% and CSP = 90.1%) For LR (CS = 21.7% and CSP = 80.5%)
6	Fedele et al, 2010[cz23]	integrating: MLR, SVM, ES-NN, and DENFIS	66	–	–	CA > 92%
7	Fiasche et al, 2010[az20]	CFS-ANN, Wrapper-SVM, PMGS-NB, and PMGS-WKNN	59	29 + 3, 29 + 4	30 + 4 30 + 3	CA for all classification models = 97%
8	Fiasche et al, 2010[az19]	integrated approach: CFS-ANN, Wrapper-SVM, PMGS-NB, and PMGS-WKNN	59	29	30:3(7), 4(7), 3(7), 4(7), 6(7), 6(7)	CA for all classification models = 97%
9	Fiasche et al, 2011[az18]	CFS-ANN, Wrapper-SVM, PMGS-NB, PMGS-WKNN, (i)PMGS- NB and (i)PMGS- WKNN	59	29	30	CA for all classification models = 97%
10	Fiasche et al, 2011[az15]	CFS-EFuNN, Wrapper-SVM, Wrapper-NB, and Wrapper- EFuNN	78	29	30	EFuNN-1—training data: CA = 97.4 EFuNN-1—test data: CA = 97.0 EFuNN-2—training data: CA = 95.0 EFuNN-2—test data: CA = 97.2
11	Paun et al, 2013[bz13]	CART		(292, 75%)	(95, 25%)	Allogeneic decision model tree training: CS = 63%, CSP = 70%, PPV = 24%, NPV = 93%; validation: CS = 28%, CSP = 68%, PPV = 17%, NPV = 80%
12	Cocho et al, 2015[az22]	SVM, SDA, and KNN	34	–	–	SVM: AUC = 99/5, CS = 100, CSP = 92/9 SDA: AUC = 95/9, CS = 92/9, CSP = 92/9 KNN: AUC = 92/9, CS = 92/9, CSP = 92/9
13	Arai et al, 2018[bz14]	ADTree	26695	70% (17,244)	30% (8,050) Validation	AUC: 0.616 and 0.623 for grades 2–4 and 3–4,
14	Lee et al, 2018[bz1]	Ensemble: FIRST STAGE (7 kinds of regression, Bayesian additive, and ANN). SECOND STAGE: LR		9651	–	AUCs' of models ranged from 0.613 + 0.640

Validation Method: [a]-LOOCV, [b]-K-fold cross-validation, and [c]-no validation. See also Supplementary Table S4 in Supplementary Material C (online only) for other abbreviations.
Features extraction: [x]-CFS-PCA, [y]-multivariate analysis, & [z]-nonfeatures extraction.

Negative Predictive Value are reported. In 57% of the studies, only the CA evaluation index has been employed.[15–21,23]

The reported quantity of each index related to different algorithms is presented in ►Table 1.

Important Predictor Variables

Biomarkers, according to Biomarkers Definitions Working Group, are a characteristic that is defined as a natural biological processes index, pathogenic processes, or drug

Table 2 Extraction of articles information based on important predictors

Study ID	First author, year of publication, reference	No. of predictors (variable /biomarker)	Important predictor[a]
1	Fiasche et al, 2009[21]	47 biomarkers	C3[1], CASP1[6], CCR4,[1] CD52[4], EGRI[1], EGR2[6], FOS[1], IL-10[7], ILI2A,[1] IL-6[2], IRFI[1], IRF7[1], MMP9[1], NFKB2[1], NOS2A,[1] PIASI[1], PTN[1], SELP[6], SER3[1], SER4[1], SLPI[4], STAT6[6], VEGF[1], BCL2A1[5], CCL7[5], CD83[5], CXCL10[5], FAS[5], ICOS[6], IL-4[5], SLP1[2], Foxp-3[3], CXCL1[5], CD52[4], IL-9[1], CCL24[1], CCL18[1], IFN-c[1], CCL2[1], EGFR[1], and NAMPT[1]
2	Fiasche et al, 2009[17]	47 biomarkers	
3	Fiasche et al, 2009[16]	47 biomarkers	
4	Cocho et al, 2015[22]	84 biomarkers	
5	Fedele et al, 2010[23]	47 biomarkers	
6	Fiasche et al, 2010[20]	47 biomarkers	
7	Fiasche et al, 2010[19]	47 biomarkers	
8	Fiasche et al, 2011[18]	47 biomarkers	
9	Fiasche et al, 2011[15]	47 biomarkers	
10	Caocci et al, 2010[24]	24 variables	**Demographic:** Patient gender, patient age, donor gender, donor age, male recipient/female donor, and male recipient/female donor, Pesaro risk class 1, and Pesaro risk class 2 **Infection:** CMV (serology, positivity) **Clinical:** Conditioning regimen with ATG **Genes:** HLA class I mismatching, presence of patient HLA-11 positivity, HLA-DPB1 no-permissive mismatch, patient KIR ligands C1/C2, Recipient KIR (activatory/inhibitory), donor KIR (activatory/inhibitory), donor homozygosity for KIR A haplotype, patient HLA-G 14-basepair, and donor HLA-G 14-basepair
11	La Nasa et al, 2009[25]		
12	Paun et al, 2013[13]	15 variables	**Transplantation:** Time from T0, graft type[2], and previous transplant Date **Demographic:** Recipient gender[2], donor gender[2], donor's age[2], recipient age, diagnosis, donor's blood type and Rh factor, donor's race, donor's weight, donor source, type (AML, ALL, MDS or CML) donor-patient female-male sex-mismatch, and patient-unrelated donor, sex mismatch, underlying disease, and Karnofsky score **Genes:** HLA compatibility, low resolution typing, DPB1 permissivity, HLA mismatch, and HLA-compatibility (8/8 or ⅞ HLA-matched) **Drugs:** TBI, GVHD prophylaxis[2] regime, and conditioning intensity (myeloablative or reduced-intensity/non-myeloablative), and Datein-vivo T cell depletion (no or yes) **Infectious:** Donor's CMV status and patient-donor CMV serology match **Other:** Disease stage, donor contact, HCV RNA positivity, median CD34 cell dose infused, and clinical impression
13	Arai et al, 2018[14]	–	
14	Lee et al, 2018[1]	12 variables	

[a]The superscripts above the biomarkers demonstrate the number of replicates in the studies.

response to therapeutic interruptions that are measured and evaluated directly.[38] As previously stated, to reduce dimensional data, feature extraction is applied, and most critical useful variables are picked out.[16,17,21,24,25] Overall, the analyzed variables of the current study are divided into two categories of biomarkers,[15–23] variables,[1,13,14] and mainly demographic-clinical variables according to oncologists and hematology specialists. The important reported biomarkers are IL-10, ICOS, STAT6, SELP, EGR2, and CASP1 that are mentioned in six or more studies. However, some biomarkers such as NAMPT, EGFR, C3, and CCL2 have only been reported in one study.[22] Different kinds of variables such as graft type, recipient gender, and low-resolution typing were reported in the studies. According to ER and AHF specialists' decisions, these variables were classified into categories such as transplantation, infection, and drugs (►**Table 2**).

Discussion

Summary of Evidence

The current study has focused explicitly on the usage of ML algorithms and their important predictors. For this purpose, after a thorough investigation, only 14 articles matched our inclusion criteria. The few available studies could be due to three reasons. One is that ML algorithms results are not coupled with clinical reasoning.[39] Next is the Black-box rule deduction mechanism of most ML algorithms,[39,40] and third is time and cost-consuming processes of obtaining some biomarkers.[41] One way to help to reduce the costs is by using a feature extraction method separately or as a combination with other methods to reduce dimensional data to eliminate the less effective predictors. One of these methods is the simplified PCA approach.[16,17,21]

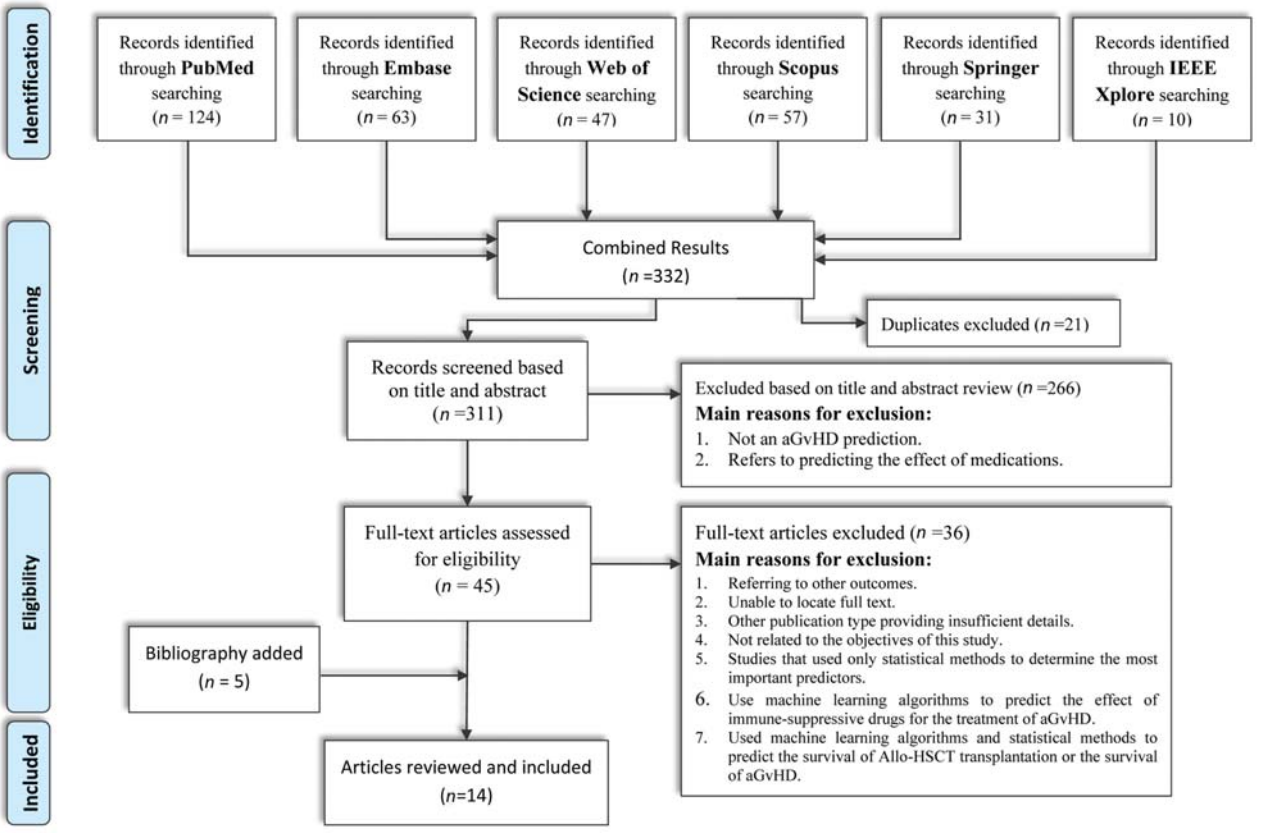

Fig. 1 Flow diagram of the included and excluded studies.[36]

Examining the results of the selected articles indicates that different predictors used as ML algorithms inputs have a distinct relationship with their evaluation indices in a way that in nine studies[15–23] out of fourteen, only genetic biomarkers are deployed, and the results demonstrate that the applied algorithm evaluation indices were above 90% whereas, in studies with mainly demographic-clinical variables,[1,13,14] it was under 70%. In two other studies,[24,25] in which genetic biomarkers and demographic-clinical variables were involved, the evaluation indices were in the middle range. Thus, it can be concluded that regardless of the type of algorithms, using biomarkers can noticeably enhance the accuracy of aGvHD prediction after Allo-HSCT. Although, in recent studies, demographic-clinical variables were cost-efficient and favoured. Hence, in future studies, it would be better to choose biomarkers after considering their cost.

Additionally, more common biomarkers should be selected, such as CASP1, IL-10, STAT6, and demographic-clinical variables. Analyzing the functional indices used to evaluate ML algorithms in the chosen studies illustrates that these indices mostly depend on the type of predictors and we cannot decide on the efficiency of algorithms merely based on a high quantity of these indices. Therefore, according to the aim of the study (extraction of the most appropriate classification algorithms of ML), most important algorithms should be introduced and then evaluated based on their strength and weaknesses.

Analysis of the results (►Table 1) revealed that SVM, ANNs, Decision Tree, and NB algorithms were famous in predicting aGvHD after Allo-HSCT. SVM algorithm works well with noisy data and provides a more accurate classification mechanism for border-line data compared with other classifier algorithms and can be entirely generalized. However, its downside is the interpretation of the generated model and its sensitivity to the proper parameters setup.[40,42] ANN algorithms, like SVM, work well with noisy data, can present linear and non-linear functions of compound structures, and process black-box data. They are also strongly dependent on setting the input parameters. NB is one of the other famous ML algorithms that has some advantages such as being fast to train, not being sensitive to irrelevant features, handling real and discrete data and streaming data properly. However, this algorithm assumes the independence of the features.[43] The latest algorithm of interest is a Decision Tree which, compared with the algorithms as mentioned earlier, not only has a high capability to distinguish final classification attributes but also, its deducted rules can be easily interpreted (white-box).[44–47]

Each of these ML algorithms has a particular characteristic and different proposed predictors. Therefore, it seems these ML algorithms have a high potential for predicting aGvHD if the process of modelling is performed correctly. Finally, applying these algorithms in the design of clinical decision support systems, not only provides awareness and guarantee for improving the therapeutic process, but it can

also assist the healthcare team, patients and their families in the process of clinical decision making.

Limitations

The first is that there are very few published articles on ML algorithms for predicting aGvHD after Allo-HSCT. To address this problem, we searched many databases such as PubMed, Embase, Web of Science, Scopus, Springer, and IEEE Xplore, as well as manually searching in other relevant databases (e.g., Google Scholar). The second limitation was that we did not publish a protocol in advance, so a systematic analysis was performed to evaluate every step of this study.

Authors' Contributions

All authors made significant contributions to the manuscript. CS developed the design of the systematic review and was involved in the data screening and extraction with FA, conducted the medical evaluation of the included studies, and wrote the manuscript. ARP and AK were involved in the medical assessment of the included studies. FA supervised and guided the project. ABH and ER categorized the biomarkers and variables that extracted from findings. All authors provided critical revision and approved the manuscript.

Conflict of Interest

None declared.

Acknowledgments

The authors would like to thank the reviewers for their valuable comments and suggestions to improve the quality of the paper. This study was part of a PhD project conducted at Shahid Beheshti University of Medical Sciences, Tehran, Iran, that was approved by Iran National Committee for Ethics in Biomedical Research with Approval ID IR.SBMU.REC.1397.128 in 2019–03–03.

References

1 Lee C, Haneuse S, Wang H-L, et al. Prediction of absolute risk of acute graft-versus-host disease following hematopoietic cell transplantation. PLoS One 2018;13(01):e0190610
2 Ali AM, DiPersio JF, Schroeder MA. The role of biomarkers in the diagnosis and risk stratification of acute graft-versus-host disease: a systematic review. Biol Blood Marrow Transplant 2016;22 (09):1552–1564
3 Bacigalupo A. Acute graft-versus-host disease. Immunotherapy 2011;3(12):1419–1422
4 Gale RP, Bortin MM, van Bekkum DW, et al. Risk factors for acute graft-versus-host disease. Br J Haematol 1987;67(04): 397–406
5 Gopalakrishnan R, Jagasia M. Pathophysiology and Management of Graft-Versus-Host Disease. Hematopoietic Cell Transplantation for Malignant Conditions. Elsevier; 2019:301–319. Available at: https://doi.org/10.1016/B978-0-323-56802-9.00022-5
6 Ferrara JL, Levine JE, Reddy P, Holler E. Graft-versus-host disease. Lancet 2009;373(9674):1550–1561
7 Blazar BR, Murphy WJ, Abedi M. Advances in graft-versus-host disease biology and therapy. Nat Rev Immunol 2012;12(06): 443–458
8 Ball LM, Egeler RM; EBMT Paediatric Working Party. Acute GvHD: pathogenesis and classification. Bone Marrow Transplant 2008; 41(Suppl 2):S58–S64
9 Martin PJ, Schoch G, Fisher L, et al. A retrospective analysis of therapy for acute graft-versus-host disease: initial treatment. Blood 1990;76(08):1464–1472
10 Iqbal N, Salzman D, Lazenby AJ, Wilcox CM. Diagnosis of gastrointestinal graft-versus-host disease. Am J Gastroenterol 2000;95 (11):3034–3038
11 Bolaños-Meade J, Jacobsohn DA, Margolis J, et al. Pentostatin in steroid-refractory acute graft-versus-host disease. J Clin Oncol 2005;23(12):2661–2668
12 Martin PJ, Rizzo JD, Wingard JR, et al. First- and second-line systemic treatment of acute graft-versus-host disease: recommendations of the American Society of Blood and Marrow Transplantation. Biol Blood Marrow Transplant 2012;18(08): 1150–1163
13 Paun O, Phillips T, Fu P, et al. Cutaneous complications in hematopoietic cell transplant recipients: impact of biopsy on patient management. Biol Blood Marrow Transplant 2013;19(08): 1204–1209
14 Arai Y, Kondo T, Fuse K, et al. Prediction of acute graft-versus-host disease following allogeneic hematopoietic stem cell transplantation using a machine learning algorithm. Blood 2018;132:3
15 Fiasché M, Verma A, Cuzzola M, Morabito FC, Irrera G, Eds. Incremental – Adaptive – Knowledge Based – Learning for Informative Rules Extraction in Classification Analysis of aGvHD. Berlin, Heidelberg: Springer Berlin Heidelberg; 2011
16 Fiasché M, Verma A, Cuzzola M, Iacopino P, Kasabov N, Morabito FC, Eds. Discovering Diagnostic Gene Targets and Early Diagnosis of Acute GVHD Using Methods of Computational Intelligence over Gene Expression Data. Berlin, Heidelberg: Springer Berlin Heidelberg; 2009
17 Fiasché M, Cuzzola M, Verma A, et al. Methods of artificial intelligence applied to gene expression data for early diagnosis of acute graft vs host disease. Haematologica 2009;94:43
18 Fiasché M, Cuzzola M, Irrera G, Iacopino P, Morabito FC. Advances in medical decision support systems for diagnosis of acute graft-versus-host disease: Molecular and computational intelligence joint approaches. Front Biol China 2011;6(04):263–273
19 Fiasche M, Cuzzola M, Iacopino P, Kasabov N, Morabito F. Personalized modeling based gene selection for acute GvHD gene expression data analysis: a computational framework proposed. Aust Intell Inf Process Syst Machine Learn Appli 2010;12(4 Part II):13–18
20 Fiasché M, Cuzzola M, Fedele R, Iacopino P, Morabito FC, Eds. Machine Learning and Personalized Modeling Based Gene Selection for Acute GvHD Gene Expression Data Analysis. Berlin, Heidelberg: Springer Berlin Heidelberg; 2010
21 Fiasche M, Cuzzola M, Cacciola M, et al , Eds. A neural network model for early diagnosis of acute GVHD based on gene expression data. 2009. IEEE Int Workshop Genomic Signal Process Stat 2009;2009:17–21
22 Cocho L, Fernández I, Calonge M, et al. Gene expression-based predictive models of graft versus host disease-associated dry eye. Invest Ophthalmol Vis Sci 2015;56(08):4570–4581
23 Fedele R, Cuzzola M, Fiasché M, et al. An evolving intelligent system for diagnostic gene targets of acute GvHD: an integrated approach. Bone Marrow Transplant 2010;45:S124–S5
24 Caocci G, Baccoli R, Vacca A, et al. Comparison between an artificial neural network and logistic regression in predicting acute graft-vs-host disease after unrelated donor hematopoietic stem cell transplantation in thalassemia patients. Exp Hematol 2010;38(05):426–433
25 La Nasa G, Caocci G, Baccoli R, et al. Artificial neural network and logistic regression in predicting acute graft-versus-host disease following unrelated hematopoietic stem cell transplantation. Haematologica 2009;94:99

26 Berner ES, La Lande TJ. Overview of Clinical Decision Support Systems. Clinical Decision Support Systems. New York, NY: Springer; 2007:3–22
27 Syeda-Mahmood T. Plenary talk: the role of machine learning in clinical decision support. SPIE Newsroom. 2015. doi:10.1117/2.3201503.29
28 Géron A. Hands-On Machine Learning with Scikit-Learn, Keras, and Tensor Flow: Concepts, Tools, and Techniques to Build Intelligent Systems. O'Reilly Media, Inc., Sebastopol, CA; 2019
29 Alpaydin E. Introduction to Machine Learning. Massachusetts London, England: MIT Press; 2014
30 Raschka S, Mirjalili V. Python Machine Learning. Packt Publishing Ltd; 2017
31 Géron A. Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, tools, and Techniques to Build Intelligent Systems. " O'Reilly Media, Inc."; 2017
32 Han J, Pei J, Kamber M. Data Mining: Concepts and Techniques. Elsevier; 2011
33 Buscema PM, Massini G, Breda M, Lodwick WA, Newman F, Asadi-Zeydabadi M. Artificial Neural Networks. Artificial Adaptive Systems Using Auto Contractive Maps. Springer International Publishing; Imprint: Springer; 2018:11–35
34 Rau C-S, Wu S-C, Chuang J-F, et al. Machine learning models of survival prediction in trauma patients. J Clin Med 2019;8(06):799
35 Ghaneei M, Ekyalimpa R, Westover L, Parent EC, Adeeb S. Customized k-nearest neighbourhood analysis in the management of adolescent idiopathic scoliosis using 3D markerless asymmetry analysis. Comput Methods Biomech Biomed Engin 2019;22(07):696–705
36 Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 2009;151(04):264–269, W64
37 Golbraikh A, Tropsha A. Predictive QSAR modeling based on diversity sampling of experimental datasets for the training and test set selection. Mol Divers 2002;5(04):231–243
38 Group BDW, Atkinson AJ Jr, Colburn WA, et al; Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther 2001;69(03):89–95
39 Sendak M, Gao M, Nichols M, Lin A, Balu S. Machine learning in health care: a critical appraisal of challenges and opportunities. EGEMS (Wash DC) 2019;7(01):1
40 Lorena AC, Jacintho LF, Siqueira MF, et al. Comparing machine learning classifiers in potential distribution modelling. Expert Syst Appl 2011;38(05):5268–5275
41 Yu J, Parasuraman S, Shah A, Narayanan S. Influence of Acute Graft-Versus-Host Disease (aGVHD) during Allogeneic Hematopoietic Stem Cell Transplantation (HSCT) Engraftment Admission on Hospital Length of Stay (LOS), Charges, and Sosts. American Society of Clinical Oncology; 2017;35(15_suppl):e18544-e. Available at: https://ascopubs.org/doi/abs/10.1200/JCO.2017.35.15_-suppl.e18544
42 Ben-Hur A, Ong CS, Sonnenburg S, Schölkopf B, Rätsch G. Support vector machines and kernels for computational biology. PLOS Comput Biol 2008;4(10):e1000173
43 Keogh E. Naive Bayes classifier. Accessed 2006;1:2018
44 Dreiseitl S, Ohno-Machado L. Logistic regression and artificial neural network classification models: a methodology review. J Biomed Inform 2002;35(5-6):352–359
45 Shouval R, Labopin M, Bondi O, et al. Prediction of allogeneic hematopoietic stem-cell transplantation mortality 100 days after transplantation using a machine learning algorithm: a European Group for Blood and Marrow Transplantation Acute Leukemia Working Party retrospective data mining study. J Clin Oncol 2015;33(28):3144–3151
46 Labopin M, Bondi O, Shamay HM, et al. Prediction of Allogeneic Hematopoietic Stem Cell Transplantation (allo-HSCT) Related Mortality in Acute Leukemia: Generation of a Machine Learning-Based Model Using the Data Set of The Acute Leukemia Working Party (ALWP) of The EBMT. Am Soc Hematology; 2013
47 Shouval R, Bondi O, Mishan H, Shimoni A, Unger R, Nagler A. Application of machine learning algorithms for clinical predictive modeling: a data-mining approach in SCT. Bone Marrow Transplant 2014;49(03):332–337