Review

Machine Learning Applications in the Diagnosis of Benign and Malignant Hematological Diseases

Ibrahim N. Muhsen[1], David Shyr[2], Anthony D. Sung[3], Shahrukh K. Hashmi[4,5,*]

[1]*Department of Medicine, Houston Methodist Hospital, Houston, TX, USA*

[2]*Division of Stem Cell Transplantation and Regenerative Medicine, Stanford School of Medicine, Palo Alto, CA, USA*

[3]*Department of Medicine, Division of Hematologic Malignancies and Cellular Therapy, Duke University School of Medicine, NC, USA*

[4]*Department of Medicine, Mayo Clinic, Rochester, MN, USA*

[5]*Department of Medicine, Sheikh Shakbout Medical City, Abu Dhabi, UAE*

ARTICLE INFO

ABSTRACT

The use of machine learning (ML) and deep learning (DL) methods in hematology includes diagnostic, prognostic, and therapeutic applications. This increase is due to the improved access to ML and DL tools and the expansion of medical data. The utilization of ML remains limited in clinical practice, with some disciplines further along in their adoption, such as radiology and histopathology. In this review, we discuss the current uses of ML in diagnosis in the field of hematology, including image-recognition, laboratory, and genomics-based diagnosis. Additionally, we provide an introduction to the fields of ML and DL, highlighting current trends, limitations, and possible areas of improvement.

1. INTRODUCTION

The fields of machine learning (ML), deep learning (DL), and artificial intelligence (AI) are speculated to change the way that we practice medicine [1–3]. The current medical literature in ML's and DL's use is exponentially expanding; however, most applications of these technologies are still limited, with a lack of prospective and real-life applications of the most proposed algorithms [4]. The integration of ML and DL in clinical care is faced with multiple ethical and logistic concerns [5]. However, it is reasonable to say that the field is promising, given the considerable amount of data generated by the current healthcare system and the ability of ML and DL to analyze, process, and identify patterns.

The field of hematology (benign and malignant) is also rapidly expanding, including increasing our molecular/genomic understanding and the number of new therapies joining the pipeline [6]. For instance, genomics has become an integral part of oncology, where it provides diagnostic, prognostic, and therapeutic values [7]. Using deep/machine learning in genomics might help overcome the challenge that a large amount of data imposes, leading to improved diagnosis, identification of genomic patterns, and increased efficiency. The integration of these newer technologies in hematology practice holds an opportunity to improve it.

In this review, we introduce AI tools, particularly ML and DL, in hematology diagnosis. We summarize the significant directions of ML and DL in the current hematology literature and hematopoietic cell transplant.

2. BASICS OF AI, ML, AND DL

2.1. Basic Definitions

Artificial intelligence is a general term that describes the use of technology in accomplishing tasks that would usually need human intelligence, for example, voice or image recognition. Machine learning, on the other hand, is a subset of artificial intelligence. The ML technology enables machines to learn from previous data using statistical approaches and algorithms [8]. The performance of ML algorithms generally improves as more data are used.

Machine learning has two main categories: Supervised and unsupervised. Supervised ML is the most widely used in the medical literature. It uses labeled input and output, which can then be applied to an algorithm to create a function that relates between the input and the output, i.e., $y = f(x)$, where y is output, and x is the input. Supervised ML is usually used to accomplish two main tasks: classification (e.g., support vector machine, naïve Bayes classifier) and regression (e.g., non-linear regression, Bayesian linear regression) [8,9]. Additionally, the incorporation of the ensembles method (combining multiple algorithms to produce a more accurate model) to create decision trees (such methods include random forest, XGBoost Bagging) is another important function of such algorithms.

On the other hand, unsupervised machine learning is used to a lesser extent in medicine. Unlike the supervised process, output

data are not needed in unsupervised learning. In addition to supervised and unsupervised learning, semi-supervised learning (or reinforcement) is another category of machine learning that uses a combination of labelled and unlabeled data [8,9]. This approach power stems from its ability to adapt through rewards-like systems to more complex environments. Figure 1 illustrates the different types of machine learning with examples on each.

Deep learning is a subset of ML which approaches problems in a fashion that is more similar to the human approach. In deep learning, algorithms can deduce the data's features/patterns

through multiple layers of processing (i.e., neural networks) [10]. Two very frequently used neural network approaches include convolutional neural network (CNN) and recurrent neural networks (RNN). CNN is a feed-forward neural network method that is mainly used in medical imaging processing, including pathologic diagnosis of hematologic cancers. In RNN, outputs of prior steps are fed as inputs in subsequent steps. This approach is mainly used in time series analysis and sequence classification (Figure 2). Similar to ML, DL algorithms can use labeled and unlabeled data. The DL popularity has increased in the medical literature in recent years.
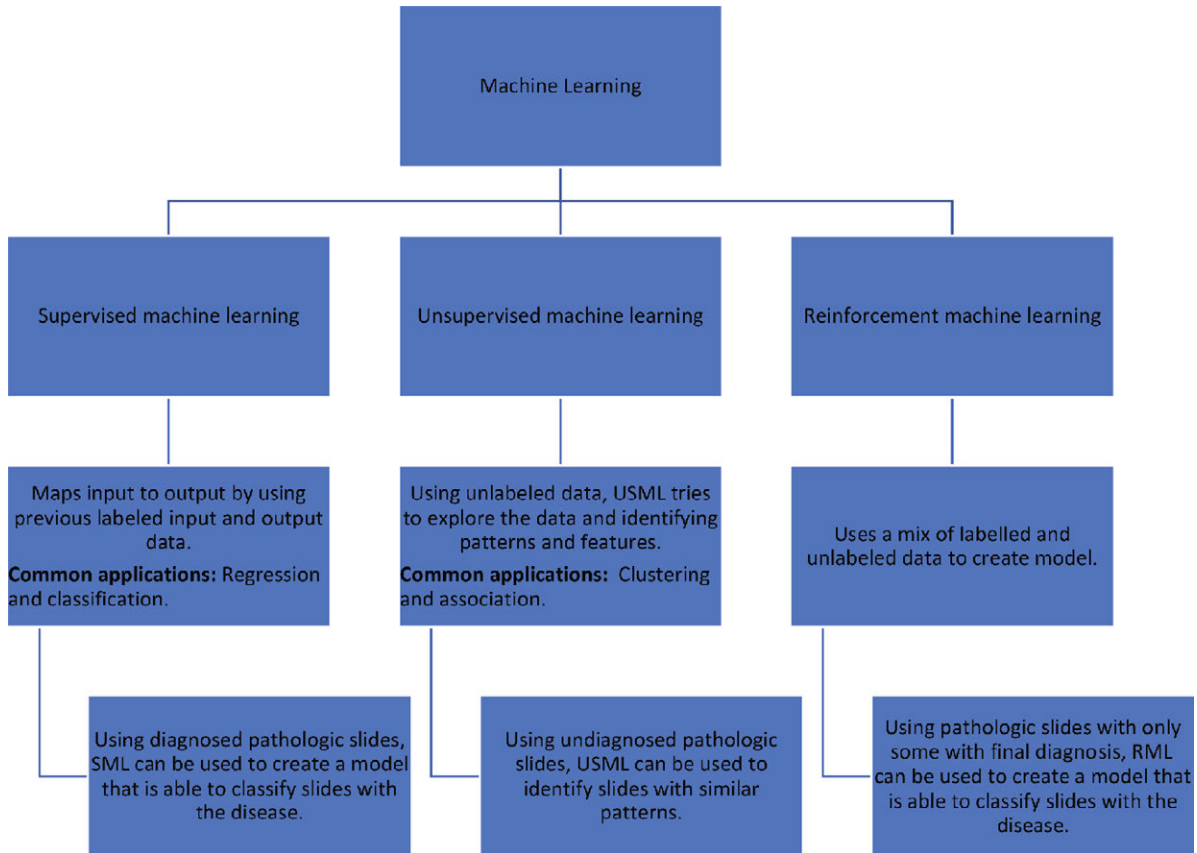

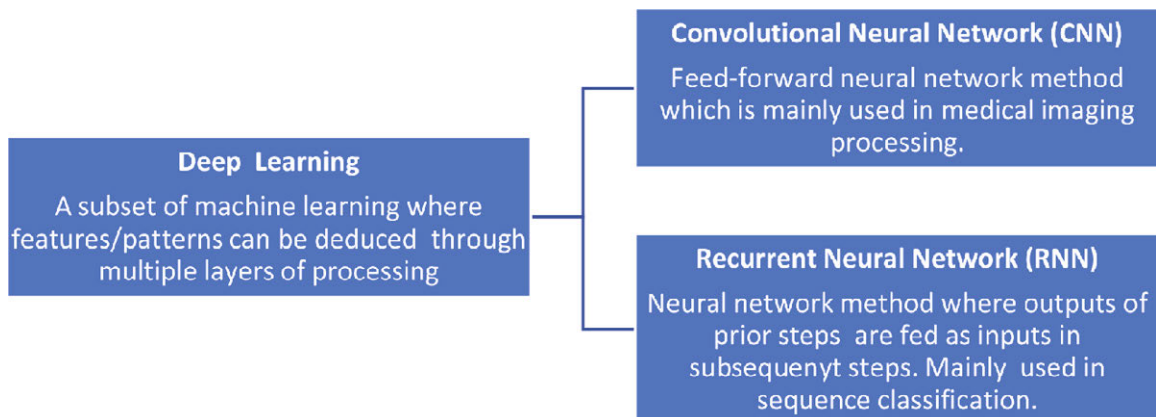**Figure 1** | Supervised machine learning and its different subtypes.

Figure 2 | Deep learning with description of two types of neural networks frequently used in medical literature.

2.2. Models' Evaluation

Applying the different ML and DL methods described above to different medical questions requires measures to ensure that the models are accurate and valid. It is important to note that every model is usually based on a set of data called 'training set.' After creating the model, it is generally validated with either another set of data, "validation set", or using the original data set itself. However, the usual approach is to start with the validation with the original data set, and then use an external "test set" sample.

The hold-out approach involves the process by which the data are usually split into training and validation sets. Typically, the validation set is a new set of data that is not a subset of the training set. This approach ensures that the data can be applied to another/new set of data. A discrepancy in accuracy between the training and validation sets indicates 'overfitting.' Overfitting is a phenomenon where the algorithm is modeled very well on the training set to a point where it is not generalizable to other data [8,11]. However, another approach to validation is using the training set itself. A commonly used method is 'cross-validation.' In cross-validation, the training set will usually be split into multiple subsets on which the model will be run to confirm accuracy. The number of subsets is generally dictated by how many k-folds of cross-validation, where k is the number of subsets [11].

Moreover, models are evaluated by different metrics. Similar to the evaluation of diagnostic tests, classification algorithms (the majority of diagnostic algorithms) can be assessed based on their ability to identify samples with a specific diagnosis (Figure 3). Thus, concepts such as 'accuracy' and 'precision' can still be used to evaluate algorithms [12,13]. On the other hand, "recall" can be calculated by the ratio of positive results identified by the algorithm to all the positive samples. Moreover, $F1$ (F-measure) is a harmonic mean

of precision and recall used to measure the not correctly classified results [11].

Using the algorithms' specificity and sensitivity, the receiver operator curve (ROC) can be drawn using a true-positive and false-positive rate. Additionally, the area under the curve (AUC) can be calculated, an operation which has been widely used for years [14] and is now used to evaluate classification algorithms with binary results. There are no exact cut-offs to interpret AUC values; however, AUC of 0.50 and below usually indicates a low classification ability. Table 1 demonstrates the different definitions.

3. ML AND DL IN THE DIAGNOSIS OF HEMATOLOGICAL DISEASES

The use of ML and DL tools has been investigated in different areas of hematologic diagnoses, including laboratory, histopathology, flow cytometry, and molecular data. The applications of ML and DL in hematology are many. In the discussion below, we highlight ML's and DL's multiple applications in diagnosing hematologic diseases. Table 2 lists selected examples of these.

3.1. Image-based Diagnosis and Recognition

Image recognition and diagnosis are essential in multiple diseases, including malignant and benign hematologic diseases. The increase in ML and DL utilization could be attributed to the increased awareness and the efforts to digitalize histopathology slides, including whole slide imaging (WSI) [15,16]. The primary ML function used in image-based diagnosis is the classification of histopathology slides.

The approach of using ML in image recognition and classification usually starts with pre-processing the images, which includes digitally labeling the slides. Slide labeling is particularly important with supervised machine learning, which generally requires the labeling of both the input and the output, a process that introduces a logistical limitations, given the required resources to label many slides [17]. Following that, images are usually segmented into different parts (i.e., cytoplasm and nucleus), followed by identifying features (feature extraction). The ML algorithm can then be applied to the sample (usually with known output), creating a model that

Prediction Model Results			
		Predicted yes	**Predicted no**	
Actual Results	**Yes**	**True positives (TP)**	**False negative (FN)**	Sensitivity= (TP)/(TP +FN)
	No	**False positive (FP)**	**True negative (TN)**	Specificity= (TN)/(TN +FP)

Figure 3 | Confusion matrix.

Table 1 | Definitions of frequently used evaluation matrices

Index	Description	Comments
Accuracy	Reflects the models' ability to achieve correct predictions	(TP + TN)/(TP + TN + FN + FP)
Precision	It is the ratio of the true positive results to the total positive results	TP/(TP + FP)
Recall	It is the ratio of the true positive results to the total results that are supposed to be positives	TP/(TP + FN)
$F1$ score	A measure that uses precision and recall to identify the not correctly classified results	$\dfrac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$
The area under the curve	It is the area under the receiver operating characteristic curve, which plots two parameters; true positive rate (sensitivity) over false-positive rate (1-specificity)—used to assess the performance of binary classification models	

TP, true positive; TN, true negative; FP, false positive; FN, false-negative.

Table 2 | Selected examples of ML and DL use in the diagnosis of hematologic disease

Hematologic disease	Application	Diagnostic modality	Methodology	References
Acute leukemia	Diagnosing AML using histopathology slides	Microscopic	ML-supervised (SVM)	Kazemi et al. [23]
	Diagnosing ALL using histopathology slides	Microscopy	ML-supervised (Multiple algorithms)	Rawat et al. [22]
	Using gene expression profiling (GEP) to diagnose AML	GEP	Multiple	Warnet-Herresthal et al. [39]
Chronic leukemia	Differentiating normal cytometry from flow cytometry indicating CLL	Flow cytometry	Multiple supervised and unsupervised algorithms	Lakoumenta et al. [37]
Myeloma	Detecting bone metastasis in myeloma patients	Radiology	ML-supervised (SVM)	Xu et al. [27]
	Diagnosis of MM	Mass spectrum	DL	Deulofeu et al. [38]
Lymphoma	Differentiating different types of lymphoma	Microscopy	DL (CNN)	Achi et al. [19]
	Grading follicular lymphoma into a high or low grade	Microscopy	ML-supervised (k-NN)	Fauzi et al. [20]
	Using GEP to classify large B-cell lymphoma (DLBCL)	GEP	ML-supervised (SVM)	Zhao et al. [40]
Myelodysplastic syndrome	Detecting MRD in AML and MDS	Flow cytometry	ML-supervised (SVM)	Ko et al. [36]
Anemia	Diagnosing anemia via fundoscopic images	Fundoscopic images	DL (CNN)	Mitani et al. [26]
	Using cell population data parameters to improve the detection of liver disease and anemia in samples with abnormal scattergrams	Laboratory data	ML-supervised (Multiple algorithms)	Bigorra et al. [34]

ALL, acute lymphocytic leukemia; AML, acute myeloid leukemia; CLL, chronic lymphocytic leukemia; MM, multiple myeloma; MDS, myelodysplastic syndrome; MRD, minimal residual disease.

can classify images based on their features. The majority of models in the slide-based diagnosis use a binary approach (diagnosis/no diagnosis), which significantly simplifies the real-life complexity of histopathology diagnosis [18]. Subsequently, all models have to be validated internally or externally to ensure applicability. The use of ML and DL is not a guarantee for developing an accurate model, as many factors play a role in achieving higher accuracy models.

Multiple studies have reported on the application of ML and DL in lymphoma. In one of these, Achi et al. [19] utilized the DL methodology, particularly CNN and WSI from two different databases. The established model was able to diagnose three different types of lymphomas (diffuse large B-cell lymphoma, Burkitt lymphoma, and small lymphocytic lymphoma) with an accuracy of 95%, illustrating the utility of ML in more complex medical questions. As a DL approach, CNN is used for image classification by analyzing each image's features in multiple deep layers. Usually, the initial layers aim at understanding the simple features of the image, subsequently moving up to more complex and sophisticated feature extraction. K-nearest neighbor (k-NN) is another type of algorithm for tackling both classification and regression problems. Fauzi et al. [20] used k-NN to create a model to grade follicular lymphoma into a high or low grade. The model achieved an acceptable AUC of 0.75.

In leukemia, several studies have reported on ML supervised approaches such as support vector machine (SVM), naïve Bayesian, and random forest. Bigorra et al. [21] and Rawat et al. [22] have used SVM for diagnosing acute lymphoid leukemia (ALL). The AUC achieved by both models was 0.74 and 0.84, respectively. SVM is one of the most commonly used ML supervised approaches. It requires

the labeled input and output data to analyze and process the input data, and to perform a classification function, as in the case of histopathology diagnosis. Kazemi et al. [23] used SVM in diagnosis acute myeloid leukemia (AML), achieving a model accuracy of 95%. CNN was primarily used in the histopathology diagnosis of ALL, with models achieving accuracy that is close to 95% [24,25].

The ML utility in image recognition goes beyond the histopathology diagnosis of hematologic diseases. With data from the UK Biobank, Mitani et al. [26] used retinal fundal imaging to diagnose anemia via DL, achieving an AUC of 0.88 when used along with other patients' data. ML was utilized for detecting bone lesions in multiple myeloma patients via analyzing PET/CT scans using various algorithms, including SVM, k-NN, RF, and CNN [27]. The latter outperformed the other ML algorithms used in that study.

Although the magnitude of practice change by such studies is questionable, they demonstrate the power of DL and ML in analyzing patterns, and the importance of having databases that can provide sufficient data to build algorithms. Furthermore, identifying patterns in imaging using AI could help develop prognostic and diagnostic models [28].

Studies that have been done so far in histopathology diagnosis and image recognition are limited in terms of the sample size, which can be attributed to the lack of multi-institutional database-based studies. However, digital histopathology is a growing field, and WSI's use may provide multiple opportunities, particularly in ML and DL utilization. The challenge in the AI literature remains to establish whether these models will achieve similar results in real-practice [21,23–25].

3.2. Laboratory-based Diagnosis and Flow Cytometry

Multiple studies have utilized ML and DL in laboratory diagnosis and flow cytometric diagnosis of benign and malignant hematological diseases. The integration of AI models in hematology, particularly laboratory hematology, is not new [29,30]. Several knowledge-based equipment types were developed to aid in simple laboratory diagnoses, for instance, iron deficiency anemia or peripheral smear analysis [31,32].

The current literature provides a good sample of the breadth of the possible applications. For instance, one study has used both regression and classification algorithms to predict ferritin levels and increase its accuracy, as it is usually affected by multiple other biologic processes [33]. The model presented achieved an AUC of 0.97.

Bigorra et al. [34] investigated the use of ML to build a model that can use cell population data parameters to improve the detection of liver disease and anemia in samples with abnormal scattergrams. Multiple algorithms have been used, including random forests, naive Bayes classifiers, *k*-NN, neural networks, and SVM. The later achieved the highest accuracy. Guncar et al. [35] have used the random forest method to create a model that is able to analyze patterns between the different blood indices (which are usually overlooked) to guide clinicians to the five most possible hematologic diagnoses (both benign and malignant) at an early stage of investigation. The model's accuracy outperformed the evaluation by internal medicine specialists, and was comparable to that of hematology specialists. The random forest algorithm is another supervised ML algorithm that combines multiple decision trees.

There are also data on the utility of flow cytometry-based ML algorithms yielding diagnostic and prognostic information [36,37]. In one study, the authors used SVM to develop multiple algorithms that can detect minimal residual disease in AML and myelodysplastic syndrome, using data from multiparameter flow cytometry [36]. The model utilized results from 5000 bone marrow samples from 1700 patients, and achieved an AUC of higher than 0.90. Other applications included using mass spectrum as in Deulofeu et al. [38] along with DL (using artificial neural networks) to achieve 100% sensitivity and 95% specificity in diagnosing multiple myeloma.

3.3. Applications of ML in Genomics

With the integration of genomics in cancer diagnosis, prognosis, and treatment, ML and DL have been utilized in improving the diagnosis of multiple cancers, including hematologic malignancies. For instance, Warnet-Herresthal et al. [39] used gene expression profiling to create a system that is able to diagnose AML with no expertise intervention. The authors have used multi-dimensional data with multiple algorithms achieving AUC and accuracy of more than 90%. However, as with many other ML articles, the approach was preliminary, and more prospective studies are needed to establish the utility of these models.

Machine learning and DL approaches have also been utilized in the classification of lymphoma using gene expression profile and DNA microarrays [40,41]. On one study, gene expression profiling was used to classify large B-cell lymphoma into its different molecular subtypes using SVM. The model was able to classify the data into germinal center B-cell like (GCB) and non-GCB with an accuracy of more than 90% in the validation cohort [40].

Given the advancements in genomics and the amount of data generated by the cutting-edge technologies, ML and DL offer the analytical tools to approach and use these data in clinical practice. However, high-quality studies and well-established databases will be needed to build models that will be clinically relevant.

3.4. Other Clinical Uses of ML and DL

The previous discussion has highlighted multiple uses of ML and DL in the diagnosis of hematologic diseases. However, ML and DL tools have been used in various other applications, including prognostic models. These applications are beyond the scope of this article; however, below is a discussion of some examples.

The utilization of AI in hematopoietic cell transplantation (HCT) is increasing [42]. The majority of ML and DL studies in HCT are for prognostication, but multiple studies have tried to tackle other problems, including pre-transplant patient's selection. Donor/recipient and transplantation characteristics were used to create algorithms to improve pre-transplant match results [43,44]. Nonetheless, AI/ML is complex and dependent on multiple factors; thus, some algorithms had suboptimal accuracy or failed validation [43,45]. Graft vs host disease (GvHD) is a significant cause of morbidity after allogeneic HCT. However, the use of ML and DL models in GvHD diagnosis/prediction is still minimal. Current studies demonstrate these approaches' potential in improving the accuracy of predicting GvHD and its associated prognosis [46,47]. In a study by Lee et al. [48], a super-learner (combining multiple algorithms) was used to create a model to predict acute GvHD. The algorithm achieved a modest AUC of around 0.60, indicating the importance of data selection in building models.

Several HCT-related databases have been established. They are being used in ML and DL studies, mostly by the European Society for Blood and Marrow Transplantation and the Center for International Blood and Marrow Transplant Research [45,49]. Although they are still in their initial phases, these databases' use and investment might improve ML and DL utilization [50].

Gal et al. [51] used ML to predict complete remission in AML patients utilizing around 75 genes. The *k*-NN algorithm was used, and the model was able to achieve an AUC of 0.81. Other examples of prognostic models were published in the literature, including the use of single-nucleotide polymorphism data for the prognosis of MM [52], gene expression to predict prognosis in Hodgkin lymphoma [53], and an ensemble algorithm to develop a model to predict the risk of infection complications in chronic lymphocytic leukemia patients [54].

4. FUTURE DIRECTIONS

The field of AI and ML is a promising field that would help to evolve our practice. This review highlights the different trends and applications in the hematology literature. AI uses are so far limited in terms of quality and quantity [4,18]. "Big data" is difficult

to obtain with no sufficient databases with internally homogenous data, particularly in histopathology diagnosis where digital histopathology is still in its childhood. New databases should be developed with governance and data collection processes that allow for optimal use of ML algorithms. In a promising step, the American Society of Hematology has announced the development of a data hub to harness big data [55].

The incorporation of ML tools in medical practice faces many challenges. Technical hurdles include data collection and the need for 'big data' to create representative algorithms. However, the challenges facing ML integration are multi-faceted. When data are available, creating a model might be technically feasible, but many other issues must be taken into consideration. For instance, the ethics of data sharing and the use of patient information is a significant issue to be addressed. A similar concern is the possibility of, models' bias, which may deepen health-care disparities by representing only certain groups and increasing racial biases, has aspects which have been previously reported [56]. These issues will require regulations and guidelines to avoid the misuse of data or an increase in bias. Lastly, to be able to meaningfully implement ML models, there is a growing need for them and their tools to be technically well-performing, transparent and explainable to both providers and patients [57].

The currently available data show a limited number of prospective studies in the field of AI, with no evidence to improve clinical outcomes. The ML models reported on different studies lack reproducibility, limiting their applicability. An accurate ML algorithm may not be reproducible on another population/dataset [58]. The costs of reproducing/replicating ML algorithms are huge, introducing a challenging aspect of implementing them in real-practice [59]. Moreover, current utilization of ML does not always yield significantly better clinical outcomes and results, as demonstrated by a recent review of algorithms in the field of heart failure, indicating that more work is needed to unlock the power of ML [60].

The field ML, particularly in hematology, holds the potential for future impact. This should be facilitated by developing databases/hubs, better governance of data collection, and improving current research practices.

CONFLICTS OF INTEREST

The authors declare they have no conflicts of interest.

AUTHORS' CONTRIBUTION

INM and SKH wrote the first draft of the manuscript. All authors vouched for the accuracy and contents of the manuscript.

DISCLOSURES

SKH has received honorarium from Mallinckrodt, Janssen, Novartis and Pfizer & travel grants from Merck, Takeda, Sanofi and Gilead. ADS has received research support from Merck & Co. and Novartis.

REFERENCES

[1] Patel VL, Shortliffe EH, Stefanelli M, Szolovits P, Berthold MR, Bellazzi R, et al. The coming of age of artificial intelligence in medicine. Artif Intell Med 2009;46;5–17.

[2] Kantarjian H, Yu PP. Artificial intelligence, big data, and cancer. JAMA Oncol 2015;1;573–4.

[3] Beam AL, Kohane IS. Big data and machine learning in health care. JAMA 2018;319;1317–18.

[4] Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019;25;44–56.

[5] Char D, Shah N, Magnus D. Implementing machine learning in health care — addressing ethical challenges. N Engl J Med 2018;378;981–3.

[6] Donald Harvey R. New drug updates in hematologic malignancies: CAR-T, targeted therapeutics, and other agents. J Adv Pract Oncol 2018;9;282–6.

[7] Zou J, Huss M, Abid A, Mohammadi P, Torkamani A, Telenti A. A primer on deep learning in genomics. Nat Genet 2018;51;12–18.

[8] Liu Y, Chen PHC, Krause J, Peng L. How to read articles that use machine learning: users' guides to the medical literature. JAMA 2019;322;1806–16.

[9] Kohli M, Prevedello L, Filice R, Geis J. Implementing machine learning in radiology practice and research. AJR Am J Roentgenol 2017;208;754–60.

[10] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015; 521;436–44.

[11] Handelman GS, Kok HK, Chandra RV, Razavi AH, Huang S, Brooks M, et al. Peering into the black box of artificial intelligence: evaluation metrics of machine learning methods. AJR Am J Roentgenol 2019;212;38–43.

[12] Jaeschke R, Guyatt G, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test, A. Are the results of the study valid? The Evidence-Based Medicine Working Group. JAMA 1994;271;389–91.

[13] Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test, B. What are the results, and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. JAMA 1994;271;703–7.

[14] Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143;29–36.

[15] Ghaznavi F, Evans A, Madabhushi A, Feldman M. Digital imaging in pathology: whole-slide imaging and beyond. Annu Rev Pathol 2013;24;331–59.

[16] Acs B, Rimm DL. Not just digital pathology, intelligent digital pathology. JAMA Oncol 2018;4;403–4.

[17] Tizhoosh HR, Pantanowitz L. Artificial intelligence and digital pathology: challenges and opportunities. J Pathol Inform 2018;9;38.

[18] Salah HT, Muhsen IN, Salama ME, Owaidah T, Hashmi SK. Machine learning applications in the diagnosis of leukemia: current trends and future directions. Int J Lab Hematol 2019;41; 717–25.

[19] Achi HE, Belousova T, Chen L, Wahed A, Wang I, Hu Z, et al. Automated diagnosis of lymphoma with digital pathology images using deep learning. Ann Clin Lab Sci 2019;49;153–60.

[20] Fauzi MFA, Pennell M, Sahiner B, Chen W, Shana'ah A, Hemminger J, et al. Classification of follicular lymphoma: the effect of computer aid on pathologists grading. BMC Med Inform Decis Mak 2015;15;115.

[21] Bigorra L, Merino A, Alférez S, Rodellar J. Feature analysis and automatic identification of leukemic lineage blast cells and reactive lymphoid cells from peripheral blood cell images. J Clin Lab Anal 2016;31;e22024.

[22] Rawat J, Singh A, Bhadauria HS, Virmani J, Devgun JS. Classification of acute lymphoblastic leukaemia using hybrid hierarchical classifiers. Multimedia Tools Appl 2017;76;19057–85.

[23] Kazemi F, Najafabadi TA, Araabi BN. Automatic recognition of acute myelogenous leukemia in blood microscopic images using K-means clustering and support vector machine. J Med Signals Sens 2016;6;183–93.

[24] Shafique S, Tehsin S. Acute lymphoblastic leukemia detection and classification of its subtypes using pretrained deep convolutional neural networks. Technol Cancer Res Treat 2018;17;1533033818802789.

[25] Rehman A, Abbas N, Saba T, Rahman S, Mehmood Z, Kolivand H. Classification of acute lymphoblastic leukemia using deep learning. Microsc Res Tech 2018;81;1310–17.

[26] Mitani A, Huang A, Venugopalan S, Corrado GS, Peng L, Webster DR, et al. Detection of anaemia from retinal fundus images via deep learning. Nat Biomed Eng 2019;4;18–27.

[27] Xu L, Tetteh G, Lipkova J, Zhao Y, Li H, Christ P, et al. Automated whole-body bone lesion detection for multiple myeloma on ^{68}Ga-Pentixafor PET/CT imaging using deep learning methods. Contrast Media Mol Imaging 2018;2018;2391925.

[28] Schenone D, Lai R, Cea M, Rossi F, Torri L, Bignotti B, et al. Radiomics and artificial intelligence analysis of CT data for the identification of prognostic features in multiple myeloma, vol. 11314. Houston, USA: Medical Imaging 2020: Computer-Aided Diagnosis; 2020, p. 113144A.

[29] Zini G. Artificial intelligence in Hematology. Hematology 2005;10;393–400.

[30] Diamond LW, Mishka VG, Seal AH, Nguyen DT. Multiparameter interpretive reporting in diagnostic laboratory hematology. Int J Biomed Comput 1994;37;211–24.

[31] Kratz A, Lee SH, Zini G, Riedl JA, Hur M, Machin S, et al. Digital morphology analyzers in hematology: ICSH review and recommendations. Int J Lab Hematol 2019;41;437–47.

[32] Dogan S, Turkoglu I. Iron-deficiency anemia detection from hematology parameters by using decision trees. Int J Sci Technol 2008;3;85–92.

[33] Luo Y, Szolovits P, Dighe AS, Baron JM. Using machine learning to predict laboratory test results. Am J Clin Pathol 2016;145;778–88.

[34] Bigorra L, Larriba I, Gutiérrez-Gallego R. Machine learning algorithms for the detection of spurious white blood cell differentials due to erythrocyte lysis resistance. J Clin Pathol 2019;72;431–7.

[35] Gunčar G, Kukar M, Notar M, Brvar M, Černelč P, Notar M, et al. An application of machine learning to haematological diagnosis. Sci Rep 2018;8;411.

[36] Ko BS, Wang YF, Li JL, Li CC, Weng PF, Hsu SC, et al. Clinically validated machine learning algorithm for detecting residual diseases with multicolor flow cytometry analysis in acute myeloid leukemia and myelodysplastic syndrome. EBioMedicine 2018;37;91–100.

[37] Lakoumentas J, Drakos J, Karakantza M, Nikiforidis GC, Sakellaropoulos GC. Bayesian clustering of flow cytometry data for the diagnosis of B-Chronic Lymphocytic Leukemia. J Biomed Inform 2009;42;251–61.

[38] Deulofeu M, Kolářová L, Salvadó V, Peña-Méndez EM, Almáši M, Štork M, et al. Rapid discrimination of multiple myeloma patients by artificial neural networks coupled with mass spectrometry of peripheral blood plasma. Sci Rep 2019;9;7975.

[39] Warnat-Herresthal S, Perrakis K, Taschler B, Becker M, Baßler K, Beyer M, et al. Scalable prediction of acute myeloid leukemia using high-dimensional machine learning and blood transcriptomics. iScience 2020;23;100780.

[40] Zhao S, Dong X, Shen W, Ye Z, Xiang R. Machine learning-based classification of diffuse large B-cell lymphoma patients by eight gene expression profiles. Cancer Med 2016;5;837–52.

[41] Hong JH, Cho SB. Lymphoma cancer classification using genetic programming with SNR features. EuroGP 2004;3003;78–88.

[42] Muhsen IN, Elhassan T, Hasmhi SK. Artificial intelligence approaches in hematopoietic cell transplantation: a review of the current status and future directions. Turk J Hematol 2018;35;152–7.

[43] Marino SR, Lee SM, Binkowski TA, Wang T, Haagenson M, Wang HL, et al. Identification of high-risk amino-acid substitutions in hematopoietic cell transplantation: a challenging task. Bone Marrow Transplant 2016;51;1342–9.

[44] Buturovic L, Shelton J, Spellman SR, Wang T, Friedman L, Loftus D, et al. Evaluation of a machine learning-based prognostic model for unrelated hematopoietic cell transplantation donor selection. Biol Blood Marrow Transplant 2018;24;1299–306.

[45] Shouval R, Labopin M, Unger R, Giebel S, Ciceri F, Schmid C, et al. Prediction of hematopoietic stem cell transplantation related mortality- lessons learned from the in-silico approach: a European Society for Blood and Marrow Transplantation Acute Leukemia Working Party data mining study. PLoS One 2016;11;e0150637.

[46] Caocci G, Baccoli R, Vacca A, Mastronuzzi A, Bertaina A, Piras E, 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;426–33.

[47] Elhassan T, Chaudhri N, Ahmed SO, Rasheed W, Alsharif F, Al Zahrani H. Predicting day-100 TRM in patients with acute leukemia who received myloablative (MA) conditioning regimen and transplanted using matched related donor Allo-HCT: a machine learning study [abstract B455]. In: 44th Annual Meeting of the European Society for Bone Marrow Transplant. Lisbon: EBMT; 2018.

[48] Lee C, Haneuse S, Wang HL, Rose S, Spellman SR, Verneris M, et al. Prediction of absolute risk of acute graft-versus-host disease following hematopoietic cell transplantation. PLoS One 2018;13;e0190610.

[49] Shouval R, Labopin M, Bondi O, Mishan-Shamay H, Shimoni A, Ciceri Jordi Esteve F, 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;3144–51.

[50] Muhsen IN, Jagasia M, Toor AA, Hashmi SK. Registries and artificial intelligence: investing in the future of hematopoietic cell transplantation. Bone Marrow Transplant 2018;54;477–80.

[51] Gal O, Auslander N, Fan Y, Meerzaman D. Predicting complete remission of acute myeloid leukemia: machine learning applied to gene expression. Cancer Inform 2019;18;1176935119835544.

[52] Waddell M, Page D, Shaughnessy J. Predicting cancer susceptibility from single-nucleotide polymorphism data: a case study in multiple myeloma. In: BIOKDD '05: Proceedings of the 5th International Workshop on Bioinformatics. Chicago, Illinois, USA: ACM; 2005, pp. 21–8.

[53] Parodi S, Manneschi C, Verda D, Ferrari E, Muselli M. Logic Learning Machine and standard supervised methods for Hodgkin's lymphoma prognosis using gene expression data and clinical variables. Health Inform J 2016;24;54–65.

[54] Agius R, Brieghel C, Andersen MA, Pearson AT, Ledergerber B, Cozzi-Lepri A, et al. Machine learning can identify newly diagnosed patients with CLL at high risk of infection. Nat Commun 2020;11;363.

[55] Hematology.org. ASH announces its vision for hematologic big data [online]. 2017. Available from: https://ashpublications.org/thehematologist/article/doi/10.1182/hem.V14.5.7608/462916/ASH-Announces-Its-Vision-for-Hematologic-Big-Data.

[56] Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science 2019;366;447–53.

[57] Lamy JB, Sekar B, Guezennec G, Bouaud J, Séroussi B. Explainable artificial intelligence for breast cancer: a visual case-based reasoning approach. Artif Intell Med 2019;94;42–53.

[58] Li J, Liu L, Le TD, Liu J. Accurate data-driven prediction does not mean high reproducibility. Nat Mach Intell 2020;2;13–15.

[59] Beam AL, Manrai AK, Ghassemi M. Challenges to the reproducibility of machine learning models in health care. JAMA 2020;323;305–6.

[60] Desai RJ, Wang SV, Vaduganathan M, Evers T, Schneeweiss S. Comparison of machine learning methods with traditional models for use of administrative claims with electronic medical records to predict heart failure outcomes. JAMA Netw Open 2020;3;e1918962.