Current Hematologic Malignancy Reports (2024) 19:9–17
https://doi.org/10.1007/s11899-023-00716-5

# Unlocking the Potential of Artificial Intelligence in Acute Myeloid Leukemia and Myelodysplastic Syndromes

**Abdulrahman Alhajahjeh**[1,2] · **Aziz Nazha**[3]

Accepted: 25 October 2023 / Published online: 24 November 2023


## Abstract

**Purpose of the Review** This review aims to elucidate the transformative impact and potential of machine learning (ML) in the diagnosis, prognosis, and clinical management of myelodysplastic syndromes (MDS) and acute myeloid leukemia (AML). It further aims to bridge the gap between current advances of ML and their practical application in these diseases.

**Recent Findings** Recent advances in ML have revolutionized prognostication, diagnosis, and treatment of MDS and AML. ML algorithms have proven effective in predicting disease progression, optimizing treatment responses, and in the stratification of patient groups. Particularly, the use of ML in genomic and epigenomic data analysis has unveiled novel insights into the molecular heterogeneity of MDS and AML, leading to better-informed therapeutic strategies. Furthermore, deep learning techniques have shown promise in analyzing complex patterns in bone marrow biopsy images, providing a potential pathway towards early and accurate diagnosis.

**Summary** While still in the nascent stages, ML applications in MDS and AML signify a paradigm shift towards precision medicine. The integration of ML with traditional clinical practices could potentially enhance diagnostic accuracy, refine risk stratification, and improve therapeutic approaches. However, challenges related to data privacy, standardization, and algorithm interpretability must be addressed to realize the full potential of ML in this field. Future research should focus on the development of robust, transparent ML models and their ethical implementation in clinical settings.

**Keywords** Artificial intelligence · Machine learning · Acute myeloid leukemia · Myelodysplastic syndromes

## Introduction

Hematologic malignancies such as acute myeloid leukemia (AML) and myelocytic dysplastic syndrome (MDS) are complex and heterogeneous diseases that present significant challenges to oncologists and researchers [1–3]. These diseases involve various clinical and molecular alterations that contribute to treatment resistance and relapse, making it difficult to understand the disease and improve patient outcomes [4, 5].

Fortunately, recent advancements in artificial intelligence (AI) and machine learning (ML) offer hope in overcoming these challenges. AI and ML are technologies that use computer algorithms to mimic human thinking and learning processes, and they have shown tremendous potential in healthcare [6, 7]. AI and ML are often confused, but ML is a branch of AI that involves models or algorithms that can learn from data and perform tasks more flexibly than being directly programmed [8]. As the volume and complexity of medical data increase, AI and ML can extract useful results from vast amounts of data, accelerate discovery, optimize patient care, and reduce human labor in the medical field [9].

In the case of hematologic malignancies, AI and ML have shown promise in diagnosis, risk stratification, predicting prognosis, and treatment and drug discovery [9]. For example, AI can analyze patient data and predict the likelihood of relapse or response to therapy, helping oncologists to make informed decisions about treatment [10]. AI can also identify genetic mutations that contribute to drug resistance, leading to the development of more effective treatments [11].

✉ Aziz Nazha
azizn38@yahoo.com

[1] Medical School, University of Jordan, Amman, Jordan

[2] Department of Internal Medicine, King Hussein Cancer Center, Amman, Jordan

[3] Department of Medical Oncology, Sidney Kimmel Cancer Center, Thomas Jefferson University, Philadelphia, PA, USA

This review article aims to explore the potential of AI in AML/MDS and how it can revolutionize the management of these complex diseases. With the growing availability of electronic medical records and genomic data, AI and ML offer exciting opportunities to transform healthcare and improve patient outcomes in hematologic malignancies.

## A Brief Introduction to AI Terminologies

### Machine Learning

The field of AI has made significant strides in recent years, with ML being a prominent area of application in healthcare. ML algorithms can process a wide range of data types, either individually or in combination, to produce outcomes that are not easily achievable through traditional methods.

Supervised and unsupervised algorithms are two types of ML algorithms that are widely used in healthcare applications [12, 13]. In supervised algorithms, the desired outcome is known, and the algorithm is trained to achieve the best results possible. This is typically accomplished through the use of regression or classification techniques. In contrast, unsupervised algorithms are used when the desired outcome is unknown, and the algorithm is trained to explore and identify new patterns in the data. Unsupervised algorithms can be used to identify novel features within histological sections to diagnose specific diseases that have not been identified previously. Despite the potential benefits, it is crucial to note that the use of unsupervised algorithms requires careful consideration by domain experts to determine if the results are meaningful or not. Nevertheless, the use of ML in healthcare continues to be an exciting and growing area of research that holds promise for improving patient outcomes in hematologic malignancies and other diseases [14].

### Deep Learning

Deep learning (DL) has been widely adopted in healthcare due to its ability to analyze complex and heterogeneous data sets, including text, images, and numerical data [15]. DL is based on neural network algorithms, inspired by the neuronal system in the human body. These algorithms consist of an input layer that receives various types of data, a hidden layer that processes the input, and an output layer that produces the desired results.

One type of neural network algorithm that has been widely used in healthcare is the convolutional neural network (CNN), which is commonly used for image analysis. CNNs are designed with convolutional layers that extract features from images, similar to the way the human visual system works. Radiologists have benefited the most from AI, with CNNs used for X-ray interpretation and the diagnosis of various radiological images [16••]. Moreover, CNNs have been used in histopathology to classify and predict the outcome of different pathologies and tumors, including evaluation of normal and abnormal cells in the bone marrow [17].

Another type of neural network algorithm is the recurrent neural network (RNN), which is typically used in natural language processing. The RNN algorithm is suitable for sequential tasks as it can remember previous inputs and use them to guide the processing of the following task. For instance, RNNs have been used to predict the development of complications or mortality in patients by analyzing electronic medical records. Moreover, RNNs have been used to predict the response to hypomethylating agents in MDS patients using 90-day complete blood count (CBC) data [18••].

### Transformer Models

Transformers are a type of neural network architecture that has become increasingly popular in natural language processing tasks such as language translation, text summarization, and question-answering. However, their application has now extended to other fields, including healthcare [19, 20]. Transformers have shown promising results in tasks such as medical image analysis, clinical diagnosis, and electronic health record analysis. These models have the ability to learn complex relationships in large datasets, making them a valuable tool for data-driven healthcare. These algorithms are the backbone of recent advances in AI that include large language models such as ChatGPT [19, 21, 22•, 23].

## Machine Learning in AML and MDS

Artificial intelligence (AI) has shown promise in this field, with potential applications for diagnosis, risk stratification, predicting prognosis, and treatment and drug discovery. With the increasing volume and complexity of medical data, AI could help extract useful results from this vast amount of data, accelerating discovery, optimizing patient care, and reducing human labor in the medical field, specifically for hematological disorders. The application of AI in AML and MDS can be summarized in several aspects as shown below.

### ML in Diagnosis

Computer vision has the potential to provide a more objective and standardized analysis of images and other types of data than traditional methods. Several studies have used computer vision to analyze bone marrow aspirate and biopsy images, as well as peripheral blood smears and flow cytometry data to improve the diagnostic accuracy of MDS and AML (Table 1).

Kimura et al. used a DL algorithm to analyze peripheral blood smear images from 3261 patients with various

**Table 1** Selected publications for the application of machine learning in acute myeloid leukemia and MDS diagnosis

| Author | Sample size | Type of data | Mulicenter vs uni-center | Method | Result |
|---|---|---|---|---|---|
| Kimura K et al. (2019) | 3261 peripheral blood smears to produce 703,970 cell images | Images | Single | CNN and XGB | Global AUC of 0.99 |
| Furong H, et al. (2020) | 104 bone marrow smears with 1322 images | Images | Multi | CNN with transfer learning Inception-V3, ResNet50, and DenseNet121 | The prediction accuracies of the normal groups, AML, ALL, and CML were 90%, 99%, 97%, and 95%, respectively |
| Dasariraju, S et al. (2020) | 18,365 images of leukocytes from the peripheral blood of 100 AML patients and 100 healthy controls | Images | Multi | Image segmentation and RF | Overall accuracy for the optimized random forest model was 93.45% |
| Mori J, et al. (2020) | 1797 images from 35 smears | Images | Multi | CNN(ResNet-152) followed by regression model | Final model has: Sensitivity: 91.0%, Specificity: 97.7%, PPV: 76.3%, NPV: 99.3%, Accuracy: 97.2% |
| Warnat-Herresthal S, et al. (2020) | 12,029 samples (sequenced for transcriptomic data) from 105 different studies | RNA seq | Multi | SVM with different kernels (linear, sigmoid, and polynomial), random forest, K-NN, linear discrimination, neural network, logistic regression, and LASSO | The mean accuracy to differentiate AML from CLL, ALL, CML and MDS was 0.99. The mean accuracy to differentiate AML subtypes ranged between 0.92–0.97 |
| Acevedo A, et al. (2021) | 20,670 images | Images | Single | CNN ("DysplasiaNet ") | 95.5% sensitivity, 94.3% specificity, 94% precision, and a global accuracy of 94.85% for diagnosing MDS |
| Radakovich N, et al. (2021) | 2697 patients | Clinical and genomic | Multi | GBM, RF | AUROC of 0.951 (0.934 to 0.966) for test cohorts |
| Eckardt N et al. (2022) | 1335 bone marrow images from patients and healthy donors | Images | Multi | CNN | Distinguishing between APL and non-APL AML as well as APL and healthy donors with an AUROC characteristic of 0.8575 and 0.9585, respectively |

*CNN*, convolutional neural network; *RF*, random forest; *GBM*, gradiant boosting model; *AUC*, area under the curve; *AUROC*, area under the receiver operating characteristic curve; *SVM*, supported vector machine; *AML*, acute myeloid leukemia; *ALL*, acute lymphoid leukemia; *CML*, chronic myeloid leukemia; *CLL*, chronic lymphoid leukemia

hematological disorders, including MDS and AML. The algorithm achieved an overall area under the curve of 0.99 [24]. Acevedo A et al. used convolutional neural network (dsyplasiaNet) to analyze 20,670 images to differentiate MDS with a sensitivity of 95.5%, specificity of 94.3%, and a global accuracy of 95% (Table 1) [25]. Another study by Eckardt et al. used a machine learning algorithm to analyze bone marrow aspirate images from AML, and APL patients as well as healthy donors. The algorithm was able to distinguish APL and AML from healthy donors with AUC of 0.86 and 0.96, respectively (Table 1) [26].

Other researchers tried to use different types of data to improve the diagnostic accuracy of MDS and AML. Radakovich N et al. used clinical data from CBC and genomic data to build a ML to distinguish MDS from other myeloid malignancies. The authors used an explainable ML approach to identify 15 clinical and genomic data that were used to build the final model. When applied to the test and validation cohorts, the model achieved AUROC of 0.951 (0.934 to 0.966) for test cohorts, and AUROC of 0.926 (0.916 to 0.937) for the training cohorts without the need to have a bone marrow biopsy data [27]. Using explainable ML, the authors also showed that when NGS data and patient sex were used as inputs, the model was able to predict whether the patient has a complex karyotype with AUROC of 0.821, normal karyotype with an AUROC of 0.790, and abnormal karyotype with an AUROC of 0.761 [27]. In another study, Warnat-Herrsethal S et al. used RNA-seq data from 105 different studies to differentiate AML and MDS from other myeloid malignancies. The final model mean accuracy was 0.99. Further, the model was able to distinguish AML subtypes with mean accuracy of 0.92–0.97 across three different datasets [28•].

## ML in Prognosis

Machine learning models that predict survival for patients with myelodysplastic syndromes and acute myeloid leukemia are becoming increasingly important in the field of hematology. These models can analyze vast amounts of data and identify prognostic factors that are often difficult to detect using traditional methods. By providing more accurate and personalized prognostic information, these models can aid in treatment decision-making and ultimately improve patient outcomes (Table 2).

Nazha A et al. developed a personalized prediction model to risk stratify patients with MDS based on their unique clinical and molecular characteristics. The researchers used a cohort of 1471 MDS patients to develop the model and validate it in multiple patient cohorts from different academic centers in the USA, which incorporated various factors such as age, cytogenetics, and gene mutations. The personalized model demonstrated a higher C-index of 0.74 in predicting overall survival compared to the commonly used Revised International Prognostic Scoring System (IPSS-R) with a C-index of 0.66. Moreover, the personalized model improved risk stratification and patient outcome prediction, especially for those with low-risk IPSS-R scores [29••].

Eckardt J et al. developed machine learning models to predict complete remission and 2-year overall survival in a large cohort of 1383 AML patients who received intensive induction therapy [30]. Nine machine learning models were used to predict the outcomes, incorporating clinical, laboratory, cytogenetic, and molecular genetic data [30]. The models identified significant predictive markers for complete remission and 2-year overall survival, including established markers of favorable or adverse risk and markers of controversial relevance. The models showed feasibility for risk stratification in AML, demonstrating the clinical applicability of machine learning as a decision support system in hematology [30]. The areas under the receiver operating characteristic curves ranged between 0.77 and 0.86 for complete remission and between 0.63 and 0.74 for 2-year overall survival in the test set, and between 0.71–0.80 and 0.65–0.75 in the external validation cohort [30].

In another study, Tazi Y et al. aimed to integrate AML molecular classes into prognostic models for clinical management [31]. The researchers compared prognostic models based on genetic features to class-based models and found that a simple model based on class membership and FLT3ITD status captures the same prognostic information as more complex genetic models [31]. They also included clinical features such as age, gender, blast, antecedent hematologic disorder, performance status, white blood cells, hemoglobin, and platelet, which achieved the highest improvement in model discrimination [31]. The study also presented a multi-state model for disease progression that provides a detailed resolution of anticipated transitions across molecular subgroups and endpoint-specific outcomes for different AML classes [31].

## ML in Treatment Selection and Drug Discovery

Machine learning plays a crucial role in predicting response to cancer treatment and drug discovery by analyzing large datasets and identifying patterns that can inform treatment decisions. Its ability to rapidly process and integrate diverse data sources, such as genomics, proteomics, and clinical records, can accelerate the development of precision medicine approaches and ultimately improve patient outcomes (Table 3).

In a recent study by Radakovich et al. (2022), machine learning approaches were used to predict the response of myelodysplastic patients to hypomethylating agents [18••]. The study analyzed serial complete blood count data over a 90-day period from 514 patients, using 5-cross-folds and multiple models including RF, GBDT, XGBoost, lightGBM,

**Table 2** Selected publications that used machine learning to predict prognosis in MDS/AML

| Author | Sample size | Type of data | Multicenter vs uni-center | Method | Result |
|---|---|---|---|---|---|
| Radhachandran A et al. (2021) | 790,470 patients | Clinical data | Multi | XGB, LR, and NN | Out of all the models tested, the XGB model proved to be the most effective at detecting MDS one year before diagnosis, with an AUROC of 0.87, sensitivity of 0.79, and specificity of 0.80 |
| Nazha A et al. (2021) | 1936 patients | Clinical, cytogenetic, and molecular data | Multi | Random survival algorithm | The final model internal cross-validation showed the c-index score of 0.74 (95% CI, 0.73 to 0.75) for OS and 0.81 (95% CI, 0.80 to 0.82) for leukemia transformation in the training cohort and 0.71 (95% CI, 0.73 to 0.75) and 0.84 (95% CI, 0.73 to 0.75) compared with IPSS 0.66 (95% CI, 0.62 to 0.67) and IPSS-R 0.67 (95% CI, 0.62 to 0.68) in the validation cohort |
| Eckardt J et al. (2021) | 1383 patients | Clinical, laboratory, molecular, and cytogenetic data | Multi | RF, Linear SVM, LR, Adaptive Boosting, Gradient Boosting, Polynomial SVM, NN, RBF-SVM, and K-NN | The AUROC for predicting complete remission ranged between 0.71 and 0.80, and the AUROC for predicting 2-year overall survival ranged between 0.65 and 0.75 on external validation cohort |
| Orgueira AM et al. (2022) | 672 patients | Genomic data | Multi | Mclust algorithm and RFS | Their model which contains 123 variables achieved C-index of 0.70 in predicting the survival in the validation cohort |
| Tazi Y et al. (2022) | 3653 patients | Clinical and molecular data | Multi | six-state Markov Model (MM), SVM, RFS, Cox, Cox boost, Cox random effect, Linear and logistic Lasso | The best model concordance index was higher than 0.7 showing that a simple model based on $FLT3^{ITD}$ status (17 features), have the same prognostic property as complex one with 154 features |

*RFS*, random forest survival; *LASSO*, least absolute shrinkage and selection operator; *GBM*, gradiant boosting model; *LR*, logistic regression model; *NN*, neural network; *AUC*, area under the curve; *AUROC*, area under the receiver operating characteristic curve; *XGB*, eXtreme Gradient Boostinglight; *SVM*, supported vector machine

**Table 3** Selected publications that used machine learning to predict response to therapies in myelodysplastic syndromes and acute myeloid leukemia

| Author | Sample size | Type of data | Multicenter vs uni-center | Method | Result |
|---|---|---|---|---|---|
| Shouval R et al. (2015) | 28,236 patients | Clinical data | Multi | ADT | The model achieved an AUROC of 0.701 in predicting 100-day mortality after allo-HSCT, outperforming the European Group for Blood and Marrow Transplantation score, which had an AUROC of 0.646 |
| Herold T et al. (2018) | 1106 patients | Genomic data | Multi | LASSO | Using 29 gene markers and cytogenetic risk, a scoring system was developed to predict resistant AML. The system demonstrated an odds ratio of 2.39 ($P$-value < 0.01) as a continuous variable, but as a dichotomous classifier, it achieved an odds ratio of 8.03 (P-value < 0.001) with an accuracy of 77% and an AUC of 0.76 on the validation Cohort |
| Xie L et al. (2018) | 978 gene expression | Transcriptomic data | Single | LR, RF, VC, GBDT, and DNN | The DNN model demonstrated excellent performance in predicting drug-protein interactions, with a mean accuracy of 90.53% and a mean $F$-score of 86.38% |
| Fuse K et al. (2019) | 217 patients | Clinical data | Multi | ADT | the model achieved a prediction accuracy of 71%, an AUC of 0.68, and a false positive rate of 0.216 in identifying relapse after allo-HSCT |
| Radakovich N et al. (2022) | 514 patients | Clinical and laboratory data | Multi | RF, XGboost, and lightGBM, | In predicting the response to hypomethylating agent based on the CBC results from the first 90 days of treatment, the RF model outperformed the other models, achieving AUROC and PR-AUC scores of 0.84 in the validation cohort |
| Arabyarmohammadi. S et al. (2022) | 874 images | Image | Single | u-net, LASSO, linear, and discriminant analysis creating risk score | A risk score was developed based on texture features extracted from the images, showing a hazard ratio of 1.57 with a 95% CI of 1.01–2.45 and a $P$-value of .044 for predicting AML relapse. The resulting signature achieved an AUC of 0.71 in the validation set to predicate the relapse |
| Wadood. A et al. (2022) | 773 ligands | Molecular data | Multi | KNN cluster, GNB, SVM, and RF | The RF model showed superior performance in predicting the binding between the legend and STAT3, with an accuracy of 97%, precision of 0.98, recall of 0.77, F1-score of 0.84, and MCC of 0.88 |

*RF*, random forest; *GBDT*, gradient boosting decision tree; *XGboost*, eXtreme Gradient Boostinglight; *GBM*, gradient boosting model; *RNN*, recurrent neural network; *SVM*, supported vector machine; *CNN*, convolutional neural network;, *ADT*, alternating decision tree; *LASSO*, least absolute shrinkage and selection operator; *LR*, logistic regression model; *VC*, voting classifier; *DNN*, deep neural network; *AUC*, area under the curve; *AUROC*, area under the receiver operating characteristic curve; *CBC*, complete blood count; *allo-HSCT*, allogeneic hematopoietic stem cell transplantation; *OR*, odds ratio

RNN, and CNN. The results showed that RF, XGBoost, and lightGBM models had higher AUROC and precision-recall AUROC values in the training/test set, with the random forest model showing the highest values [18••].The independent validation set also confirmed the robustness of these models, with improved AUROC and precision-recall AUROC values. However, due to poor performance, the RNN and CNN models were excluded from the analysis [18••]. These findings suggest that machine learning approaches can be valuable tools for predicting patient response to cancer treatment, potentially leading to improved clinical outcomes.

Fuse et al. (2019) aimed to develop a machine learning algorithm to predict relapse in acute leukemia patients who had undergone allogeneic hematopoietic stem cell transplantation, while accounting for various prognostic factors [32]. The researchers used an alternating decision tree model and found that the algorithm achieved an accuracy of 78.4%, and AUC of 0.746, in the training set [32]. However, the performance of the model decreased in the validation set, with an accuracy of 71.0%, and AUC of 0.667. The model also identified the branching point of patients, indicating the optimal time to adjust treatment plans and improve patient management [32].

The study conducted by Shouval et al. (2015) aimed to predict the 100-day post-HSCT mortality using machine learning techniques in a large cohort of 28,236 patients, with a validation cohort of 19,765 patients and a test cohort of 8471 patients [33].The study employed an alternating decision tree model, which achieved an AUC of 0.697 for predicting the 100-day mortality, comparable to the Cox regression model [33]. Moreover, the machine learning model achieved an AUC of 0.648 for predicting the 2-year overall survival, close to the AUC of 0.653 obtained by the Cox regression model [33]. Herold et al. (2018) developed a machine learning classifier to predict resistance to AML treatment using a combination of clinical and laboratory variables [34]. The LASSO model identified several significant predictors, including PS29MRCdic, age, *NPM1*, *RUNX1*, and *TP53* mutations, with PS29MRCdic having the highest predictive power [34]. The classifier achieved an accuracy of 77% in categorizing AML patients as high or low risk for treatment resistance, which could improve risk stratification and ultimately lead to better treatment outcomes [34].

In another study, Nazha et al. developed a novel framework to explore the association of multiple mutations with resistance to hypomethylating agents (HMAs) in patients with MDS [35]. The approach is analogous to recommender systems used in commerce, in which customers who buy products A and B are likely to buy C [35]. The authors screened a cohort of 433 patients with MDS who received HMAs for the presence of common myeloid mutations in 29 genes obtained before therapy. The Apriori market basket analysis algorithm was used to assess the association between mutations and response. The authors identified several genomic combinations that were highly associated with no response [35]. These molecular signatures were present in 30% of patients with three or more mutations per sample and had an accuracy rate of 87% in the training cohort and 93% in the validation cohort [35].

## Challenges and Limitations of AI in Healthcare

While the application of ML in healthcare holds immense potential for improving diagnostics, treatment planning, and patient outcomes, several significant challenges and limitations persist. A primary constraint is the quality of data utilized in the predictive models. Inaccurate, incomplete, or biased data can lead to flawed predictions, potentially jeopardizing patient's outcomes. Additionally, the lack of inclusion of socioeconomic factors in these models often results in solutions that are not universally applicable, potentially reinforcing health inequities. This is because these models typically fail to consider how variables such as income, education, and geography might influence health outcomes. On the ethical and legal front, using information derived from ML models presents another challenge. The use of patient data raises concerns about privacy and consent, and the opacity of some machine learning processes (often referred to as the "black box" problem) may lead to decision-making processes that are not transparent or explainable. Furthermore, the legal responsibility when AI-driven decisions lead to incorrect diagnosis or treatment remains a largely unexplored and contentious issue. Balancing these challenges with the potential benefits of ML is a crucial task for healthcare professionals, data scientists, ethicists, and policymakers alike.

### Specific Challenges for the Application of AI in AML/MDS

The application of AI in the research and clinical realms of AML and MDS presents a multifaceted array of challenges. Notably, the limited datasets available for these conditions can hinder the development and refinement of AI models. The scarcity of data becomes especially pronounced when considering the intricate nuances and subtypes of these malignancies. Furthermore, the diagnosis of MDS based on histological slides is inherently challenging due to the subtle morphological changes that characterize the condition. Employing computer vision algorithms to identify blasts or dysplastic cells can lead to misleading results given the nuanced variations that even experienced hematopathologists sometimes grapple with. Additionally, there is a pertinent risk associated with biases in the available data. If datasets used to train AI models predominantly represent

certain patient demographics or disease subtypes, the resultant models can produce skewed or non-generalizable outcomes. As such, while AI offers promise in revolutionizing AML and MDS research, it is imperative to approach its integration with a discerning and critical lens.

### Future Directions for AI and Machine Learning in Healthcare and Oncology

The future of AI and ML in healthcare appears promising, with the potential to reshape various aspects of care delivery, disease prevention, and health promotion. The integration of large language models (LLMS) can significantly contribute to this transformation. LLMS, with their capacity to learn and adapt over time, can enhance AI's potential in healthcare, allowing it to provide dynamic solutions that evolve with new data and changing contexts. This could lead to improved prediction and diagnosis of diseases, personalized treatment plans, and the optimization of healthcare operations.

The continuous learning feature of LLMS could help address one of the key challenges in healthcare: data heterogeneity and temporality. These algorithms could accommodate and learn from the constantly evolving nature of patient data, therefore refining their predictive models over time. This evolution could lead to more precise, personalized care that adjusts to patients' changing health conditions.

To optimize the outcome of using AI in healthcare, several next steps should be considered. Firstly, ensuring the quality of data inputted into the models should be prioritized, as the performance of AI and ML models heavily relies on the accuracy and completeness of the data they are trained on. Moreover, to address the problem of model interpretability or the "black box" issue, efforts should be directed towards developing explainable AI models. This would allow healthcare professionals to understand and validate the predictions made by these models, thereby building trust and promoting their wider adoption. Lastly, it's crucial to establish legal and ethical guidelines for the use of AI and ML in healthcare. These should include procedures for obtaining informed consent from patients, safeguards to protect patient privacy, and regulations defining the responsibilities of different stakeholders when AI-driven decisions lead to medical errors.

The future of AI in the realms of MDS and AML is undeniably promising. Envisioning a new era of precision medicine and large language models that are poised to enhance diagnostic accuracy by processing vast amounts of medical literature, patient data, and clinical insights. More revolutionary, however, is the emergence of multimodal AI approaches, which can use image-based, clinical, genomic, and other types of data. By synthesizing information from histopathological slides, patient clinical histories, and genomic markers, these models offer unparalleled granularity in diagnosis and prognosis. As the fields of hematology and AI converge, a new paradigm of patient-centric, data-driven care emerges, holding the potential to radically transform the management of MDS and AML.

## Conclusion

In summary, the article discusses the application of AI in AML and MDS and the potential benefits it can offer not only in these diseases but across many other specialties in healthcare. These benefits include improved accuracy and efficiency in diagnoses, personalized treatment plans, and enhanced patient outcomes.

Further, the article emphasizes that AI has enormous potential to revolutionize healthcare by improving the quality and efficiency of care. However, careful consideration and planning are necessary to ensure that AI is integrated responsibly and effectively into healthcare systems. This requires collaboration between healthcare providers, data scientists, policymakers, and patients to address the challenges and limitations of AI and leverage its potential to improve healthcare outcomes for all.

### Compliance with Ethical Standards

**Conflict of Interest** AN is an employee at Incyte Pharma and owns stock at Incyte.

**Human and Animal Rights and Informed Consent** This article does not contain any studies with human or animal subjects performed by any of the authors.

## References

Papers of particular interest, published recently, have been highlighted as:
- • Of importance
- •• Of major importance

1. Cazzola M. Myelodysplastic syndromes. N Engl J Med. 2020;383:1358–74.
2. Koenig KL, Sahasrabudhe KD, Sigmund AM, Bhatnagar B. AML with Myelodysplasia-related changes: development, challenges, and treatment advances. Genes (Basel). 2020;11:845.
3. Vardiman JW, Thiele J, Arber DA, Brunning RD, Borowitz MJ, Porwit A, et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. Blood. 2009;114:937–51.
4. Newell LF, Cook RJ. Advances in acute myeloid leukemia. BMJ. 2021;375:n2026.
5. Lachowiez CA, Long N, Saultz J, Gandhi A, Newell LF, Hayes-Lattin B, et al. Comparison and validation of the 2022 European LeukemiaNet guidelines in acute myeloid leukemia. Blood Adv. 2023;7:1899–909.

6. Currie G, Hawk KE, Rohren E, Vial A, Klein R. Machine learning and deep learning in medical imaging: intelligent imaging. J Med Imaging Radiat Sci. 2019;50:477–87.
7. Koski E, Murphy J. AI in Healthcare. Stud Health Technol Inform. 2021;284:295–9.
8. Janiesch C, Zschech P, Heinrich K. Machine learning and deep learning. Electron Mark. 2021;31:685–95.
9. Bhalla S, Laganà A. Artificial intelligence for precision oncology. Adv Exp Med Biol. 2022;1361:249–68.
10. Eckardt J-N, Röllig C, Metzeler K, Kramer M, Stasik S, Georgi J-A, et al. Prediction of complete remission and survival in acute myeloid leukemia using supervised machine learning. Haematologica. 2023;108:690–704.
11. Duchmann M, Wagner-Ballon O, Boyer T, Cheok M, Fournier E, Guerin E, et al. Machine learning identifies the independent role of dysplasia in the prediction of response to chemotherapy in AML. Leukemia. 2022;36:656–63.
12. Roohi A, Faust K, Djuric U, Diamandis P. Unsupervised machine learning in pathology. Surg Pathol Clin. 2020;13:349–58.
13. Ghazal TM, Al Hamadi H, Umar Nasir M, Atta-Ur-Rahman, Gollapalli M, Zubair M, Adnan Khan M, Yeob Yeun C. Supervised machine learning empowered multifactorial genetic inheritance disorder prediction. Comput Intell Neurosci. 2022;2022:1051388.
14. Habehh H, Gohel S. Machine learning in healthcare. Curr Genomics. 2021;22:291–300.
15. Miotto R, Wang F, Wang S, Jiang X, Dudley JT. Deep learning for healthcare: review, opportunities and challenges. Brief Bioinform. 2018;19:1236–46.
16.•• Radakovich N, Nagy M, Nazha A. Artificial intelligence in hematology: current challenges and opportunities. Curr Hematol Malig Rep. 2020;15:203–10. **This paper underscores the profound impact of Artificial Intelligence as a transformative tool in the field of medicine at large, with a special emphasis on its significance in hematology.**
17. Liu J, Yuan R, Li Y, Zhou L, Zhang Z, Yang J, et al. A deep learning method and device for bone marrow imaging cell detection. Ann Transl Med. 2022;10:208–208.
18•• Radakovich N, Sallman DA, Buckstein R, Brunner A, Dezern A, Mukerjee S, et al. A machine learning model of response to hypomethylating agents in myelodysplastic syndromes. iScience. 2022;25:104931. **This paper showcases the remarkable capabilities of Machine Learning within the realm of Myelodysplastic Syndromes by successfully predicting treatment outcomes approximately midway through the regimen of hypomethylating agents.**
19. Nath S, Marie A, Ellershaw S, Korot E, Keane PA. New meaning for NLP: the trials and tribulations of natural language processing with GPT-3 in ophthalmology. Br J Ophthalmol. 2022;106:889–92.
20. Li Y, Rao S, Solares JRA, Hassaine A, Ramakrishnan R, Canoy D, et al. BEHRT: Transformer for electronic health records. Sci Rep. 2020;10:7155.
21. Cunningham AR, Behm HE, Ju A, et al. Long-term survival of patients with glioblastoma of the pineal gland: a ChatGPT-assisted, updated case of a multimodal treatment strategy resulting in extremely long overall survival at a site with historically poor outcomes. Cureus. 2023;15(3):e36590.
22.• Johnson SB, King AJ, Warner EL, Aneja S, Kann BH, Bylund CL. Using ChatGPT to evaluate cancer myths and misconceptions: artificial intelligence and cancer information. JNCI Cancer Spectr. 2023;7(2):pkad015. **This paper highlights the substantial value of ChatGPT, one of the most cutting-edge and widely embraced AI tools of our time, in delivering accurate and reliable information regarding prevalent cancer myths and misconceptions.**
23. Sallam M. ChatGPT Utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. Healthcare. 2023;11:887.
24. Kimura K, Tabe Y, Ai T, Takehara I, Fukuda H, Takahashi H, et al. A novel automated image analysis system using deep convolutional neural networks can assist to differentiate MDS and AA. Sci Rep. 2019;9:13385.
25. Acevedo A, Merino A, Boldú L, Molina Á, Alférez S, Rodellar J. A new convolutional neural network predictive model for the automatic recognition of hypogranulated neutrophils in myelodysplastic syndromes. Comput Biol Med. 2021;134:104479.
26. Eckardt J-N, Schmittmann T, Riechert S, Kramer M, Sulaiman AS, Sockel K, et al. Deep learning identifies acute promyelocytic leukemia in bone marrow smears. BMC Cancer. 2022;22:201.
27. Radakovich N, Meggendorfer M, Malcovati L, Hilton CB, Sekeres MA, Shreve J, et al. A geno-clinical decision model for the diagnosis of myelodysplastic syndromes. Blood Adv. 2021;5:4361–9.
28.• 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. **This paper underscores the immense potential of machine learning in harnessing transcriptomics data to effectively classify and sub-categorize Acute Myeloid Leukemia.**
29.•• Nazha A, Komrokji R, Meggendorfer M, Jia X, Radakovich N, Shreve J, et al. Personalized prediction model to risk stratify patients with myelodysplastic syndromes. J Clin Oncol. 2021;39:3737–46. **This was the first paper to demonstrate the superiority of Machine Learning models over the existing prognostic models.**
30. Eckardt J-N, Röllig C, Metzeler K, Kramer M, Stasik S, Georgi J-A, et al. Prediction of complete remission and survival in acute myeloid leukemia using supervised machine learning. Haematologica. 2022;108:690–704.
31. Tazi Y, Arango-Ossa JE, Zhou Y, Bernard E, Thomas I, Gilkes A, et al. Unified classification and risk-stratification in acute myeloid leukemia. Nat Commun. 2022;13:4622.
32. Fuse K, Uemura S, Tamura S, Suwabe T, Katagiri T, Tanaka T, et al. Patient-based prediction algorithm of relapse after allo-HSCT for acute Leukemia and its usefulness in the decision-making process using a machine learning approach. Cancer Med. 2019;8:5058–67.
33. Shouval R, Labopin M, Bondi O, Mishan-Shamay H, Shimoni A, Ciceri 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.
34. Herold T, Jurinovic V, Batcha AMN, Bamopoulos SA, Rothenberg-Thurley M, Ksienzyk B, et al. A 29-gene and cytogenetic score for the prediction of resistance to induction treatment in acute myeloid leukemia. Haematologica. 2018;103:456–65.
35. Nazha A, Sekeres MA, Bejar R, Rauh MJ, Othus M, Komrokji RS, Barnard J, Hilton CB, Kerr CM, Steensma DP, DeZern A, Roboz G, Garcia-Manero G, Erba H, Ebert BL, Maciejewski JP. Genomic biomarkers to predict resistance to hypomethylating agents in patients with myelodysplastic syndromes using artificial intelligence. JCO Precis Oncol. 2019;3:PO.19.00119.