

Application of machine learning in the management of acute myeloid leukemia: current practice and future prospects

Jan-Niklas Eckardt,¹ Martin Bornhäuser,¹⁻³ Karsten Wendt,⁴ and Jan Moritz Middeke¹

¹Department of Internal Medicine I, University Hospital Carl Gustav Carus, Dresden, Germany; ²National Center for Tumor Diseases, Dresden (NCT/UCC), Dresden, Germany;

³German Consortium for Translational Cancer Research, DKFZ, Heidelberg, Germany; and ⁴Institute of Circuits and Systems, Technical University Dresden, Dresden, Germany

Machine learning (ML) is rapidly emerging in several fields of cancer research. ML algorithms can deal with vast amounts of medical data and provide a better understanding of malignant disease. Its ability to process information from different diagnostic modalities and functions to predict prognosis and suggest therapeutic strategies indicates that ML is a promising tool for the future management of hematologic malignancies; acute myeloid leukemia (AML) is a model disease of various recent studies. An integration of these ML techniques into various applications in AML management can assure fast and accurate diagnosis as well as precise risk stratification and optimal therapy. Nevertheless, these techniques come with various pitfalls and need a strict regulatory framework to ensure safe use of ML. This comprehensive review highlights and discusses recent advances in ML techniques in the management of AML as a model disease of hematologic neoplasms, enabling researchers and clinicians alike to critically evaluate this upcoming, potentially practice-changing technology.

Introduction

Despite recent research efforts, acute myeloid leukemia (AML) still poses a challenge in diagnosis and treatment alike, with curative options limited to a minority of cases.¹ In the past, numerous preclinical and clinical studies, often with multicenter cohorts of patients, have led to a better understanding of AML pathogenesis and classification and, subsequently, to improved treatment options. The rise of genomics has further improved our understanding of AML and resulted in novel modes of risk stratification² that were adopted in the European LeukemiaNet classification of AML.¹

The first studies of ML techniques in the diagnosis of hematologic malignancies were conducted 2 decades ago. They started with the recognition of leukemic cells from blood samples,^{3,4} flow cytometry,^{5,6} and the evaluation of genetic data,^{7,8} establishing the groundwork of ML methods in the investigation of hematologic malignancies. However, computational power was limited, and an integration of different diagnostic modalities on multidimensional data sets seemed out of immediate reach. From the first theoretical introduction of an artificial neuron by McCulloch and Pitts in 1943,⁹ the refinement of computational methods and ML approaches in the last decades, especially in neural networks, has opened up a variety of integrative approaches in the field of hematology. The ever-growing body of data from clinical studies, as well as new insights from preclinical models, poses a challenge for researchers and clinicians alike to organize and interpret said data to improve patient care.

It has been shown that ML is well suited for dealing with large amounts of complex data and may prove to be a powerful tool in understanding and overcoming disease.¹⁰⁻¹² Classically, diagnostic tests and patient data are interpreted by experienced clinicians who rely on years of medical education and training. However, ML algorithms have recently been shown to be on par with experts in a variety of tasks, from initial diagnosis, to prognosis estimation and prediction of treatment complications, to relapse

monitoring in hematologic malignancies. However, many ML approaches have still not found their way into everyday clinical practice due to a variety of hurdles and pitfalls.

The current comprehensive review provides an overview of recent studies of ML in AML diagnostics, prognostication, and treatment allocation. It discusses current challenges and pitfalls to improve studies of ML in AML and foster safe and informed clinical use of the presented techniques in the future.

Diagnosis

Currently, the initial diagnosis of AML relies on 4 pillars: cytomorphology, cytogenetics, molecular genetics, and immunophenotyping.¹ Given the evolving treatment stratification based on cytogenetic and molecular results, assigning patients to the best available treatment option seems appropriate.¹³ Therefore, precise characterization and classification of AML with high levels of accuracy are crucial for adequate therapy. Although our understanding of cancer in general has improved with the ever-increasing amount of genetic and genomic data, we still struggle to implement these large and complex data sets into clinical practice. ML approaches have shown tremendous potential in the analysis of complex genetic data.¹⁴

Analyzing >12 000 samples from >100 different studies, Warnat-Herresthal et al¹⁵ combined transcriptomic and genomic data with ML to develop classifiers that accurately detect AML in a near-automated and low-cost method. However, not every center interested in research in the applications of ML in cancer has such large data sets at hand. Fortunately, various freely accessible online data sets are available for multiple research purposes; these include the Leukemia Gene Atlas,¹⁶ Beat-AML,¹⁷ and The Cancer Genome Atlas.¹⁸ They allow researchers all across the globe to evaluate genetic risk profiles or identify novel genetic targets for individualized cancer therapy with the aid of ML techniques.¹⁹ This option could prove useful in the development of prospective basket trials of specific cancer-type overlapping mutations identified by ML algorithms. Support vector machines (SVMs), an ML technique that delineates data points in a coordinate system by calculating a hyperplane between distinct data sets, can be used for classification of high-dimensional data sets.²⁰ They can be applied to classify subtypes in large genomic data sets once pre-processing steps such as filtering for biomarker signatures or gene alterations have been performed to organize multidimensional data sets; the SVMs use these for classification, thereby revealing potential targets for therapy,^{21,22} and detect leukemic stem cells by genetic profiling.²³

A well-known disease-causing mutation of the FMS-like tyrosine kinase 3 (*FLT3*) occurs in almost one-third of AML cases, with internal tandem duplication (ITD) representing the most common *FLT3* mutation.²⁴ The combination of RNA-sequencing and genotyping with ML can distinguish malignant cell types and identify prototypic genetic lesions and an association of *FLT3*-ITD with progenitor-like cells.²⁵ SVM and random forest (RF), a combination of decision trees in which each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest,²⁶ are able to identify feature genes with the capacity to predict the mutation status of *FLT3*-ITD.²⁷

Deep neural networks (DNNs) can identify critical proteins associated with *FLT3*-ITD.²⁸ DNNs are a subset of ML that imitate the neuronal structure of the brain by creating interconnected artificial neural networks that can be applied for computer vision purposes, especially object detection, image segmentation, and classification.²⁹ After adequate pre-processing of image data, DNNs can be used in computer-aided diagnosis in cytomorphology. Key steps in DNN-based assessment of bone marrow and peripheral blood smears are cell segmentation, extraction, quantification of cell-specific features, and subsequent cell classification.³⁰ Especially in leukemia, precise recognition of white blood cells with various segmentation techniques (filtering, enhancement, edge detection, feature extraction, and classification)³¹ is crucial for correctly distinguishing between leukemic and non-leukemic cells.³²⁻³⁴ ML can use these techniques to analyze whole slides with automated focusing.³⁵ Classification of leukemia subtypes (AML, acute lymphoblastic leukemia, chronic myeloid leukemia, and chronic lymphocytic leukemia) can be achieved by a variety of ML approaches such as DNN,³⁶ SVM, and k-means-clustering (an unsupervised ML technique in which similar data points are grouped into *k* clusters according to their distance to a cluster mean).^{37,38}

Another essential part of the diagnostic process in AML is flow cytometry,³⁹ which can aid in the detection of relapse with a higher sensitivity than cytomorphology alone.⁴⁰ ML can be used to precisely distinguish between samples from AML patients and healthy individuals.⁴¹⁻⁴³ Computer-driven analysis of flow cytometry using clustering techniques (eg, FlowSOM) in combination with ML techniques (eg, SVM and RF) increases diagnostic accuracy in various hematologic malignancies⁴⁴ and correctly classifies rare cells.⁴⁵ FlowSOM is based on self-organizing maps to analyze flow or mass cytometry data, providing an overview of large sets of markers,⁴⁶ and it thereby aids in phenotyping leukemia and assessing measurable residual disease (MRD).⁴⁷ ML may thus provide an automated classification of data generated by flow cytometry and aid clinicians in their analysis and interpretation of the data by providing them with various differential diagnoses and their respective likelihood based on the given data. Hence, an integration of all diagnostic modalities in the evaluation of AML by the combination of different ML techniques seems feasible and provides a fast, automated, data-driven overview of each individual suspected case of AML for the medical professional to evaluate and verify.

Treatment and prognosis

The European Leukemia Net 2017 risk stratification divides patients with AML into favorable, intermediate, and adverse risk groups with distinct therapeutic implications and outcomes.¹ ML is advantageous in the early detection of potentially high-risk leukemias based on their individual genetic profile. Morita et al⁴⁸ analyzed bone marrow samples of 868 patients with myeloid leukemias (AML, myelodysplastic syndrome [MDS], chronic myelomonocytic leukemia, and myeloproliferative neoplasm) and generated an ML-based model that accurately predicts clinical phenotype based on somatic mutation data. Siddiqui et al⁴⁹ proposed an ML model based on clinical parameters known before treatment that predicts mortality rates for patients undergoing chemotherapy, thereby enabling clinicians to identify patients suitable for intensive induction regimens. DNN approaches have been shown to accurately predict

AML prognosis based on cytogenetics, mutational status, and age.⁵⁰ Gerstung et al⁵¹ have reported that large data sets combining clinical and genomic data in the form of knowledge banks can therefore be used to guide clinicians to precisely tailor a treatment approach for the individual patient; this method provides an accurate prediction of relapse, remission, and overall survival.

ML may even yield higher levels of accuracy compared with current standards. Fleming et al⁵² used RF and decision trees to predict survival prognosis in >2000 cases of non-APL/AML and reported a lower error rate for their model compared with the European Leukemia Net 2017 score. Similarly, Shreve et al⁵³ devised an ML model based on clinical, cytogenetic, and mutational data to predict personalized outcomes for the individual patient and reported a significantly better performance than the European Leukemia Net classification. Li et al⁵⁴ developed an algorithm for automatic classification of AML, MDS, and healthy samples based on >2000 patients limiting the number of flow cytometry markers while maintaining high levels of accuracy.

ML can be used to develop novel prognostic indices or refine the understanding of already established prognostic mutational markers. *NPM1* mutations are among the most commonly found mutations in AML, representing a distinct entity in the World Health Organization classification.⁵⁵ Patkar et al⁵⁶ identified genomic aberrations in *NPM1*^{mut} AML and developed a scoring system classifying *NPM1*^{mut} AML into 3 prognostic subgroups. Wagner et al⁵⁷ used an ML approach to associate a 3-gene expression signature consisting of *CALCRL*, *CD109*, and *LSP1* with overall survival, resulting in a prognostic score that includes gene expression levels and clinical data.

However, AML therapy remains challenging, and refractory disease poses a substantial threat for patient outcome.⁵⁸ ML can predict the likelihood of complete response in pediatric AML patients who received induction therapy based on gene expression patterns obtained through RNA sequencing.⁵⁹ Based on proteomics, ML can divide patients with AML into different treatment response groups, although combined use with clinical data may be essential.⁶⁰ MRD is an important marker for risk stratification and decision-making concerning therapeutic adjustments.^{61,62} Measurement of MRD can be improved by ML techniques such as SVM,⁶³⁻⁶⁵ and the evaluation of MRD is of growing importance in clinical decision-making in the management of AML. However, MRD evaluation is not available at all sites because it requires a high level of accuracy and technical expertise that can thus far only be achieved by specially equipped laboratories. ML techniques help implement MRD assessment in clinical practice by providing a highly standardized and data-driven approach based on the evaluation of large multicenter MRD data sets. Therefore, collaborative prospective studies between experienced laboratories and hematologic centers are needed to establish ML models that can accurately assess MRD; the goal is to provide to the clinical practice high-quality, standardized and automated MRD assessment verified by field experts.

Recently, a variety of novel therapeutic agents have been approved by both the US Food and Drug Administration and the European Medicines Agency for frontline treatment of patients with AML; long-term benefits remain uncertain, however, and study design as well as eligibility criteria may be flawed.⁶⁶⁻⁶⁸ ML provides the means

to improve patient selection, recruitment, and monitoring in clinical trials by assessing eligibility criteria, scanning electronic health records for suitable patients, or predicting the likelihood of failure or success in a trial.⁶⁹ ML models have been established in drug discovery and development, and DNNs especially show tremendous potential in identifying biomarkers and druggable targets and in the assessment of potential therapeutic molecules.^{70,71} The National Cancer Institute and the Dialogue on Reverse Engineering Assessment and Methods (DREAM) have launched challenges to develop ML tools to discover novel treatment strategies and detect drug-sensitive targets from genomic data.⁷² ML can use these large genomic data sets to predict targets for therapeutic agents. Lee et al⁷³ identified *SMARCA4* as a marker and driver of sensitivity to the topoisomerase II inhibitors mitoxantrone and etoposide, showing increased drug sensitivity both in ML models and in *in vitro* assays. Chen et al⁷⁴ used ML to assess potential *STAT3* inhibitors in AML and MDS. Janssen et al⁷⁵ developed drug discovery maps based on *t*-distributed stochastic neighbor embedding to predict novel inhibitors of *FLT3*, and Cutler and Fridman⁷⁶ generated an ML model that predicts high sensitivity to FLX925, a small molecule inhibitor of *FLT3*, in AML.

Despite the advent of targeted therapy, in the majority of AML cases, a curative treatment approach is still often limited to allogeneic stem cell transplantation, which harbors various risks. These risks include high treatment toxicity, infectious complications, graft-versus-host-disease (GVHD), transplant failure, and relapse.⁷⁷

Evaluating suitable patients for transplantation and patients at risk for complications is therefore crucial before starting conditioning therapy. Shouval et al⁷⁸ identified key variables to predict overall survival 100 days after transplantation in an analysis of >25 000 leukemia patients from the European Society for Blood and Marrow Transplantation with various ML techniques; they validated the scoring system in a prospective cohort study of 1848 patients from the Italian national transplant network.⁷⁹ The choice of conditioning regimen and post-grafting immunosuppression may therefore be guided by ML algorithms to design a personally tailored approach for the individual patient based on large databases of specific immunogenetic environments of patients undergoing allogeneic hematopoietic stem cell transplantation.⁸⁰ Relapse after transplantation can be estimated by using alternating decision trees.⁸¹ ML can also be used to predict development of acute GVHD after allogeneic transplantation⁸² and stratify outcomes in chronic GVHD, revealing novel groups at risk based on clinical phenotypes more accurately than current approaches based on cumulative severity.⁸³

Discussion

ML has already proven to be a versatile, precise, and robust tool in the diagnostic and therapeutic evaluation of AML, with a variety of challenges for future research as summarized in Table 1.

The efficiency of ML algorithms greatly depends on the quality and quantity of data they are trained on, as well as the selected end points and outcomes that researchers use. Therefore, large data sets are needed to construct and train such models.⁸⁴ With publicly available data sets, researchers have access to large amounts of training data for the development of ML tools, providing even small centers with the opportunity to conduct research in ML. Nevertheless, it is questionable how well ML models developed on online

Table 1. Applications and challenges of ML in the management of AML

Application	Cytomorphology/histology	Immunophenotyping	Clinical data	Cytogenetics/molecular genetics	New therapies/prognostic scores
Improvements needed	Precision of image segmentation (eg, detection of cell boundaries)	MRD evaluation (eg, standardization of cutoffs to form a decision boundary)	Integration of different high-dimensional data sets	Availability of data	Prospective studies for validation
			Number values such as laboratory results as well as written text are better evaluated by different ML techniques (integrative models needed)	Models trained on online data (eg, Beat-AML or The Cancer Genome Atlas) may not be accurate on regional data	Majority of ML studies are only retrospective; models have to be evaluated in a prospective manner to evaluate their translational application in patient care
	Feature extraction (eg, relation of nucleus to cytoplasm)	Classification methods	Standardization of clinical reports	Pre-processing of high-dimensional data	
		Which combination of different ML techniques shows the most accurate results?	Uniformity of clinical reports (eg, with standardized vocabulary) makes natural language processing easier	Accurate filtering of biosignatures is needed before classification	
	Cell classification				
	Labeling by field experts needed				
Outlook	• Integrated workflow of various ML techniques to guide clinical decision-making • Strict legal and regulatory framework to ensure patient safety • Prospective clinical trials to verify robustness of ML models • Physicians with basic knowledge in ML techniques to optimally implement ML into clinical practice				

data sets can perform in a regional setting. ML algorithms can automatize narrow repetitive tasks and thereby aid clinicians in accurately diagnosing AML as well as cutting time and effort in diagnostic steps such as the assessment of genetic, mutational, cytomorphologic, and flow cytometry data. The combination of different diagnostic modalities is crucial for the correct diagnosis of AML. An implementation of different ML techniques linking, for example, the results of flow cytometry, cytomorphology, and cytogenetics can provide clinicians with integrated ML tools to evaluate each suspected case of AML faster and with higher precision because adaptive ML tools are able to learn with every new case they are trained on, thus improving accuracy.

Furthermore, ML algorithms will help to better understand the complex interaction of distinct molecular subgroups in AML by identifying specific markers delineating specific groups of patients. These markers can also be evaluated in different hematologic entities (eg, MDS), and discoveries of overlapping disease-driving genetic alterations provide the opportunity for the development of prospective basket trials to create a directed therapy against disease-causing genes. Prognostic indices based on patient features derived by ML promise an unbiased view of potential markers of risk and adverse outcome and may refine current standards. Clinicians benefit from such tools as they partially free them from handling large amounts of data and equip them with methods to match individual patients with an ideal treatment.¹² ML therefore provides a powerful tool in the advent of precision medicine by identifying disease-specific genetic alterations and simultaneously recommending molecular structures that may be used to target these mutations for the individual patient. Once established, these algorithms could even provide hematologic expertise to regions without immediate access to large medical centers and help general practitioners to adequately screen for patients in need of hematologic assessment or treatment.

Nevertheless, tight regulation and oversight are crucial for the proper application of computer-aided diagnosis and treatment

allocation. A sophisticated structure of regulatory oversight, legal frameworks, and monitoring systems, adaptive to the fast pace of the current developments, is of the highest importance to ensure safe development and use of ML in everyday hematologic practice. The fast development of ML not only in the field of hematology but overall medical practice shows that computer skills are essential in medical training. We argue that basic knowledge of ML techniques, especially their potential in diagnostics and therapeutics as well as their limitations therein and potential for bias, is a key skill in medical education preparing for current and future developments and changes in practice. Future physicians should be taught to be critical users of ML in their practice, understanding how different models work, and what is in and out of reach of ML; this would enable them to properly integrate these methods into their practice and critically analyze and evaluate the data and recommendations provided by these tools.

As promising as these first results of ML in hematology may be, however, there is still a long road ahead. Clinicians and researchers should be aware of common pitfalls in ML when designing new studies.⁸⁵ For instance, many techniques require splitting research data into a training set and a testing set, and sometimes inadequate splits or hidden trends in data sets can falsify results. Researchers should also be wary for seemingly insignificant hidden variables that can influence ML models (eg, the placing of the scale bar in microscopic images). Furthermore, ML models can be overfit by feeding them biased data or making them catch “noise” instead of actual features, which results in a model that does suspiciously well in training sets but often cannot generalize well in test sets.⁸⁶ There is no gold standard in model selection, and every ML set-up depends on the use-case and research question at hand. It is also important to report on negative results and difficulties, given the variability of different ML set-ups, to work out collectively which approaches are more promising in which specific use-case. Many studies of ML in AML and cancer in general are still only retrospective and the underlying code of the ML model is often

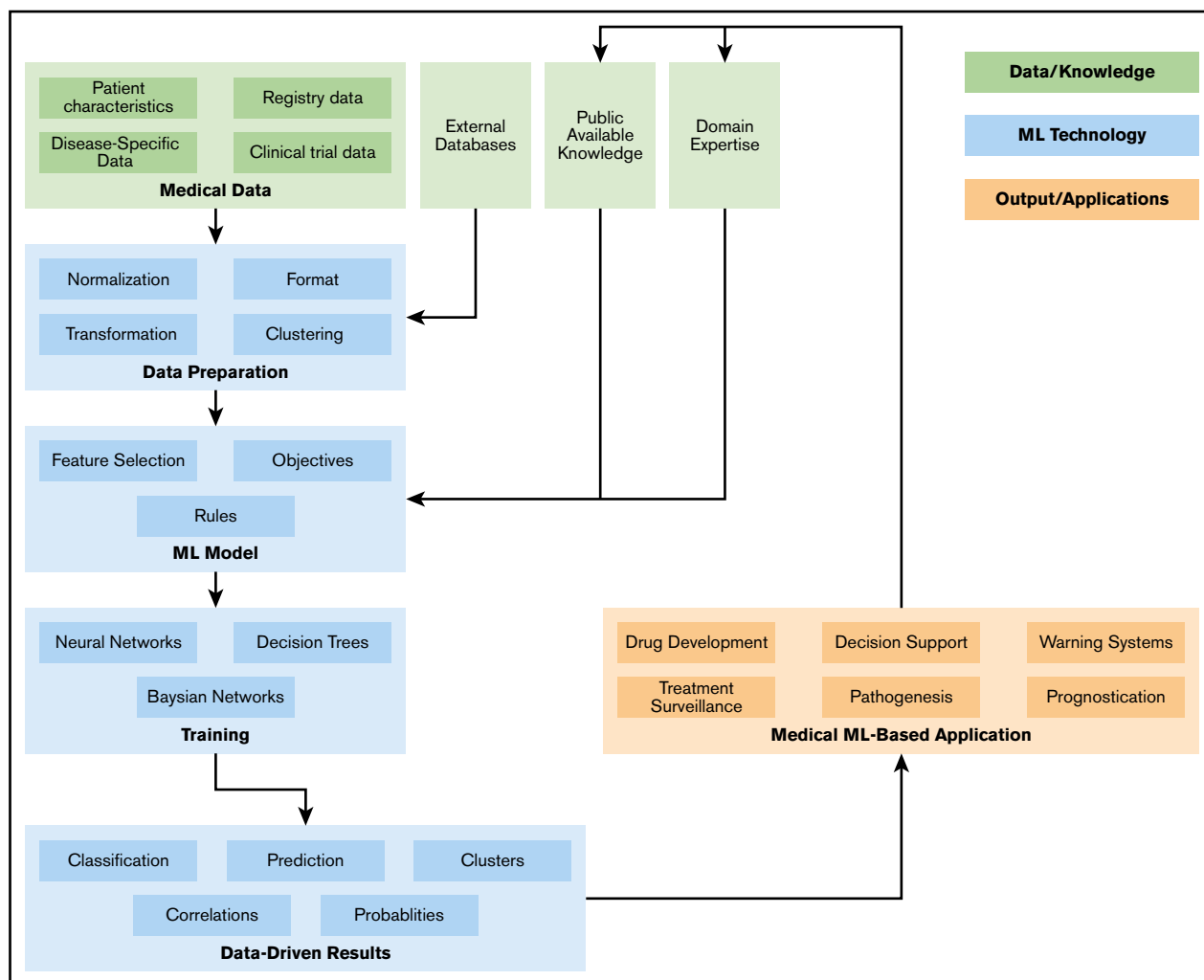


Figure 1. Overview of ML in the management of AML.

not reported, and therefore reports are at risk of bias.⁸⁷ Prospective validation in a proof-of-concept fashion is needed. ML-derived scoring systems, as well as individually tailored treatment approaches, can be verified in prospective trials proving or disproving the robustness of different models. Stand-alone ML tools based only on retrospective data are insufficient for widespread clinical use.

From a software point of view, prospective validation poses a challenge, because the large variety of available ML technology offers not only a vast amount of models to choose from but also may lead to flaws in study design because the optimal model for a distinct research question might not be chosen. At first, when conducting ML-based research, it is usually unknown which technology and parameterization lead to sufficient quality, especially for smaller data sets.⁸⁸ Therefore, software solutions should provide an iterative workflow to improve in a step-wise manner the ML set-up and integrate a broad spectrum of technology in a uniform way to support technological variety and progress.^{89,90}

As shown in Figure 1, an ML workflow consists of several data preprocessing and postprocessing steps, as well as meta-mechanics to optimize parameters and track objectives.^{91,92} To increase the

transparency of the approach and results, and to ensure reproducibility as well as comparability, a more abstract technical workflow description is required (eg, based on attribute grammars or model-based development approaches). Finally, domain experts (ie, physicians) need to have direct access to the ML workflow, as the translation of medical requirements, knowledge, and objectives by technicians implies obstacles and sources of errors. Hence, this should be minimized by adaptive, context-sensitive, and customizable user interfaces. Cooperation between study groups and a pooling of data sets may yield even more robust results.

In conclusion, ML in AML introduces a variety of novel and deeper insights in disease development and has the potential to significantly improve prognostication, personalized treatment, and patient surveillance. Close cooperation between computer scientists, data scientists, software developers, basic medical researchers, and physicians is imperative for sustained success and regulatory oversight. Legal frameworks are needed for safe and standardized development and use of ML tools for medical practice. The awareness of potential pitfalls of ML techniques and the knowledge gained from recent studies should lead to a more informed design of ML research. The goal is to create integrative tools that can

analyze and interpret data from multidimensional diagnostic modalities to further aid the clinician in everyday practice in diagnosis, prognostication, and treatment allocation of patients with AML, ultimately improving patient outcome.

Authorship

Contribution: J.-N.E. performed the literature search and wrote the draft; J.M.M., M.B., and K.W. edited the manuscript; and all authors revised and approved the manuscript.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

ORCID profiles: J.-N.E., 0000-0002-3649-2823; K.W., 0000-0001-9373-6751.

Correspondence: Jan-Niklas Eckardt, Department of Internal Medicine I, University Hospital Carl Gustav Carus, Fetscherstr 74, 01307 Dresden, Germany; e-mail: jan-niklas.eckardt@uniklinikum-dresden.de.

References

- Döhner H, Estey E, Grimwade D, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood*. 2017;129(4):424-447.
- Bullinger L, Döhner K, Döhner H. Genomics of acute myeloid leukemia diagnosis and pathways. *J Clin Oncol*. 2017;35(9):934-946.
- Zini G, d'Onofrio G. Neural network in hematopoietic malignancies. *Clin Chim Acta*. 2003;333(2):195-201.
- Adjouadi M, Ayala M, Cabrerizo M, Zong N, Lizarraga G, Rossman M. Classification of leukemia blood samples using neural networks. *Ann Biomed Eng*. 2010;38(4):1473-1482.
- Zong N, Adjouadi M, Ayala M. Optimizing the classification of acute lymphoblastic leukemia and acute myeloid leukemia samples using artificial neural networks. *Biomed Sci Instrum*. 2006;42:261-266.
- Kothari R, Cuaing H, Balachander T. Neural network analysis of flow cytometry immunophenotype data. *IEEE Trans Biomed Eng*. 1996;43(8):803-810.
- Lyons-Weiler J, Patel S, Bhattacharya S. A classification-based machine learning approach for the analysis of genome-wide expression data. *Genome Res*. 2003;13(3):503-512.
- Berrar DP, Downes CS, Dubitzky W. Multiclass cancer classification using gene expression profiling and probabilistic neural networks. *Pac Symp Biocomput*. 2003;5-16.
- McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. 1943. *Bull Math Biol*. 1990;52(1-2):99-115, NaN-97.
- Sajda P. Machine learning for detection and diagnosis of disease. *Annu Rev Biomed Eng*. 2006;8(1):537-565.
- Al-Jarrah OY, Yoo PD, Muhaidat S, Karagiannidis GK, Taha K. Efficient machine learning for big data: a review. *Big Data Res*. 2015;2(3):87-93.
- Rajkomar A, Dean J, Kohane I. Machine learning in medicine. *N Engl J Med*. 2019;380(14):1347-1358.
- Röllig C, Kramer M, Schliemann C, et al. Time from diagnosis to treatment does not affect outcome in intensively treated patients with newly diagnosed acute myeloid leukemia. *Blood*. 2019;134(suppl 1):13.
- Libbrecht MW, Noble WS. Machine learning applications in genetics and genomics. *Nat Rev Genet*. 2015;16(6):321-332.
- Warnat-Herresthal S, Perrakis K, Taschler B, et al. Scalable prediction of acute myeloid leukemia using high-dimensional machine learning and blood transcriptomics. *iScience*. 2020;23(1):100780.
- Hebestreit K, Gröttrup S, Emden D, et al. Leukemia gene atlas—a public platform for integrative exploration of genome-wide molecular data. *PLoS One*. 2012;7(6):e39148.
- Beat AML Functional Genomic Study—National Cancer Institute. Beat AML 1.0: a collaborative program for functional genomic data integration. <https://www.cancer.gov/about-nci/organization/ccg/blog/2019/beataml>. Accessed 1 September 2020.
- The Cancer Genome Atlas Program—National Cancer Institute. The Cancer Genome Atlas Program. <https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga>. Accessed 1 September 2020.
- Benard B, Gentles AJ, Köhnke T, Majeti R, Thomas D. Data mining for mutation-specific targets in acute myeloid leukemia. *Leukemia*. 2019;33(4):826-843.
- Noble WS. What is a support vector machine? *Nat Biotechnol*. 2006;24(12):1565-1567.
- Vasighizaker A, Sharma A, Dehzangi A. A novel one-class classification approach to accurately predict disease-gene association in acute myeloid leukemia cancer. *PLoS One*. 2019;14(12):e0226115.
- Huang S, Cai N, Pacheco PP, Narrandes S, Wang Y, Xu W. Applications of support vector machine (SVM) learning in cancer genomics. *Cancer Genomics Proteomics*. 2018;15(1):41-51.
- Li J, Lu L, Zhang YH, et al. Identification of leukemia stem cell expression signatures through Monte Carlo feature selection strategy and support vector machine. *Cancer Gene Ther*. 2020;27(1-2):56-69.
- Daver N, Schlenk RF, Russell NH, Levis MJ. Targeting FLT3 mutations in AML: review of current knowledge and evidence. *Leukemia*. 2019;33(2):299-312.
- van Galen P, Hovestadt V, Wadsworth LH, et al. Single-cell RNA-Seq reveals AML hierarchies relevant to disease progression and immunity. *Cell*. 2019;176(6):1265-1281.e24.

26. Breiman L. Random forests. *Mach Learn*. 2001;45(1):5-32.
27. Li C, Zhu B, Chen J, Huang X. Feature genes predicting the FLT3/ITD mutation in acute myeloid leukemia. *Mol Med Rep*. 2016;14(1):89-94.
28. Liang CA, Chen L, Wahed A, Nguyen AND. Proteomics analysis of FLT3-ITD mutation in acute myeloid leukemia using deep learning neural network. *Ann Clin Lab Sci*. 2019;49(1):119-126.
29. Guo Y, Liu Y, Oerlemans A, Lao S, Wu S, Lew MS. Deep learning for visual understanding: a review. *Neurocomputing*. 2016;187:27-48.
30. Rodellar J, Alf  rez S, Acevedo A, Molina A, Merino A. Image processing and machine learning in the morphological analysis of blood cells. *Int J Lab Hematol*. 2018;40(suppl 1):46-53.
31. Umamaheswari D, Geetha S. Review on image segmentation techniques incorporated with machine learning in the scrutinization of leukemic microscopic stained blood smear images. In: Pandian D, Fernando X, Baig Z, Shi F, eds. *Proceedings of the International Conference on ISMAC in Computational Vision and Bio-Engineering 2018 (ISMAC-CVB)*. Lecture Notes in Computational Vision and Biomechanics. Cham, Switzerland: Springer International Publishing; 2019:1773-1791.
32. Chandradevan R, Aljudi AA, Drumheller BR, et al. Machine-based detection and classification for bone marrow aspirate differential counts: initial development focusing on nonneoplastic cells. *Lab Invest*. 2020;100(1):98-109.
33. Su J, Liu S, Song J. A segmentation method based on HMRF for the aided diagnosis of acute myeloid leukemia. *Comput Methods Programs Biomed*. 2017;152:115-123.
34. Bigorra L, Merino A, Alf  rez 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*. 2017;31(2):e22024.
35. Rai Dastidar T, Ethirajan R. Whole slide imaging system using deep learning-based automated focusing. *Biomed Opt Express*. 2019;11(1):480-491.
36. Ahmed N, Yigit A, Isik Z, Alpkocak A. Identification of leukemia subtypes from microscopic images using convolutional neural network. *Diagnostics (Basel)*. 2019;9(3):E104.
37. Jagadev P, Virani HG. Detection of leukemia and its types using image processing and machine learning. In: *2017 International Conference on Trends in Electronics and Informatics (ICEI)*, Tirunelveli, India. New York, NY: IEEE; 2017:522-526.
38. Paswan S, Rathore YK. Detection and classification of blood cancer from microscopic cell images using SVM KNN and NN classifier. *Int J Adv Res Ideas Innov Technol*. 2017;3(6):315-324.
39. Chen X, Cherian S. Acute myeloid leukemia immunophenotyping by flow cytometric analysis. *Clin Lab Med*. 2017;37(4):753-769.
40. Zhou Y, Wood BL, Walter RB, et al. Is there a need for morphologic exam to detect relapse in AML if multi-parameter flow cytometry is employed? *Leukemia*. 2017;31(11):2536-2537.
41. Angeletti C. A method for the interpretation of flow cytometry data using genetic algorithms. *J Pathol Inform*. 2018;9(1):16.
42. Manninen T, Huttunen H, Ruusuvaari P, Nykter M. Leukemia prediction using sparse logistic regression. *PLoS One*. 2013;8(8):e72932.
43. Biehl M, Bunte K, Schneider P. Analysis of flow cytometry data by matrix relevance learning vector quantization. *PLoS One*. 2013;8(3):e59401.
44. Duetz C, Bachas C, Westers TM, van de Loosdrecht AA. Computational analysis of flow cytometry data in hematological malignancies: future clinical practice? *Curr Opin Oncol*. 2020;32(2):162-169.
45. Qiu P. Computational prediction of manually gated rare cells in flow cytometry data. *Cytometry A*. 2015;87(7):594-602.
46. Van Gassen S, Callebaut B, Van Helden MJ, et al. FlowSOM: using self-organizing maps for visualization and interpretation of cytometry data. *Cytometry A*. 2015;87(7):636-645.
47. Lacombe F, Dupont B, Lechevalier N, Vial JP, B  n   MC. Innovation in flow cytometry analysis: a new paradigm delineating normal or diseased bone marrow subsets through machine learning. *HemaSphere*. 2019;3(2):e173.
48. Morita K, Wang F, Makishima H, et al. Pan-myeloid leukemia analysis: machine learning-based approach to predict phenotype and clinical outcomes using mutation data. *Blood*. 2018;132(suppl 1):1801.
49. Siddiqui NS, Klein A, Godara A, Varga C, Buchsbaum RJ, Hughes MC. Supervised machine learning algorithms using patient related factors to predict in-hospital mortality following acute myeloid leukemia therapy. *Blood*. 2019;134(suppl 1):3435.
50. Lin M, Jaitly V, Wang I, et al. Application of deep learning on predicting prognosis of acute myeloid leukemia with cytogenetics, age, and mutations. *ArXiv181013247 Cs Q-Bio Stat*. <http://arxiv.org/abs/1810.13247>. Accessed 13 April 2020.
51. Gerstung M, Papaemmanuil E, Martincorena I, et al. Precision oncology for acute myeloid leukemia using a knowledge bank approach. *Nat Genet*. 2017;49(3):332-340.
52. Fleming S, Tsai CH, D  hner H, et al. Use of machine learning in 2074 cases of acute myeloid leukemia for genetic risk profiling. *Blood*. 2019;134(suppl 1):1392.
53. Shreve J, Meggendorfer M, Awada H, et al. A personalized prediction model to risk stratify patients with acute myeloid leukemia (AML) using artificial intelligence. *Blood*. 2019;134(suppl 1):2091. doi:10.1182/blood-2019-128066
54. Li J, Wang Y, Ko B, Li C, Tang J, Lee C. Learning a cytometric deep phenotype embedding for automatic hematological malignancies classification. In: *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. New York, NY: IEEE; 2019:1733-1736.
55. Heath EM, Chan SM, Minden MD, Murphy T, Shlush LI, Schimmer AD. Biological and clinical consequences of NPM1 mutations in AML. *Leukemia*. 2017;31(4):798-807.
56. Patkar N, Shaikh AF, Kakirde C, et al. A novel machine-learning-derived genetic score correlates with measurable residual disease and is highly predictive of outcome in acute myeloid leukemia with mutated NPM1. *Blood Cancer J*. 2019;9(10):79.

57. Wagner S, Vadakekolathu J, Tasian SK, et al. A parsimonious 3-gene signature predicts clinical outcomes in an acute myeloid leukemia multicohort study. *Blood Adv*. 2019;3(8):1330-1346.
58. Rashidi A, Weisdorf DJ, Bejanyan N. Treatment of relapsed/refractory acute myeloid leukaemia in adults. *Br J Haematol*. 2018;181(1):27-37.
59. 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.
60. Cheboub L, Miannay B, Boughaci D, Guziolowski C. Discriminate the response of acute myeloid leukemia patients to treatment by using proteomics data and answer set programming. *BMC Bioinformatics*. 2018;19(suppl 2):59.
61. Schuurhuis GJ, Heuser M, Freeman S, et al. Minimal/measurable residual disease in AML: a consensus document from the European LeukemiaNet MRD Working Party. *Blood*. 2018;131(12):1275-1291.
62. Hourigan CS, Gale RP, Gormley NJ, Ossenkoppele GJ, Walter RB. Measurable residual disease testing in acute myeloid leukaemia. *Leukemia*. 2017;31(7):1482-1490.
63. Ko BS, Wang YF, Li JL, 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.
64. Ni W, Hu B, Zheng C, et al. Automated analysis of acute myeloid leukemia minimal residual disease using a support vector machine. *Oncotarget*. 2016;7(44):71915-71921.
65. Voigt AP, Eidenschink Brodersen L, Pardo L, Meshinchi S, Loken MR. Consistent quantitative gene product expression: #1. Automated identification of regenerating bone marrow cell populations using support vector machines. *Cytometry A*. 2016;89(11):978-986.
66. Estey E, Karp JE, Emadi A, Othus M, Gale RP. Recent drug approvals for newly diagnosed acute myeloid leukemia: gifts or a Trojan horse? *Leukemia*. 2020;34(3):671-681.
67. Estey EH, Gale RP, Sekeres MA. New drugs in AML: uses and abuses. *Leukemia*. 2018;32(7):1479-1481.
68. Estey E, Gale RP. Acute myeloid leukemia therapy and the chosen people. *Leukemia*. 2017;31(2):269-271.
69. Harrer S, Shah P, Antony B, Hu J. Artificial intelligence for clinical trial design. *Trends Pharmacol Sci*. 2019;40(8):577-591.
70. Vamathevan J, Clark D, Czodrowski P, et al. Applications of machine learning in drug discovery and development. *Nat Rev Drug Discov*. 2019;18(6):463-477.
71. Jing Y, Bian Y, Hu Z, Wang L, Xie XQ. Deep learning for drug design: an artificial intelligence paradigm for drug discovery in the big data era [published correction appears in *AAPS J*. 2018;20(4):79]. *AAPS J*. 2018;20(3):58.
72. Costello JC, Heiser LM, Georgii E, et al; NCI DREAM Community. A community effort to assess and improve drug sensitivity prediction algorithms. *Nat Biotechnol*. 2014;32(12):1202-1212.
73. Lee SI, Celik S, Logsdon BA, et al. A machine learning approach to integrate big data for precision medicine in acute myeloid leukemia. *Nat Commun*. 2018;9(1):42.
74. Chen X, Chen HY, Chen ZD, Gong JN, Chen CYC. A novel artificial intelligence protocol for finding potential inhibitors of acute myeloid leukemia. *J Mater Chem B Mater Biol Med*. 2020;8(10):2063-2081.
75. Janssen APA, Grimm SH, Wijdeven RHM, et al. Drug discovery maps, a machine learning model that visualizes and predicts kinome-inhibitor interaction landscapes. *J Chem Inf Model*. 2019;59(3):1221-1229.
76. Cutler G, Fridman JS. A machine-learning analysis suggests that FLX925, a FLT3/CDK4/6 kinase inhibitor, is potent against FLT3-wild type tumors via its CDK4/6 activity. *Blood*. 2016;128(22):3520.
77. Loke J, Malladi R, Moss P, Craddock C. The role of allogeneic stem cell transplantation in the management of acute myeloid leukaemia: a triumph of hope and experience. *Br J Haematol*. 2020;188(1):129-146.
78. Shouval R, Labopin M, Unger R, 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(3):e0150637.
79. Shouval R, Bonifazi F, Fein J, et al. Validation of the acute leukemia-EBMT score for prediction of mortality following allogeneic stem cell transplantation in a multi-center GITMO cohort. *Am J Hematol*. 2017;92(5):429-434.
80. Bornhäuser M. Conditioning intensity and antilymphocyte globulin: towards personalized transplant strategies? *Haematologica*. 2019;104(6):1101-1102.
81. Fuse K, Uemura S, Tamura S, 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(11):5058-5067.
82. Arai Y, Kondo T, Fuse K, et al. Using a machine learning algorithm to predict acute graft-versus-host disease following allogeneic transplantation. *Blood Adv*. 2019;3(22):3626-3634.
83. Gandelman JS, Byrne MT, Mistry AM, et al. Machine learning reveals chronic graft-versus-host disease phenotypes and stratifies survival after stem cell transplant for hematologic malignancies. *Haematologica*. 2019;104(1):189-196.
84. Radakovich N, Nagy M, Nazha A. Artificial intelligence in hematology: current challenges and opportunities. *Curr Hematol Malig Rep*. 2020;15(3):203-210.
85. Riley P. Three pitfalls to avoid in machine learning. *Nature*. 2019;572(7767):27-29.
86. Cawley GC, Talbot NLC. On over-fitting in model selection and subsequent selection bias in performance evaluation. *J Mach Learn Res*. 2010;11:2079-2107.

87. Nagendran M, Chen Y, Lovejoy CA, et al. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. *BMJ*. 2020;368:m689.
88. Zhang Y, Ling C. A strategy to apply machine learning to small datasets in materials science. *Npj Comput Math*. 2018;4(1):25.
89. Baltrušaitis T, Ahuja C, Morency LP. Multimodal machine learning: a survey and taxonomy. *IEEE Trans Pattern Anal Mach Intell*. 2019;41(2):423-443.
90. Nguyen G, Dlugolinsky S, Bobák M, et al. Machine learning and deep learning frameworks and libraries for large-scale data mining: a survey. *Artif Intell Rev*. 2019;52(1):77-124.
91. Bottou L, Curtis FE, Nocedal J. Optimization methods for large-scale machine learning. *SIAM Rev*. 2018;60(2):223-311.
92. Chandrashekar G, Sahin F. A survey on feature selection methods. *Comput Electr Eng*. 2014;40(1):16-28.