JAMA Oncology | Review | AI IN ONCOLOGY

Emerging Artificial Intelligence Technologies for Risk Assessment and Management in Acute Myeloid Leukemia
A Review

Mohammad Amin Ansarian, MBBS; Mahsa Fatahichegeni, MBBS; Rui Xu, MD; Ying Chen, MD, PhD; Xiaoning Wang, MD, PhD; Juan Ren, MD; Huasheng Liu, MD, PhD

IMPORTANCE Acute myeloid leukemia (AML) is a severe hematologic cancer with complex genetic heterogeneity necessitating personalized treatment approaches. Artificial intelligence (AI) technologies may revolutionize risk stratification, diagnosis enhancement, and treatment planning in addressing critical gaps in AML management, particularly in low-resource health care environments.

OBSERVATIONS This narrative review synthesizes existing AI applications in 3 primary areas of AML management. Machine learning algorithms integrating clinical, cytogenetic, and molecular data demonstrate greater prognostic accuracy than conventional European LeukemiaNet (ELN) guidelines. Deep learning approaches to image analysis yield excellent results for AML subtype identification from bone marrow smears (area under the receiver operating characteristic curve [AUROC]: 0.97) and genetic variant prediction (eg, *NPM1* status [AUROC: 0.92]). AI-driven genomic analysis reveals novel prognostic signatures and therapeutic targets through advanced pattern recognition, with high-dimensional machine learning achieving greater than 99% accuracy in AML classification from transcriptomic data. Explainable AI models overcome the black box limitation through interpretable algorithms with Shapley Additive Explanations values and local interpretable model-agnostic explanation techniques. Federated learning approaches enable multi-institutional collaboration with protection of patient privacy, with 96.5% accuracy in leukemia classification on heterogeneous datasets.

CONCLUSIONS AND RELEVANCE AI technologies hold potential to improve AML treatment through enhanced risk stratification, early detection capabilities, and individualized treatment optimization. The transition toward explainable AI models is essential to clinical readiness, with federated learning architectures resolving data scarcity concerns. Seamless integration requires harmonized data standards, robust regulatory frameworks, and equitable access to technology to fully realize the transformative potential of AI in improving outcomes for patients with AML globally.

JAMA Oncol. doi:10.1001/jamaoncol.2025.3601
Published online November 6, 2025.

Supplemental content

CME at jamacmelookup.com

Author Affiliations: Department of Hematology, The First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, Shaanxi, People's Republic of China.

Corresponding Author: Huasheng Liu, MD, PhD, Department of Hematology, The First Affiliated Hospital of Xi'an Jiaotong University, 277 W Yanta Rd, Xi'an, Shaanxi, PR China (lhs681995@126.com).

European LeukemiaNet (ELN) genetic risk categorization integrates molecular markers for improved acute myeloid leukemia (AML) prognosis, enabling automated risk stratification and personalized therapeutic strategies.[1-3] However, treatment remains challenging, particularly in low-income and middle-income nations with limited access to advanced therapies and diagnostic infrastructure.[4] This review examines AI applications in AML management with 3 objectives: to (1) evaluate the efficacy of artificial intelligence (AI) techniques in risk stratification and early detection, (2) analyze the contribution of AI to image-based diagnostics and genomic profiling, and (3) investigate challenges and opportunities in AI-driven AML management.

The 2024 ELN recommendations highlighted the integration of novel genetic markers with traditional prognostic factors like cytogenetics, age, and disease status to refine risk stratification for patients with AML.[5] AI applications may improve AML management by enabling genetic feature prediction from routine bone marrow smears, reducing reliance on expensive genetic testing.[1,6,7] Deep neural networks achieve accuracy in bone marrow cell morphology analysis.[8] Despite these advancements, obstacles persist concerning data consistency, model scalability, and equitable AI technology access in health care. Current challenges include insufficient sample sizes, poor predictive performance, and a lack of generalization capabilities.[6] Recent research highlights how federated learning (FL) and foundation models (FMs) address these challenges by enabling collaborative AI model development without sharing sensitive patient data, particularly valuable in hematologic cancers.[9,10]

Table 1. Machine Learning Methods and Their Effectiveness

AI method	Key features	Main findings	Reference
Multilayer perceptron neural network	Dataset: 3687 patients with AML (IC, 3030; AZA, 657); 52 diagnostic variables	68.5% Accuracy for IC cohort prediction; 62.1% accuracy for AZA cohort prediction	Didi et al,[19] 2024
Graph neural network	Used first 6 mo of clinical data; integrated electronic health records and pathology data	Comparable AUROC to ELN criteria; identified high-risk cases without direct variant data	Sinha et al,[18] 2023
Decision tree	Combined datasets from Taiwan (n = 759) and Germany/Austria (n = 1315); used recursive partitioning and random forest	Reclassified 45% of cases vs ELN 2017; reduced error rate from 45% to 35%	Fleming et al,[24] 2019

Abbreviations: AI, artificial intelligence; AML, acute myeloid leukemia; AUROC, area under the receiver operating characteristic curve; AZA, azacitidine; ELN, European LeukemiaNet; IC, intensive chemotherapy.

AI Model Validation and Standards

Cross-validation helps determine AI model reliability and clinical usefulness. Our literature synthesis provides a validation framework with complementary approaches. The most widely used performance measures are accuracy and sensitivity estimates, often complemented by receiver operating characteristic (ROC) curves and area under the curve (AUC or area under the receiver operating characteristic curve [AUROC]). ROC curves plot true positive rates vs false positive rates at different thresholds[11] whereas AUROC ($0 \leq \text{AUROC} \leq 1$) provides a single measurement of classification effectiveness, with 1 representing perfect classification.[12] Combined with precision values, these measures offer deeper insights into model performance.

External validation determines model reliability across varied populations and clinical environments.[13] Research consistently shows worse model performance on external validation,[14] highlighting the requirement for multicenter validation and diverse patient groups to determine AI model applicability in clinical practice.

Models must demonstrate reproducibility. Several research groups addressed this by releasing code, models, and deidentified datasets to the research community. However, substantial obstacles exist in standardizing data acquisition, processing, and reporting protocols across institutions and research groups. Image acquisition protocol variability, data preprocessing algorithm differences, and model architecture specifications affect reproducibility and must be considered when testing AI models for use.

To address standardization challenges, we provide data requirements and formatting guidelines for future AML AI research (eTable 1 in the Supplement), including imaging data formats, genomic variant standards, clinical metadata, Health Insurance Portability and Accountability Act–compliant[15] deidentification procedures, and dataset versioning protocols.

With these reproducibility and external validation principles, we describe AI techniques for AML risk stratification, focusing on machine learning (ML), image analysis, and genomic applications. A summary of key studies, methods, data inputs, cohort sizes, validation approaches, and performance metrics appears in eTable 2 in the Supplement.

Quality and Bias Assessment

We selected 24 studies using multistage screening for quality assessment using QUADAS-AI[16] (eFigure 1 in the Supplement). The detailed assessment framework is included in eTable 3 in the Supplement. Many studies showed high risk or unclear risk of bias for patient selection (17 of 24 [70.8%]) and index tests (14 of 24 [58.3%]), often failing to provide clear patient descriptions or lacking adequate external evaluation. More studies exhibited low risk of bias for reference standards (12 of 24 [50.0%]) and workflow timing (14 of 24 [58.3%]), indicating appropriate reference standard selection and reasonable timing. Only 5 studies (20.8%) demonstrated low risk of bias across all domains, suggesting the need for improved methodological quality in future AI-based hematological analysis studies.

AI Methods in AML Risk Assessment

AI methods can reshape AML management through ML models that integrate clinical/molecular data, deep learning (DL) algorithms that detect subtypes from images, and genomic analyses that identify prognostic signatures. This integrated approach substantially refines risk stratification by moving beyond traditional criteria to generate a more dynamic and patient-specific risk profile

ML

ML algorithms from gradient boosting to graph neural networks (GNNs) consistently exhibit enhanced predictive accuracy for survival outcomes and treatment responses.[17-19] ML algorithms predict 30-day mortality, complete remission rates, and 2-year survival by integrating clinical, cytogenetic, and molecular data accurately.[20-23]

The transition from single-model to ensemble approaches has been associated with substantially improved prediction accuracy. Multilayer perceptron established an early benchmark with 68.5% accuracy, while GNNs achieved comparable performance with conventional clinical parameters without directly incorporating variant data. The decision tree model's ability to reclassify 45% of cases compared with ELN 2017 criteria, reducing error rates from 45% to 35%, demonstrates tangible benefits of advanced analytical approaches in personalizing treatment planning[19] (**Table 1**[18,19,24]).

Incorporating patient age alongside molecular and genetic data enhances AML risk stratification. Ensemble models integrating clinical and genetic data with ELN 2022 guidelines demonstrated improved prognostic capabilities.[23] Hierarchical FL frameworks enable multi-institutional collaboration without compromising privacy. These privacy-preserving techniques achieved excellent leukemia classification performance (96.5% accuracy) while maintaining data confidentiality, which is particularly valuable for rare genetic subtypes for which institutions may have limited cases.[9] ML approaches refined ELN risk groups by incorporating gene variants and variant allele fractions, identifying heterogeneous outcomes

Table 2. Image Analysis Applications

AI method	Approach	Results	Reference
Deep learning platform	Analyzed bone marrow smears; automatic cell segmentation	High precision in APL detection; ROC of 0.86 for APL vs non-APL	Eckardt et al,[27] 2022
Multistep deep learning	Automated cell segmentation; morphological feature analysis	AUROC of 0.97 for AML detection; AUROC of 0.92 for *NPM1* variant prediction	Eckardt et al,[22] 2022
Convolutional neural network	171 374 Bone marrow images; 627 validation images	Automated classification of 22 leukocyte types; high accuracy in cell type identification	Matek et al,[29] 2021

Abbreviations: AI, artificial intelligence; AML, acute myeloid leukemia; APL, acute promyelocytic leukemia; AUROC, area under the receiver operating characteristic curve; ROC, receiver operating characteristic.

within favorable and intermediate groups that were previously indistinguishable.[25]

AI Techniques for Image Analysis

DL has been associated with improved image analysis for AML diagnosis and classification. Neural networks achieved marked accuracy in bone marrow cell morphology classification,[7] while computer-aided systems achieved improved cell classification and blast counting performance.[26] These technologies demonstrate accuracy in AML detection from bone marrow smears,[22,27] AML subtype classification (like acute promyelocytic leukemia [APL]),[27] and genetic characteristic prediction, such as *NPM1* variant status.[7,22]

Modern AI systems automatically segment cells, extract morphological features, and accurately classify leukocytes. Multiple-instance learning models using whole-slide images without manual annotations attained high accuracy (AUC of 0.90 for *NPM1* and 0.81 for *FLT3-ITD*) in predicting key genetic alterations.[2] Holotomography integration with DL enabled *NPM1* variant detection in AML with excellent performance (AUC, 0.9375) through 3-dimensional cell morphology analysis.[28] These methods facilitate rapid diagnosis, particularly in settings without advanced genetic testing capabilities.

The DL platform[27] attains an ROC of 0.86 in detecting APL. Multistep DL approaches[22] exhibit significant performance, achieving AUROC values of 0.97 for AML detection and 0.92 for predicting *NPM1* variant status, integrating morphological analysis with genetic profiling. CNN classification of 22 leukocyte categories from more than 170 000 images demonstrated the precision and scalability of contemporary AI systems in complex cellular analysis.[29] ResNet-50 (Microsoft) implementations exhibited efficacy in detecting blast cells for AML diagnosis, with markedly improved sensitivity compared with traditional approaches.[30] Dual-attention network augmentations enhanced classification performance, achieving accuracy rates exceeding 99% in classifying white blood cells across multiple heterogeneous datasets[31] (**Table 2**[22,27,29]).

Adami et al[26] developed an end-to-end DL pipeline that automated bone marrow smear analysis. This method integrates region-of-interest detection using U-Net and cell classification through Faster R-CNN (Microsoft), resulting in 87% overall accuracy across 20 cell types, including myeloblasts and promyelocytes. This approach is associated with decreased processing time and interobserver variability, rendering it beneficial in resource-limited environments. StyleGAN2 (NVIDIA) adaptive discriminator augmentation synthetic bone marrow images showed promise in mitigating data scarcity and privacy concerns.[32] Synthetic images facilitate training highly accurate classification models (AUC > 0.95) while preserving patient privacy, enhancing the potential of AI to improve diagnostic workflows without advanced genetic testing.

AI in Genomics and Transcriptomics

ML and AI applications significantly advanced AML research through genomic and transcriptomic data analysis, yielding enhanced prognostic capabilities.[1,33] Researchers created models that predicted factors associated with patient survival,[34] recognized typical transcriptomic signatures with corresponding genomic subtypes,[35] and classified AML based on gene expression profiles.[33] These approaches potentially represent paths toward personalized therapeutic strategies that many improve patient outcomes.[35]

FMs represent a paradigm shift in AI development for genomic analysis in AML. Domain-specific FMs leverage contrastive learning, generative learning, and fine-tuning to support diverse health care applications from diagnosis to treatment recommendation.[36] These models process multimodal data, including genomic sequences, medical images, and clinical texts, with unprecedented efficiency. Parameter-efficient, fine-tuning techniques enable adaptation of powerful models to specific hematologic contexts with limited data or computational resources through low-rank adaptation.[37,38] Explainable GNN frameworks demonstrated capability in integrating multiomics information, offering insights into cancer development and heterogeneity that were previously unattainable with conventional methods[39] (**Table 3**[6,33,40]).

Multicenter Data Validation

Rigorous validation across independent medical centers ensures AI model generalizability and robustness for AML. The Artificial Intelligence Prediction of Acute Leukemia model underwent validation across 6 independent university hospital databases in France.[41] This XGBoost model evaluated 19 routine laboratory parameters from 1410 patients to differentiate between AML, APL, and ALL.

Multicenter validation exhibited good performance, attaining AUC values of 0.97 for APL, 0.90 for ALL, and 0.89 for AML across all validation cohorts. This validation approach addresses a substantial AI implementation challenge by verifying model performance across patient populations and clinical environments.

AI-driven risk assessment offers benefits compared with conventional ELN methods. An ensemble ML model that incorporated ELN 2022 recommendations enhanced the concordance index from 0.61 to 0.64 and differentiated risk levels, with hazard ratios improved by 18% to 50%.[42] A comprehensive approach that integrated 7 ML algorithms demonstrated enhanced risk stratification compared with ELN criteria alone. AI models identify intricate patterns in AML data that traditional systems overlook, providing enhanced personalized prognostication. Integrating diverse data sources (genomic, imaging, and clinical information) enables AI to enhance disease risk understanding beyond traditional methods, potentially resulting in more informed treatment decisions and improved patient outcomes.

Table 3. Artificial Intelligence (AI) in Genomics and Transcriptomics

AI method	Key features	Main findings	Reference
High-dimensional machine learning with least absolute shrinkage and selection operator regression	Dataset: 12 029 samples from 105 different studies; transcriptomic-based machine learning; data-driven, genome-wide predictors	Greater than 99% accuracy in AML identification; 97.6% sensitivity and 99.5% specificity for M3 AML vs other subtypes; scalable with low marginal cost	Warnat-Herresthal et al,[40] 2020
Cuproptosis-related prognostic model with stacking	Used GSE37642 as training data; GSE12417 and TCGA-LAML for validation; random survival forests algorithm with stacking	AUC values of 0.840, 0.876, 0.892 at 1, 2, and 3 y, respectively (training); AUC values of 0.741, 0.754, and 0.783 (external validation); superior performance vs simple predictor models	Wang et al,[6] 2024
Multi-PCD pathway machine learning biomarker screening	13 PCD pathways, including cuproptosis, ferroptosis, and pyroptosis; 73 ML combinations from 10 algorithms; TCGA-AML, Tyner, GSE37642 cohorts	6-Gene PPCDI signature; high PPCDI associated with worse prognosis; robust prognostic nomograms with clinical variables	Qin et al,[33] 2024

Abbreviations: AML, acute myeloid leukemia; AUC, area under the curve; ML, machine learning; PCD, programmed cell death; PPCDI, pan-programmed cell death index.

AI in Early Detection and Treatment Planning

Molecular profiling enables patient stratification based on cytogenetic markers, with next-generation sequencing identifying variants in *TET2*, *PHF6*, *KIT*, and *NPM1* as prognostic indicators.[43,44]

ML algorithms excel at early risk stratification by integrating diverse data types to forecast outcomes. Promising applications include venetoclax/azacitidine therapy prediction,[45,46] with DL models analyzing bone marrow smear images, achieving an AUROC of 0.84 for separating responders from nonresponders.[45] These systems serve as rapid point-of-diagnosis decision aids while minimizing interobserver variability in blast counting.[47]

For treatment selection, AI models demonstrate increasing sophistication in choosing between intensive chemotherapy and less aggressive alternatives.[46] Multidimensional module optimization generates interpretable drug response predictions from genetic markers, while other systems provide therapy recommendations with high predictive performance.

DL models analyze longitudinal data to predict outcomes and identify resistance, while explainable AI (XAI) is associated with improved clinical interpretation of predictions.[48,49] These AI-driven approaches show particular promise for democratizing AML care in resource-constrained environments where sophisticated molecular diagnostics remain inaccessible, with point-of-care systems achieving diagnostic accuracy that is comparable with expensive conventional methods.[50] While these applications show promise, clinical translation necessitates resolving challenges associated with data standardization, model scalability, and equitable access to these technologies.[51,52]

Challenges and Future Directions

In the evolving AML treatment landscape, AI offers potential alongside considerable challenges. Due to limited extensive research, we enhanced our literature analysis with recent abstracts from prominent hematology conferences.

AI in AML Management: Challenges and Opportunities

AI in AML management presents promise and challenging obstacles. Disease complexity, model transparency, data limits, regulatory concerns, and implementation challenges are obstacles to clinical adoption.[26,53]

The diverse genetic foundation of AML poses serious difficulties for AI researchers. Each patient's genomic profile represents a distinct constellation of changes that evolve as disease progresses. Integrating various data types, from genomic sequences to clinical reports, requires advanced methods to capture significant trends as disease evolves.[19,44]

From Black Box to Transparent Decision Support

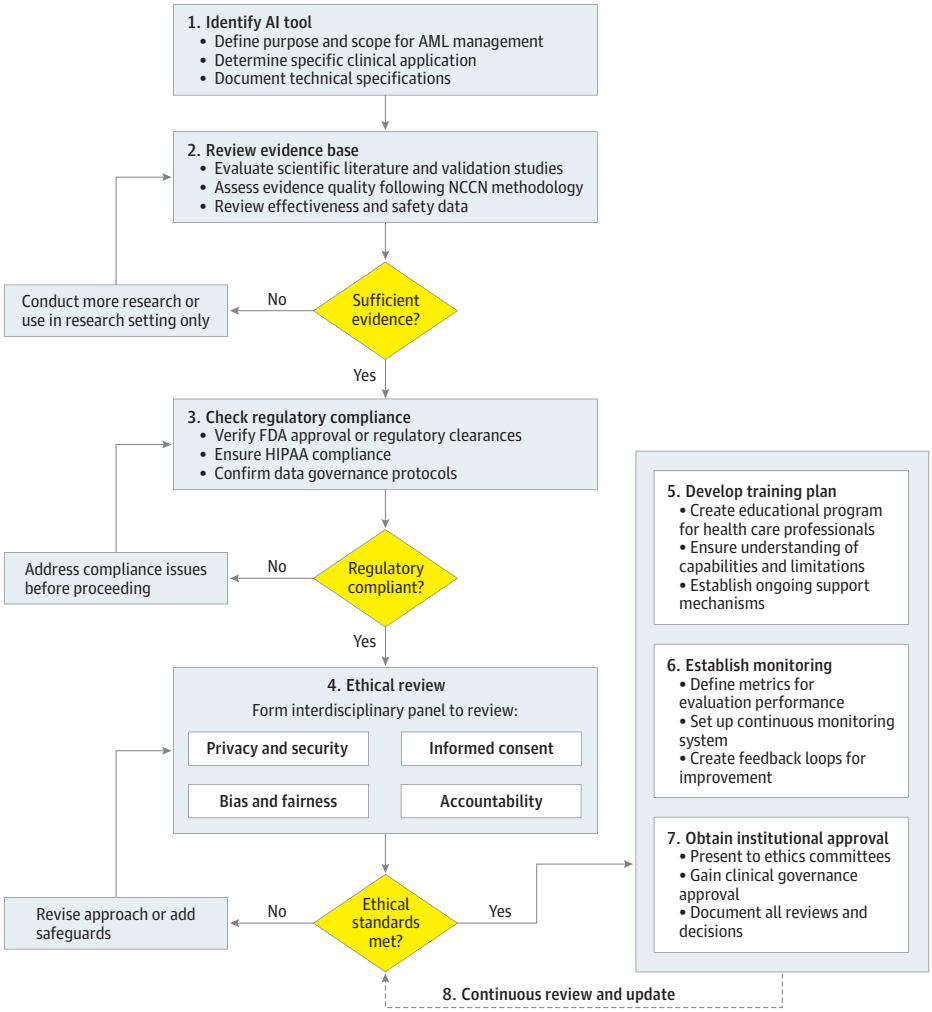
The black box nature of AI systems necessitates transparent, interpretable models for clinical acceptance.[48,54] Risk models integrate XAI tools, such as Shapley Additive Explanations values[26] and local interpretable model-agnostic explanations to display feature contributions to predictions. These methods are vital in bridging the research-clinical application gap because they allow physicians to validate AI reasoning against clinical judgment. This distinction becomes crucial, as expert interpretation of clinical trial data varies substantially, with studies showing that even experienced clinicians sometimes incorrectly interpret trial results,[55] highlighting the value of AI in standardizing evidence synthesis while preserving clinical expertise in decision-making.

A representative case example (eFigure 2 in the Supplement) demonstrates how Shapley Additive Explanations plots can guide venetoclax/azacitidine therapy decisions by visualizing the positive association of *NPM1* variants alongside negative factors like a high blast percentage, enabling clinicians to make targeted therapy adjustments based on patient-specific risk factors.

Ethical Framework and Regulatory Pathways

While AI is able to organize complex multimodal data and identify prognostic patterns invisible to traditional analysis, its role should distinguish between data organization/interpretation vs direct treatment decision-making, for which the latter remains more contentious among clinicians.[56] A study found that 70% of IBM Watson for Oncology's therapeutic options for breast, lung, colon, and rectal cancers were either identical to or acceptable alternatives to the treatment decisions made by oncologists,[57] underscoring the need for clear role delineation. AI applications in AML must adhere to intricate regulatory frameworks that were primarily established for software as a medical device and digital health products, focusing on data security, quality, and algorithm verification.[58]

Figure 1. Artificial Intelligence (AI) Ethics Review Process for Acute Myeloid Leukemia (AML) Management



Structured ethics review process for AI implementation in AML management based on National Comprehensive Cancer Network (NCCN) standards. The sequential workflow included 5 phases: AI tool identification, evidence evaluation, regulatory compliance verification, ethical review, and clinical implementation. Each phase included specific checkpoints and decision points to ensure responsible AI use in clinical practice. FDA indicates US Food and Drug Administration; HIPAA, Health Insurance Portability and Accountability Act.

AI's Role Spectrum in Clinical Care Domains

AI applications in AML management operate across distinct clinical care domains: data organization and interpretation (pattern recognition, risk stratification, variant prediction), clinical decision support (treatment recommendations with physician oversight), and administrative optimization.[59] This challenge is compounded by variation in how experts interpret identical clinical trial data for individual patient scenarios, highlighting the need for frameworks that systematically integrate clinical expertise with AI-derived insights.[60] Rather than replacing oncologists, AI should be positioned as an assistant that systematically integrates clinical expertise with algorithmic insights,[56] which is particularly valuable for organizing dozens of approved treatment approaches now available when rational choice requires quantitative biomarkers and decision algorithms.[61]

Implementation necessitates continuous risk management, transparent data governance, and suitable human oversight.[62] Ethical considerations encompass patient confidentiality, informed consent frameworks that account for AI complexity, data stewardship, and preventing algorithmic bias. The lack of established AI-specific guidelines in hematologic cancers requires collaboration among regulatory agencies, researchers, and clinicians to develop comprehensive ethical standards and governance frameworks[63] (**Figure 1**).

Data Scarcity and Representation Challenges

Limited sample sizes, insufficient demographic representation, and incomplete disease spectrum data constrain AI model generalizability and may generate biases that undermine clinical utility.[64] Technical innovations in FL have emerged as transformative solutions to address these constraints. Federated Healthcare Benchmark, a comprehensive benchmark for evaluating FL in realistic health care scenarios, along with EPFL, a personalized FL framework integrating low-rank adaptation, optimize computational and communication efficiency while maintaining model performance across diverse medical modalities.[37] These approaches balance global knowledge sharing with client-specific adaptation, enabling institutions with limited data or computational resources to benefit from state-of-the-art models without compromising patient privacy.[37,65,66] Advanced security measures, including differential privacy, homomorphic encryption, and secure aggregation, have been integrated into these frameworks, addressing concerns about model inversion and

membership inference attacks while maintaining competitive accuracy (91.2%-97.6%).[66] These democratized approaches to AI access help address health care disparities by extending advanced diagnostic capabilities to resource-constrained environments.

Bridging Research and Clinical Practice

Despite encouraging studies, most AML AI applications remain restricted to research settings[67] due to technical difficulties in integrating with existing hospital infrastructure, resistance from health care clinicians who lack AI tool experience, and a lack of standardized protocols for integrating AI suggestions into practice.[68] The combination of genomic medicine and ML is promising for novel therapeutic targets and combination therapy design,[69,70] but successful implementation requires addressing these practical challenges.

Advances in AI-Enhanced Diagnostics

AI-enhanced imaging and biomarker analysis have transformed leukemia diagnosis and management by facilitating automated blood smear analysis, interpreting bone marrow biopsy specimens, and assessing molecular markers. AI systems detect subtle morphological or molecular patterns that may elude human observation, potentially indicating disease recurrence or progression prior to clinical manifestation.[54] The amalgamation of imaging, molecular diagnostics, and clinical data offers a comprehensive disease status and progression perspective.

Case Studies in XAI

Two case studies illustrated the potential of AI in diagnostic workflows while preserving explainability. University of Münster researchers developed a DL system that predicted genetic features from high-resolution bone marrow smear images.[7] This system uses unsupervised learning to extract cells from intricate bone marrow environments, subsequently applying a neural network analysis to detect morphological anomalies that suggest genetic alterations. The system's XAI mechanisms emphasize particular visual features associated with predictions, enhancing clinical trust.

Similarly, single cell–based explainable multiple instance learning algorithm classifies AML subtypes from routine blood smears.[71] The algorithm was trained on more than 80 000 white blood cell photographs from 129 patients with AML and 60 healthy controls. It achieved perfect discrimination between patients with AML and healthy controls in validation testing and accurately detected APL subtype (mean [SD] F1 score, 0.86 [0.05]). Both systems highlight explainability, allowing clinicians to understand and approve AI-based conclusions, which is a key factor for clinical adoption.

The Horizon: Precision Medicine Through AI

The most exciting AI application in AML extends precision medicine frontiers, leveraging pattern recognition of genomic and molecular data to predict treatment responses and outcomes. This enables increasingly specific treatment regimens for each patient's disease biology and genetics. The convergence of ML and genomic medicine has potential for new therapeutic target discovery and rational combination regimens.[69,71]

Research Gaps and Future Directions

Existing AI use in AML is subject to substantial limitations. Limited sample sizes and low demographic heterogeneity hinder the ability of AI models to generalize across broad patient populations. Our review discovered a lack of published literature regarding AI use in AML, which required inclusion of conference abstracts to capture current information. Although useful for trend identification, this dependence on preliminary data highlights the necessity for full peer-reviewed studies. Most current research is retrospective, without control groups or experimental design.

Overcoming these limitations through standardized data collection procedures, validation research, and interpretable AI models will be instrumental in advancing the field. Despite these limitations, the literature shows progress in AI application in AML management, indicating potential for improving patient outcomes by enhancing diagnosis accuracy, risk stratification, and treatment planning.

AI-Enabled Decision Support Interface for AML Management

Our team proposes a unified AI-supported decision system for AML management that preserves physician autonomy while providing comprehensive patient analysis. The prototype interface demonstrates how AI-supported diagnosis and risk stratification can be integrated into clinical workflows.

The interface organizes data logically into analogous information groups following a diagnostic workflow, overlaying patient demographic characteristics with AI-derived risk analyses, contrasting conventional ELN stratification with improved AI prediction. XAI metrics embedded in the interface are more interpretable; morphological analysis explains individual cellular characteristics that are associated with algorithmic interpretations, and the genetic profile section displays confidence levels for anticipated variants.

The treatment recommendation module lays out primary and alternate therapeutic options, outcome predictions, and their respective confidence intervals to enable well-informed clinical decisions and better patient engagement. The accept and document workflow ensures physician review of all AI-driven recommendations, fulfilling regulatory requirements and promoting trust. This concept illustrates how algorithmic analysis can augment clinical knowledge while preserving established workflows, providing a pragmatic vision for applying AI technologies to AML management (**Figure 2**).

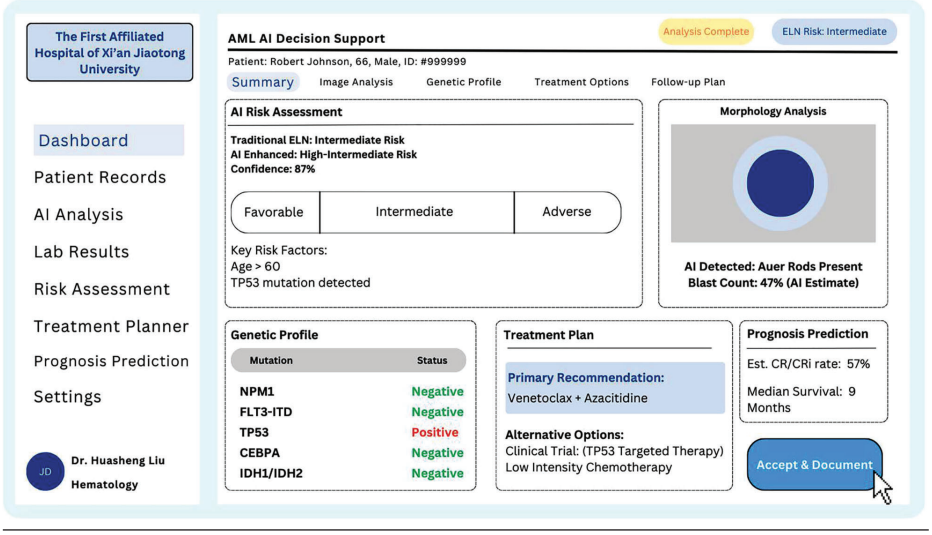
Limitations

AI use in AML faces challenges, including small patient cohorts with limited demographic diversity and a lack of prominent publications, which require reliance on conference abstracts. Most studies are retrospective and did not have control groups or randomized designs, leaving substantial gaps in understanding how AI can detect early AML. Data inconsistency across health care centers and a lack of long-term follow-up data challenge the development of universally applicable AI systems. Despite these limitations, the literature reveals that advancements in AI applications for AML control are encouraging. To advance the field, standardization of data acquisition, validation studies, and development of more interpretable AI models will be required.

Conclusions

AI technologies transform AML management through accurate risk stratification, early detection, and personalized treatment by inte-

Figure 2. Proposed Artificial Intelligence (AI)–Enhanced Clinical Decision Support Interface for Acute Myeloid Leukemia (AML) Management



This conceptual interface demonstrates the integration of AI-assisted diagnosis and risk stratification into the AML clinical workflow. The dashboard displays patient information alongside traditional European LeukemiaNet and AI-enhanced risk assessments, morphological analysis with blast count estimation, genetic profile with variant status, treatment recommendations with alternatives, and survival predictions. The interface maintains physician oversight through an accept and document workflow, ensuring human verification of all AI recommendations while providing transparent explanations of algorithmic interpretations.

grating clinical, morphological, and genetic data. Essential advances in explainable AI address the black box issue, making algorithmic recommendations more transparent and clinically meaningful. Although promising, fulfillment requires addressing substantial data availability and quality gaps, providing definitive regulatory frameworks and ensuring equitable access across diverse health care settings.

AML serves as a model in demonstrating how AI can address complex clinical challenges across cancer types, with applications stratified across clinical care domains: data organization and pattern recognition (in which AI excels), clinical decision support (in which AI assists but does not replace clinical judgment), and treatment selection optimization (in which AI provides quantitative frameworks for increasingly complex therapeutic landscapes).[72] This approach is particularly valuable in low-income and middle-income countries where AI can compensate for limited access to advanced diagnostic infrastructure.[73] The integration of FL and FMs proves particularly valuable in resource-poor environments, where AI-assisted image analysis can compensate for limited access to high-end genetic testing facilities.

Multiple-instance learning approaches using whole-slide images without manual annotations demonstrated high accuracy in predicting key genetic alterations,[2] while integration of holotomography with DL enabled detection of *NPM1* variants with exceptional performance through 3-dimensional cell morphology analysis.[28] As these technologies mature, they may improve AML diagnosis, risk stratification, and treatment planning while respecting patient privacy and regulatory requirements.[9,10,37,66] As the discipline moves forward, standardized data collection, rigorous validation studies, and ongoing collaboration between clinicians, data scientists, and regulatory bodies will be required to deliver these emerging technologies to the bedside and improve outcomes for patients with AML.

ARTICLE INFORMATION

Accepted for Publication: June 20, 2025.

Published Online: November 6, 2025. doi:10.1001/jamaoncol.2025.3601

Author Contributions: Dr Liu had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Concept and design: Ansarian, Fatahichegeni, Xu, Chen, Wang, Liu.
Acquisition, analysis, or interpretation of data: Ansarian, Fatahichegeni, Ren, Liu.
Drafting of the manuscript: Ansarian, Fatahichegeni, Chen, Wang, Liu.
Critical review of the manuscript for important intellectual content: Ansarian, Fatahichegeni, Xu, Ren, Liu.
Statistical analysis: Ansarian, Fatahichegeni, Chen, Wang, Liu.
Obtained funding: Liu.
Administrative, technical, or material support: Xu, Ren, Liu.
Supervision: Liu.

Conflict of Interest Disclosures: None reported.

Funding/Support: This study received funding from the University Joint Project of Shaanxi Provincial Key R&D Program (grant 2020GXLH-Y-006), the Clinical Research Project of the First Affiliated Hospital of Xi'an Jiaotong University (grant HX201826), and the Clinical Research Award of the First Affiliated Hospital of Xi'an Jiaotong University, China (grant XJTU1AF-CRF-2022-033).

Role of the Funder/Sponsor: The funding organizations had no role in the design and conduct of the study; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Additional Contributions: We thank Chen Li, PhD, School of Computer Science and Technology, Xi'an Jiaotong University, for his guidance on artificial intelligence combined with acute myeloid leukemia and critical feedback on the manuscript draft. He was not compensated for his contributions.

REFERENCES

1. Alhajahjeh A, Nazha A. Unlocking the potential of artificial intelligence in acute myeloid leukemia and myelodysplastic syndromes. *Curr Hematol Malig Rep*. 2024;19(1):9-17. doi:10.1007/s11899-023-00716-5

2. Wei BH, Tsai XCH, Sun KJ, et al. Annotation-free deep learning for predicting gene mutations from whole slide images of acute myeloid leukemia. *NPJ Precis Oncol*. 2025;9(1):35. doi:10.1038/s41698-025-00804-0

3. Aparicio-Pérez C, Prados de la Torre E, Sanchez-Garcia J, et al. Evolving risk classifications in AML in a real-life scenario: after changes upon changes, is it more and more adverse? *Cancers (Basel)*. 2023;15(5):1425. doi:10.3390/cancers15051425

4. Al-Khabori MK, Alhuraiji A, Alam AR, et al. Consensus-based guidelines for acute myeloid leukemia management in Gulf Cooperation Council countries: addressing unmet regional needs and the changing AML landscape. *JCO Glob Oncol*. 2024;10:e2300165. doi:10.1200/GO.23.00165

5. Döhner H, DiNardo CD, Appelbaum FR, et al. Genetic risk classification for adults with AML receiving less-intensive therapies: the 2024 ELN recommendations. *Blood*. 2024;144(21):2169-2173. doi:10.1182/blood.2024025409

6. Wang X, Sun H, Dong Y, et al. Development and validation of a cuproptosis-related prognostic model for acute myeloid leukemia patients using machine learning with stacking. *Sci Rep*. 2024;14(1):2802. doi:10.1038/s41598-024-53306-7

7. Kockwelp J, Thiele S, Bartsch J, et al. Deep learning predicts therapy-relevant genetics in acute myeloid leukemia from Pappenheim-stained bone marrow smears. *Blood Adv*. 2024;8(1):70-79. doi:10.1182/bloodadvances.2023011076

8. MoradiAmin M, Yousefpour M, Samadzadehaghdam N, Ghahari L, Ghorbani M, Mafi M. Automatic classification of acute lymphoblastic leukemia cells and lymphocyte subtypes based on a novel convolutional neural network. *Microsc Res Tech*. 2024;87(7):1615-1626. doi:10.1002/jemt.24551

9. Pervez K, Sohail SI, Parwez F, Zia MA. Towards trustworthy AI-driven leukemia diagnosis: a hybrid hierarchical federated learning and explainable AI framework. *Inform Med Unlocked*. 2025:101618. doi:10.1016/j.imu.2025.101618

10. Lhasnaoui C, Ait-Mlouk A, Agouti T, Sadgal M. *Bridging AI and Privacy: Federated Learning for Leukemia Diagnosis*. IEEE; 2024.

11. Chang PW, Newman TB. Receiver operating characteristic (ROC) curves: the basics and beyond. *Hosp Pediatr*. 2024;14(7):e330-e334. doi:10.1542/hpeds.2023-007462

12. Richardson E, Trevizani R, Greenbaum JA, Carter H, Nielsen M, Peters B. The receiver operating characteristic curve accurately assesses imbalanced datasets. *Patterns (N Y)*. 2024;5(6):100994. doi:10.1016/j.patter.2024.100994

13. Collins GS, Dhiman P, Ma J, et al. Evaluation of clinical prediction models (part 1): from development to external validation. *BMJ*. 2024;384:e074819. doi:10.1136/bmj-2023-074819

14. Ban JW, Abel L, Stevens R, Perera R. Research inefficiencies in external validation studies of the Framingham Wilson coronary heart disease risk rule: a systematic review. *PLoS One*. 2024;19(9):e0310321. doi:10.1371/journal.pone.0310321

15. US Department of Health and Human Services. Guidance regarding methods for de-identification of protected health information in accordance with the health insurance portability and accountability act (HIPAA) privacy rule. Accessed May 25, 2025. https://www.hhs.gov/hipaa/for-professionals/special-topics/de-identification/index.html

16. Sounderajah V, Ashrafian H, Rose S, et al. A quality assessment tool for artificial intelligence-centered diagnostic test accuracy studies: QUADAS-AI. *Nat Med*. 2021;27(10):1663-1665. doi:10.1038/s41591-021-01517-0

17. Park S, Park YH, Huh J, Baik SM, Park DJ. Deep learning model for differentiating acute myeloid and lymphoblastic leukemia in peripheral blood cell images via myeloblast and lymphoblast classification. *Digit Health*. 2024;10:20552076241258079. doi:10.1177/20552076241258079

18. Sinha R, Schwede M, Viggiano B, et al. Harnessing artificial intelligence for risk stratification in acute myeloid leukemia (AML): evaluating the utility of longitudinal electronic health record (EHR) data via graph neural networks (oral abstract). *Blood*. 2023;142:960. doi:10.1182/blood-2023-190151

19. Didi I, Alliot JM, Dumas PY, et al. Artificial intelligence-based prediction models for acute myeloid leukemia using real-life data: A DATAML registry study. *Leuk Res*. 2024;136:107437. doi:10.1016/j.leukres.2024.107437

20. Siddiqui NS, Klein A, Godara A, Buchsbaum RJ, Hughes MC. Predicting in-hospital mortality after acute myeloid leukemia therapy: through supervised machine learning algorithms. *JCO Clin Cancer Inform*. 2022;6:e2200044. doi:10.1200/CCI.22.00044

21. Lee H, Han JH, Kim JK, et al. Machine learning predicts 30-day outcome among acute myeloid leukemia patients: a single-center, retrospective, cohort study. *J Clin Med*. 2023;12(18):5940. doi:10.3390/jcm12185940

22. Eckardt JN, Middeke JM, Riechert S, et al. Deep learning detects acute myeloid leukemia and predicts NPM1 mutation status from bone marrow smears. *Leukemia*. 2022;36(1):111-118. doi:10.1038/s41375-021-01408-w

23. Chang MS, Tsai XCH, Chou WC, Tien HF, Hou HA, Chen CY. An ensemble model for acute myeloid leukemia risk stratification recommendations by combining machine learning with clinical guidelines. *medRxiv*. Published online April 2, 2024. doi:10.1101/2024.01.08.24301018

24. Fleming S, Tsai CH, Döhner H, et al. Use of machine learning in 2074 cases of acute myeloid leukemia for genetic risk profiling. *Blood*. 2019;134:1392. doi:10.1182/blood-2019-128243

25. Patkar N, Achrekar AR, Ugale PR, et al. Machine learning identifies gene mutations and variant allele fractions that refine the 2022 European Leukemianet risk stratification for acute myeloid leukemia (oral abstract). *Blood*. 2024;144:852. doi:10.1182/blood-2024-193534

26. Adami D, Riechert S, Schmittmann T, et al. Autonomous region-of-interest detection, cell segmentation and cell classification in bone marrow cytomorphology using deep learning. *Blood*. 2024;144(suppl 1):4975-4975. doi:10.1182/blood-2024-205235

27. Eckardt JN, Schmittmann T, Riechert S, et al. Deep learning identifies acute promyelocytic leukemia in bone marrow smears. *BMC Cancer*. 2022;22(1):201. doi:10.1186/s12885-022-09307-8

28. Kim H, Kim G, Park H, Lee MJ, Park Y, Jang S. Integrating holotomography and deep learning for rapid detection of *NPM1* mutations in AML. *Sci Rep*. 2024;14(1):23780. doi:10.1038/s41598-024-75168-9

29. Matek C, Krappe S, Münzenmayer C, Haferlach T, Marr C. Highly accurate differentiation of bone marrow cell morphologies using deep neural networks on a large image data set. *Blood*. 2021;138(20):1917-1927. doi:10.1182/blood.2020010568

30. Jawahar M, Anbarasi LJ, Narayanan S, Gandomi AH. An attention-based deep learning for acute lymphoblastic leukemia classification. *Sci Rep*. 2024;14(1):17447. doi:10.1038/s41598-024-67826-9

31. Khan S, Sajjad M, Abbas N, Escorcia-Gutierrez J, Gamarra M, Muhammad K. Efficient leukocytes detection and classification in microscopic blood images using convolutional neural network coupled with a dual attention network. *Comput Biol Med*. 2024;174:108146. doi:10.1016/j.compbiomed.2024.108146

32. Eckardt JN, Srivastava I, Wang Z, et al. Synthetic bone marrow images augment real samples in developing acute myeloid leukemia microscopy classification models. *NPJ Digit Med*. 2025;8(1):173. doi:10.1038/s41746-025-01563-9

33. Qin Y, Pu X, Hu D, Yang M. Machine learning-based biomarker screening for acute myeloid leukemia prognosis and therapy from diverse cell-death patterns. *Sci Rep*. 2024;14(1):17874. doi:10.1038/s41598-024-68755-3

34. Mosquera Orgueira A, Perez Encinas MM, Diaz Varela N, et al. Validation of the artificial intelligence prognostic scoring system for myelodysplastic syndromes in chronic myelomonocytic leukaemia: a novel approach for improved risk stratification. *Br J Haematol*. 2024;204(4):1529-1535. doi:10.1111/bjh.19341

35. Eckardt JN, Hahn W, Röllig C, et al. Mimicking clinical trials with synthetic acute myeloid leukemia patients using generative artificial intelligence. *NPJ Digit Med*. 2024;7(1):76. doi:10.1038/s41746-024-01076-x

36. He Y, Huang F, Jiang X, et al. Foundation model for advancing healthcare: challenges, opportunities and future directions. *IEEE Rev Biomed Eng*. 2025;18:172-191.

37. Wang P, Chen Q, Zhang T, Zhang Y, Lu W, Chen Y. FHBench: towards efficient and personalized federated learning for multimodal healthcare. *arXiv*. 2025:250410817.

38. Li X, Peng L, Wang YP, Zhang W. Open challenges and opportunities in federated foundation models towards biomedical healthcare. *BioData Min*. 2025;18(1):2. doi:10.1186/s13040-024-00414-9

39. Li H, Han Z, Sun Y, et al. CGMega: explainable graph neural network framework with attention mechanisms for cancer gene module dissection. *Nat Commun*. 2024;15(1):5997. doi:10.1038/s41467-024-50426-6

40. 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. doi:10.1016/j.isci.2019.100780

41. Alcazer V, Le Meur G, Roccon M, et al. Evaluation of a machine-learning model based on laboratory parameters for the prediction of acute leukaemia subtypes: a multicentre model development and validation study in France. *Lancet Digit Health*. 2024;6(5):e323-e333. doi:10.1016/S2589-7500(24)00044-X

42. Rausch C, Rothenberg-Thurley M, Dufour A, et al. Validation and refinement of the 2022 European LeukemiaNet genetic risk stratification of acute myeloid leukemia. *Leukemia*. 2023;37(6):1234-1244. doi:10.1038/s41375-023-01884-2

43. Li J, Wang Y, Dong C, Luo L. Advancements in leukemia management: bridging diagnosis, prognosis and nanotechnology (review). *Int J Oncol*. 2024;65(6):112. doi:10.3892/ijo.2024.5700

44. Dou X, Dan C, Zhang D, et al. Genomic mutation patterns and prognostic value in de novo and secondary acute myeloid leukemia:

a multicenter study from China. *Int J Cancer*. 2024;155(12):2253-2264. doi:10.1002/ijc.35125

45. Eckardt JN, Srivastava I, Xu Y, et al. Deep learning predicts response to venetoclax and azacitidine in patients with newly diagnosed acute myeloid leukemia solely based on bone marrow smear image data (poster abstract). *Blood*. 2024;144:2917. doi:10.1182/blood-2024-206776

46. Islam N, Reuben JS, Dale J, et al. Predictive models for long term survival of AML patients treated with venetoclax and azacitidine or 7+3 based on post treatment events and responses: retrospective cohort study. *JMIR Cancer*. 2024;10(1):e54740. doi:10.2196/54740

47. Zhang Z, Arabyarmohammadi S, Leo P, et al. Automatic myeloblast segmentation in acute myeloid leukemia images based on adversarial feature learning. *Comput Methods Programs Biomed*. 2024;243:107852. doi:10.1016/j.cmpb.2023.107852

48. Gimeno M, San José-Enériz E, Villar S, et al. Explainable artificial intelligence for precision medicine in acute myeloid leukemia. *Front Immunol*. 2022;13:977358. doi:10.3389/fimmu.2022.977358

49. Metta C, Beretta A, Pellungrini R, Rinzivillo S, Giannotti F. Towards transparent healthcare: advancing local explanation methods in explainable artificial intelligence. *Bioengineering (Basel)*. 2024;11(4):369. doi:10.3390/bioengineering11040369

50. Im H, Pathania D, McFarland PJ, et al. Design and clinical validation of a point-of-care device for the diagnosis of lymphoma via contrast-enhanced microholography and machine learning. *Nat Biomed Eng*. 2018;2(9):666-674. doi:10.1038/s41551-018-0265-3

51. Bernardi S, Vallati M, Gatta R. Artificial intelligence–based management of adult chronic myeloid leukemia: where are we and where are we going? *Cancers (Basel)*. 2024;16(5):848. doi:10.3390/cancers16050848

52. Castro GA, Almeida JM, Machado-Neto JA, Almeida TA. A decision support system to recommend appropriate therapy protocol for AML patients. *Front Artif Intell*. 2024;7:1343447. doi:10.3389/frai.2024.1343447

53. Eckardt JN, Hahn W, Ries RE, et al. Age-stratified game-theory–informed machine learning of molecular alterations unveils prognostic divergence in 3062 pediatric and adult acute myeloid leukemia patients (poster abstract). *Blood*. 2024;144:2211. doi:10.1182/blood-2024-198861

54. Li J, Pavlov S, Stakhov O. Expert systems for analysis of biomedical information in the diagnosis of acute leukemia. *Information Technology and Computer Engineering*. 2024;58(3):84-93. doi:10.31649/1999-9941-2023-58-3-84-93

55. D'Arrigo G, El Hafeez SA, Mezzatesta S, et al. Common mistakes in biostatistics. *Clin Kidney J*. 2024;17(7):sfae197. doi:10.1093/ckj/sfae197

56. Wang L, Chen X, Zhang L, et al. Artificial intelligence in clinical decision support systems for oncology. *Int J Med Sci*. 2023;20(1):79-86. doi:10.7150/ijms.77205

57. Suwanvecho S, Suwanrusme H, Jirakulaporn T, et al. Comparison of an oncology clinical decision-support system's recommendations with actual treatment decisions. *J Am Med Inform Assoc*. 2021;28(4):832-838. doi:10.1093/jamia/ocaa334

58. Reddy S. Global harmonization of artificial intelligence-enabled software as a medical device regulation: addressing challenges and unifying standards. *Mayo Clin Proc Digit Health*. 2024;3(1):100191. doi:10.1016/j.mcpdig.2024.100191

59. Shapiro MA, Stuhlmiller TJ, Wasserman A, et al. AI-augmented clinical decision support in a patient-centric precision oncology registry. *AI in Precision Oncology*. 2024;1(1):58-68. doi:10.1089/aipo.2023.0001

60. Thirard R, Ascione R, Blazeby JM, Rogers CA. Integrating expert opinions with clinical trial data to analyse low-powered subgroup analyses: a bayesian analysis of the VERDICT trial. *BMC Med Res Methodol*. 2020;20(1):300. doi:10.1186/s12874-020-01178-6

61. Kather JN, Abernethy AP. How AI will transform cancer care. *Ann Oncol*. 2024;35(12):1085-1087. doi:10.1016/j.annonc.2024.08.2335

62. Esmaeilzadeh P. Challenges and strategies for wide-scale artificial intelligence (AI) deployment in healthcare practices: a perspective for healthcare organizations. *Artif Intell Med*. 2024;151:102861. doi:10.1016/j.artmed.2024.102861

63. Albahri AS, Duhaim AM, Fadhel MA, et al. A systematic review of trustworthy and explainable artificial intelligence in healthcare: assessment of quality, bias risk, and data fusion. *Inf Fusion*. 2023;96:156-191. doi:10.1016/j.inffus.2023.03.008

64. Al-Obeidat F, Hafez W, Rashid A, et al. Artificial intelligence for the detection of acute myeloid leukemia from microscopic blood images; a systematic review and meta-analysis. *Front Big Data*. 2025;7:1402926. doi:10.3389/fdata.2024.1402926

65. D'Amico S, Dall'Olio L, Rollo C, et al. Multi-modal analysis and federated learning approach for classification and personalized prognostic assessment in myeloid neoplasms. *Blood*. 2022;140(suppl 1):9828-9830. doi:10.1182/blood-2022-166802

66. Teo ZL, Jin L, Li S, et al. Federated machine learning in healthcare: a systematic review on clinical applications and technical architecture. *Cell Rep Med*. 2024;5(2):101419. doi:10.1016/j.xcrm.2024.101419

67. Macheka S, Ng PY, Ginsburg O, Hope A, Sullivan R, Aggarwal A. Prospective evaluation of artificial intelligence (AI) applications for use in cancer pathways following diagnosis: a systematic review. *BMJ Oncol*. 2024;3(1):e000255. doi:10.1136/bmjonc-2023-000255

68. Marco-Ruiz L, Hernández MÁT, Ngo PD, et al. A multinational study on artificial intelligence adoption: Clinical implementers' perspectives. *Int J Med Inform*. 2024;184:105377. doi:10.1016/j.ijmedinf.2024.105377

69. Lin PH, Li HY, Fan SC, et al. A targeted next-generation sequencing in the molecular risk stratification of adult acute myeloid leukemia: implications for clinical practice. *Cancer Med*. 2017;6(2):349-360. doi:10.1002/cam4.969

70. Khanna N, Ramya B, Ramarajan N. Impact of artificial intelligence (AI) in predicting risk and improving patient selection for charitable funding: experience of Indian Cancer Society Cancer Cure Fund (ICS-CCF). *Blood*. 2024;144:7642. doi:10.1182/blood-2024-212171

71. Hehr M, Sadafi A, Matek C, et al. Explainable AI identifies diagnostic cells of genetic AML subtypes. *PLOS Digit Health*. 2023;2(3):e0000187. doi:10.1371/journal.pdig.0000187

72. Lotter W, Hassett MJ, Schultz N, Kehl KL, Van Allen EM, Cerami E. Artificial intelligence in oncology: current landscape, challenges, and future directions. *Cancer Discov*. 2024;14(5):711-726. doi:10.1158/2159-8290.CD-23-1199

73. Lawal AO, Ogunniyi TJ, Oludele OI, et al. Innovative laboratory techniques shaping cancer diagnosis and treatment in developing countries. *Discov Oncol*. 2025;16(1):137. doi:10.1007/s12672-025-01877-w