MACHINE LEARNING IN THE DIAGNOSIS AND PROGNOSTIC ASSESSMENT OF ACUTE MYELOID LEUKEMIA

Tashkent State Medical University

General treatment faculty

Mavlonkhonova Odinakhon

Scientific Supervisor: Tokhtayeva Sitora Akmalovna

Abstract:

Acute Myeloid Leukemia (AML) is a heterogeneous hematological malignancy characterized by the uncontrolled proliferation of immature myeloid cells in the bone marrow and peripheral blood. Early diagnosis and accurate prognostic assessment are essential for selecting appropriate treatment strategies and improving patient survival. Machine learning enables the analysis of complex clinical, laboratory, genomic, and imaging data with high predictive accuracy. This study investigates the application of machine learning models in the diagnosis and prognostic assessment of AML. The findings demonstrate that advanced machine learning algorithms improve diagnostic accuracy, relapse prediction, and individualized risk stratification.

Keywords: Acute Myeloid Leukemia, machine learning, diagnosis, prognosis, artificial intelligence, deep learning, predictive modeling.

Introduction

Acute Myeloid Leukemia is one of the most aggressive hematological malignancies and accounts for the majority of acute leukemias in adults. The disease develops as a result of genetic and epigenetic abnormalities that disrupt the normal differentiation of hematopoietic stem cells, leading to the uncontrolled proliferation of immature myeloid blasts. Progressive accumulation of leukemic cells suppresses normal hematopoiesis and results in anemia, thrombocytopenia, neutropenia, recurrent infections, and severe hemorrhagic complications.

AML is characterized by remarkable biological heterogeneity. Patients with similar clinical manifestations often exhibit different therapeutic responses because of variations in cytogenetic abnormalities and molecular mutations. Therefore, accurate diagnosis and reliable prognostic assessment are fundamental for selecting personalized treatment strategies and improving clinical outcomes.

Recent advances in artificial intelligence have introduced machine learning as an effective computational approach for precision hematology. Unlike conventional statistical methods, machine learning algorithms automatically identify complex relationships among multidimensional clinical and molecular variables. These techniques have demonstrated high performance in blood smear analysis, flow cytometry interpretation, genomic classification, survival prediction, and treatment response evaluation.

Materials and Methods

This study was conducted according to the IMRAD methodology through a comprehensive analytical review of scientific literature concerning machine learning applications in Acute Myeloid Leukemia.

The analyzed datasets included complete blood count parameters, bone marrow morphology, flow cytometry findings, cytogenetic abnormalities, genomic mutation profiles, demographic characteristics, treatment response, relapse status, and overall survival. Machine learning models including Random Forest, Extreme Gradient Boosting, Support Vector Machine, Logistic Regression, Artificial Neural Networks, and Convolutional Neural Networks were comparatively evaluated using accuracy, sensitivity, specificity, precision, F1-score, and area under the ROC curve.

The objective of the study was to determine the effectiveness of machine learning models in improving the diagnosis and prognostic assessment of Acute Myeloid Leukemia.

Results

The analysis demonstrated that machine learning models substantially improve the diagnostic accuracy of Acute Myeloid Leukemia by integrating hematological, clinical, cytogenetic, and molecular information. Compared with conventional statistical approaches, machine learning algorithms more effectively identified hidden relationships among biological variables and produced more reliable diagnostic classifications.

Random Forest achieved excellent performance in distinguishing AML from other hematological disorders. The model accurately classified patients by simultaneously evaluating white blood cell count, hemoglobin concentration, platelet count, blast percentage, and cytogenetic abnormalities. Its ensemble structure reduced overfitting and maintained stable predictive performance across different patient groups.

Extreme Gradient Boosting produced the highest prognostic accuracy. The algorithm successfully predicted overall survival, relapse probability, and treatment response using demographic characteristics, laboratory findings, and genomic mutations. Patients carrying FLT3-ITD and TP53 mutations were consistently assigned to the adverse-risk category, whereas NPM1 and biallelic CEBPA mutations were associated with favorable prognostic outcomes.

Support Vector Machine demonstrated high sensitivity in the classification of AML subtypes using flow cytometry and immunophenotypic markers. The model effectively differentiated morphologically similar leukemia variants, improving diagnostic precision in complex clinical cases.

Convolutional Neural Networks significantly improved automated microscopic diagnosis through digital bone marrow image analysis. The deep learning model accurately detected leukemic blast cells, segmented cellular structures, and minimized observer-dependent variability. Automated image interpretation reduced diagnostic time while maintaining high reproducibility and consistency.

Artificial Neural Networks successfully modeled nonlinear interactions among molecular biomarkers and clinical variables. Integration of genomic alterations with laboratory parameters enhanced individualized prognostic assessment and supported personalized therapeutic decision-making.

Overall, multimodal machine learning models consistently outperformed conventional diagnostic methods by simultaneously analyzing morphological, laboratory, genomic, and clinical data. These findings indicate that machine learning provides a reliable approach for early diagnosis, accurate risk stratification, relapse prediction, and prognostic assessment in patients with Acute Myeloid Leukemia.

Discussion

The findings of this study demonstrate that machine learning has become an effective approach for improving both the diagnosis and prognostic assessment of Acute Myeloid Leukemia. The integration of hematological, cytogenetic, genomic, and clinical data enables algorithms to recognize complex biological patterns that are difficult to detect using conventional diagnostic methods. As a result, diagnostic precision and risk classification are considerably enhanced.

Among the evaluated algorithms, ensemble learning models showed the greatest overall performance. Random Forest provided reliable diagnostic classification by reducing model variance and maintaining high sensitivity, while Extreme Gradient Boosting generated the most accurate prognostic predictions through the analysis of molecular mutations and clinical variables. These results indicate that ensemble methods are highly suitable for clinical decision-support systems in AML.

Deep learning also demonstrated substantial clinical value in digital pathology. Convolutional Neural Networks accurately identified leukemic blast cells from bone marrow microscopic images and reduced observer-dependent variability. Automated image interpretation improves laboratory efficiency and supports standardized morphological evaluation, particularly in high-volume diagnostic centers.

The incorporation of genomic biomarkers further strengthened prognostic assessment. Mutations in FLT3, TP53, NPM1, IDH1/2, and CEBPA significantly influenced survival prediction and relapse estimation. Machine learning models successfully integrated these molecular alterations into individualized risk stratification, providing an important foundation for precision medicine and personalized therapeutic planning.

Despite these advantages, several challenges remain. The availability of large, balanced, and high-quality clinical datasets is essential for developing reliable predictive models. In addition, external validation across different healthcare populations and improved model interpretability are necessary before widespread implementation in routine clinical practice.

Conclusion

Machine learning significantly improves the diagnosis and prognostic assessment of Acute Myeloid Leukemia by analyzing complex clinical and molecular datasets with high accuracy. Ensemble learning and deep learning algorithms provide reliable diagnostic classification, precise risk stratification, and effective prediction of survival and relapse.

Random Forest and XGBoost demonstrated superior performance in diagnostic and prognostic modeling, while Convolutional Neural Networks enhanced automated bone marrow image analysis. The integration of genomic biomarkers with clinical variables enables personalized therapeutic decision-making and supports precision hematology.

The results confirm that machine learning represents a promising clinical tool for the early diagnosis, individualized prognosis, and optimization of treatment strategies in patients with Acute Myeloid Leukemia.

References

1. Döhner H., Wei A.H., Appelbaum F.R. Diagnosis and management of acute myeloid leukemia in adults. Blood, 2022.
2. Short N.J., Rytting M.E., Cortes J.E. Acute Myeloid Leukaemia. The Lancet, 2018.
3. Papaemmanuil E., Gerstung M., Bullinger L. Genomic Classification and Prognosis in Acute Myeloid Leukemia. New England Journal of Medicine, 2016.
4. Arber D.A., Orazi A., Hasserjian R. The 2022 WHO Classification of Myeloid Neoplasms and Acute Leukemia. Leukemia, 2022.
5. Kourou K., Exarchos T.P., Exarchos K.P. Machine Learning Applications in Cancer Prognosis and Prediction. Computational and Structural Biotechnology Journal, 2015.