

REVIEW OPEN ACCESS

A Review on Biomarker-Enhanced Machine Learning for Early Diagnosis and Outcome Prediction in Ovarian Cancer Management

Somayyeh Hormaty^{1,2} | Anwar Nather Seiwan³ | Bushra H. Rasheed⁴ | Hanieh Parvaz⁵ | Ali Gharahzadeh⁶ | Hamid Ghaznavi⁷ 

¹Stem Cell and Tissue Engineering Department, Istinye University, Istanbul, Turkey | ²Department of Chemistry and Polymer Science, Istanbul Technical University, Istanbul, Turkey | ³Department of Biology, College of Science, Basrah University, Basrah, Iraq | ⁴College of Education Ibn Al-Haytham, University of Baghdad, Baghdad, Iraq | ⁵Department of Computer Engineering, Faculty of Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Iran | ⁶Department of Operating Room, Torbat Jam Faculty of Medical Sciences, Torbat Jam, Iran | ⁷Department of Computer Engineering, Social and Biological Network Analysis Laboratory, University of Kurdistan, Sanandaj, Iran

Correspondence: Hamid Ghaznavi (ghaznavi2691@gmail.com)

Received: 25 March 2025 | **Revised:** 29 July 2025 | **Accepted:** 29 August 2025

Funding: The authors received no specific funding for this work.

Keywords: early detection | machine learning | ovarian cancer | precision medicine | prognosis biomarkers

ABSTRACT

Background: Ovarian cancer (OC) remains the most lethal gynecological malignancy, largely due to its late-stage diagnosis and nonspecific early symptoms. Advances in biomarker identification and machine learning offer promising avenues for improving early detection and prognosis. This review evaluates the role of biomarker-driven ML models in enhancing the early detection, risk stratification, and treatment planning of OC.

Methods: We analyzed literature spanning clinical, biomarker, and ML studies, emphasizing key diagnostic and prognostic biomarkers (e.g., CA-125, HE4) and ML techniques (e.g., Random Forest, XGBoost, Neural Networks). The review synthesizes findings from 17 investigations that integrate multi-modal data, including tumor markers, inflammatory, metabolic, and hematologic parameters, to assess ML model performance.

Findings: Biomarker-driven ML models significantly outperform traditional statistical methods, achieving AUC values exceeding 0.90 in diagnosing OC and distinguishing malignant from benign tumors. Ensemble methods (e.g., Random Forest, XGBoost) and deep learning approaches (e.g., RNNs) excel in classification accuracy (up to 99.82%), survival prediction (AUC up to 0.866), and treatment response forecasting. Combining CA-125 and HE4 with additional markers like CRP and NLR enhances specificity and sensitivity. However, limitations such as small sample sizes, lack of external validation, and exclusion of imaging/genomic data hinder clinical adoption.

Conclusion: Biomarker-driven ML represents a transformative approach for OC management, improving diagnostic precision and personalized care. Future research should prioritize multi-center validation, multi-omics integration, and explainable AI to overcome current challenges and enable real-world implementation, potentially reducing OC mortality through earlier detection and optimized treatment.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Cancer Medicine* published by John Wiley & Sons Ltd.