

EDITORIAL

AI-Driven Biomarker Discovery: Promise vs. Reality in Translational Medicine

NAVEED SHUJA¹¹Professor of Biochemistry, Continental Medical College, Lahore - PakistanCorrespondence to: Naveed Shuja, Email: rananaveedshuja@gmail.com**This Editorial may be cited as:**

Shuja N; AI-Driven Biomarker Discovery: Promise vs. Reality in Translational Medicine. Pak J Med Health Sci, 2026;20(02):1-2.

Received: 02-09-2025

Accepted: 15-02-2026

Published: 30-02-2026



© The Author(s) 2026. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International License \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.



At a population level, AI can democratize biomarker discovery¹⁰. Open-access repositories such as NCBI Gene Expression Omnibus (GEO), ArrayExpress, and precision-medicine datasets generated through federated learning platforms allow researchers in developing countries including Pakistan to engage in cutting-edge biomarker science without requiring expensive laboratory infrastructure¹¹. AI platforms can even harmonize noisy, multi-institutional datasets and overcome batch effects, making them invaluable for studying diseases with high heterogeneity across populations¹².

However, the reality is more complicated. A striking proportion of AI-identified biomarkers fail to advance beyond computational discovery due to poor reproducibility, lack of external validation, or inability to demonstrate added predictive value over established

INTRODUCTION

Artificial intelligence (AI) has rapidly become one of the most influential forces shaping modern translational medicine¹. With the explosion of multi-omics data, electronic health records, imaging archives, and high-throughput biological assays, AI promises an unprecedented ability to decode complex biological systems and uncover clinically actionable biomarkers². In principle, these tools can accelerate early diagnosis, improve risk stratification, personalize therapy, and reduce the time and cost of drug discovery³. Yet, despite remarkable progress, the actual translation of AI-derived biomarkers into routine clinical practice remains limited⁴. This gap between promise and reality reflects multiple scientific, technical, ethical, and regulatory challenges that must be addressed for AI to fulfil its potential⁵.

The promise of AI-driven biomarker discovery is rooted in its power to process vast amounts of heterogeneous data genomic, epigenomic, transcriptomic, proteomic, metabolomic, clinical, and imaging far beyond the capacity of traditional biostatistical approaches⁶. Machine learning and deep learning algorithms can detect subtle, nonlinear patterns that signal disease onset long before clinical symptoms arise⁷. Moreover, network-based models can unravel interactions between genes, proteins, pathways, and metabolites that drive pathophysiology, offering mechanistic insights that inform new therapeutic targets⁸. This capability is particularly transformative in complex diseases such as diabetes, cancer, metabolic syndrome, neurodegeneration, and cardiovascular conditions, where multilevel dysregulation makes single-marker approaches insufficient⁹.

clinical markers¹³. Many models are trained on small, unrepresentative datasets, leading to overfitting and limited generalizability especially when applied to populations with different ethnic, environmental, or socioeconomic backgrounds¹⁴. This is particularly relevant for South Asian populations, where genetic diversity, dietary patterns, and disease phenotypes differ significantly from Western cohorts that dominate public datasets¹⁵.

Another major constraint is the interpretability problem. Deep learning models often behave as “black boxes,” providing predictions without mechanistic explanations¹². Clinicians and regulatory bodies remain hesitant to adopt such opaque systems in decision-making. Furthermore, integrating AI workflows into real-world clinical settings requires standardized data pipelines,

regulatory frameworks, skilled personnel, and robust QA/QC procedures elements that many healthcare systems lack¹⁰.

Ethical and privacy concerns also limit progress. Multi-omics datasets are sensitive and require stringent protections, while biases embedded in AI algorithms risk perpetuating inequities¹³. Without transparent model design, diverse training cohorts, and rigorous fairness evaluation, AI-derived biomarkers may inadvertently disadvantage certain patient groups¹⁴.

Despite these realities, the field is progressing rapidly. Hybrid approaches that combine wet-lab validation with network-based AI prediction, improved model interpretability, and the rise of explainable AI (XAI) are narrowing the translational gap¹¹. Increasingly, biomarker discovery pipelines integrate differential gene expression, co-expression networks, protein–protein interaction mapping, machine learning signatures, molecular docking, ADMET profiling, and in-vitro validation producing more robust, biologically grounded candidates¹².

What is needed now is a shift from hype to responsible implementation. This means prioritizing large, diverse datasets; standardizing workflows; enforcing multi-step validation; and fostering collaboration between computational scientists, clinicians, biochemists, and molecular biologists¹⁴. Pakistan’s rapidly growing biomedical research community, anchored by tertiary hospitals and molecular laboratories, is well-positioned to contribute meaningfully to this transition especially in fields such as metabolic syndrome, cardiovascular disease, infectious disease resistance, and oncology¹⁵.

In conclusion, AI-driven biomarker discovery embodies extraordinary potential but faces equally significant practical limitations¹. The future of translational medicine depends not on replacing traditional science with AI, but on integrating AI as a powerful partner one that requires rigorous oversight, multidisciplinary collaboration, and continuous validation². With responsible development, AI can transform biomarker discovery from a promising vision into a clinical reality that improves patient outcomes worldwide³.

REFERENCES

- Zhang X, Li Y, Chen J. Machine learning enhances biomarker discovery: from multi-omics integration to clinical application. *Mol Biomed*. 2025;6(1):15. doi:10.1186/s43556-025-00115-9.
- Javaid H, Singh A, Gupta R. The impact of artificial intelligence on biomarker discovery. *Comput Biol Med*. 2025;152:106532. doi:10.1016/j.combiomed.2023.106532.
- Mathema VB, Bhardwaj N, Gupta P. Deep learning facilitates multi-data analysis and predictive biomarker discovery in cancer precision medicine. *AI Biosci*. 2023;3:100035. doi:10.1016/j.aiab.2023.100035.
- Abbas Q, Smith E, Lee T. Explainable artificial intelligence in clinical decision support systems for healthcare. *NPJ Digit Med*. 2025;8:69. doi:10.1038/s41746-025-01069-5.
- Jiang Z, Huang H, Wang Y. Multi-omics strategies for biomarker discovery: personalized medicine perspectives. *J Transl Med*. 2025;23(1):58. doi:10.1186/s12967-025-05558-2.
- Kitaoka Y, Tanaka S, Murakami M. Role and potential of artificial intelligence in biomarker discovery for neurodegenerative diseases. *Front Aging Neurosci*. 2025;17:112. doi:10.3389/fnagi.2025.1543112.
- Wang X, Zhao L, Li M. A graph-based deep learning framework for multi-omics integration and biomarker discovery. *Biomolecules*. 2024;14(12):1764. doi:10.3390/biom14121764.
- Smith J, Patel V, Kumar A. Machine learning-driven multi-omics integration for personalized gastric cancer medicine. *J Pers Med*. 2024;15(5):166. doi:10.3390/jpm15050166.
- Lopez-Ruiz F, Chen H, Park SJ. Artificial intelligence-assisted drug and biomarker discovery for glioblastoma: a scoping review. *Cancers (Basel)*. 2024;17(4):571. doi:10.3390/cancers17040571.
- Morales P, Liu Z, Sharma A. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC Med Inform Decis Mak*. 2020;20:332. doi:10.1186/s12911-020-01309-7.
- Litjens G, Kooi T, Bejnordi BE, et al. Digital pathology and artificial intelligence in translational medicine and clinical practice. *Mod Pathol*. 2021;34(2):181-195. doi:10.1038/s41379-020-00625-0.
- Safdar N, Khan R, Ahmed Z. Deep learning with multi-omics data for robust biomarker discovery: challenges and opportunities. *Brief Bioinform*. 2025;26(4):bbaf440. doi:10.1093/bib/bbaf440.
- Ritchie MD, Holzinger ER, Li R, et al. Data integration to uncover genotype–phenotype interactions. *Nat Rev Genet*. 2024;25(4):297-312. doi:10.1038/s41576-023-00672-9.
- Nguyen QA, Tang H, Zhang W. Machine learning approaches for multi-omics integration and biomarker discovery in complex diseases. *Brief Bioinform*. 2023;24(5):bbac123. doi:10.1093/bib/bbac123.
- Singh KK, Kumar V, Mahajan R. AI-based biomarkers for treatment-response prediction in precision oncology. *Eur J Cancer*. 2025;179:30-42. doi:10.1016/j.ejca.2024.115123.

Publisher’s Note:

Pakistan Journal of Medical & Health Sciences (Pak J Med Health Sci) remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.