Romanian Journal of Oral Rehabilitation
Vol. 17, No.1 January-March 2025

THE IMPACT OF ARTIFICIAL INTELLIGENCE IN MEDICAL BIOSTATISTICAL STUDIES: ARE THE STATISTICIANS NO LONGER NECESSARY? A LITERATURE REVIEW

Cristina-Gena Dascălu[1], Norin Forna[1]*, Andrei Georgescu[2], Claudiu Topoliceanu[2]

1 Faculty of Medicine, University of Medicine and Pharmacy Grigore T.Popa Iasi
2 Faculty of Dental Medicine, University of Medicine and Pharmacy Grigore T.Popa Iasi

Corresponding author: Lecturer Norin Forna; e-mail: norin.forna@gmail.com

ABSTRACT
Aim. The aim of this review is to explore and critically assess the current applications of artificial intelligence (AI) in medical biostatistical research, encompassing both theoretical perspectives and real-world applications. AI techniques, such as deep learning, support vector machines, decision trees or clustering, expands upon traditional biostatistics to enable systematic processing of complex and high-dimensional unstructured biomedical datasets. AI has revolutionized the research methodologies employed in traditional medical biostatistics, providing new tools for data analysis, predictive modeling, and integration of heterogeneous data sources. AI promises to add value in many medical domains - disease risk prediction, clinical trial analysis, omics integration, and precision medicine, to name a few. Combined with traditional statistical techniques, AI provides more complex and comprehensive insights into the structure of medical information, having the potential to respond optimally to the expanding medical needs for swift, precise, and personalized treatments.

Key words: medical biostatistics, artificial intelligence, research

INTRODUCTION

Background on Biostatistical Research in Medicine

Biostatistical research, a backbone of clinical research, provides quantitative tools for studies design, data analysis and results interpretation in clinical and public health settings. Developed from the principles of probability theory, biostatistics allows to researchers to obtain valid conclusions about complex biological or medical data – that is why nowadays biostatistics is included in most postgraduate medical curricula due to the need to foster critical thinking in graduate healthcare education for evidence-based practice [1-3]. The need to develop competences in biostatistics and research methodology during medical specialty training is recognized as biostatistics represents a critical methodological skill for medical researchers [4-7].

Biostatistics is essential in clinical research for multiple purposes: the design of randomized controlled trials (RCTs), cohort and case-control studies, the setting up of working hypotheses, the samples' size calculation and optimal selection, as well as the assessment of the appropriate calculation procedures that accommodate optimally for diversity and confounding variables. In biomedical studies, well-conducted statistical assessments are critical to guarantee that the study findings and conclusions are appropriate and trustworthy. Unfortunately, medical literature often reveals a certain heterogeneity regarding the studies design and the statistical methods selected to reach the proposed objectives; that is why a major task in research is to promote the best practices in statistics, in order to elevate the results consistency and quality, to streamline the methodology and to enhance the findings reporting [8].

Artificial Intelligence emergence in biomedical sciences

AI in medicine is represented by virtual and physical approaches. The virtual approach encompasses computerized methods based on

DOI : 10.62610/RJOR.2025.1.17.82

deep learning techniques, aiming to improve health data management and control. The physical approach refers to specific devices, such as robots designed to assist the elderly patient or the resident surgeon, as well as targeted nanorobots [9].

Artificial intelligence in healthcare is mostly applied in clinical trials, preliminary diagnosis, automated image diagnosis and robot-guided surgery. The typical end-users of AI in biomedical sciences are pharmaceutical & biotechnology companies, hospital & diagnostic centers and academic & research laboratories. NVIDIA and Canon Medical Equipment are the most active companies in AI-based healthcare domain. NVIDIA created Clara Holoscan MGXTM in 2022 as a platform for the medical device industry to develop and deploy real-time AI applications at the edge. This platform supplies medical-grade reference architecture and long-term software support that accelerates the creation of medical devices. Canon Medical Systems Corporation introduced Altivity, a new brand dedicated to the development of innovative technologies powered by machine and deep-learning, redefining the role of AI in healthcare. Nowadays the most important players in the AI-based healthcare market are Recursion Pharmaceuticals, Medtronic, CloudmedX, Oncora Medical, and Google [10].

The combination of electronic health records, high-throughput technologies (e.g., genomics, proteomics), and large-scale epidemiological databases has enormously increased the volume and complexity of biomedical data. Thus, the availability of high-dimensional or unstructured data has led to an explosion of data that challenges traditional biostatistical approaches, relying on assumptions not verified in the case of high dimensional or unstructured data. In this context, there is an increasing demand for incorporating artificial intelligence (AI) into biostatistical research. AI can increase speed, flexibility and accuracy in processing large databases, providing robust techniques for pattern recognition, predictive modeling and data integration. Integration of AI with classical statistical methods enhances the analytics capabilities in biomedical research. The challenges to address the interface between medical statistics and AI concern population inference vs. prediction, generalizability, reproducibility and evidence interpretation, as well as the stability and statistical guarantees [9-11].

AIM OF REVIEW

The aim of this review is to explore and critically assess the current applications of artificial intelligence (AI) in medical biostatistical research, encompassing both theoretical perspectives and real-world applications. We proposed to achieve the following objectives in our analysis:

- to explore the opportunities for integration of AI tools in study design, data analysis, and predictive modeling;
- to give relevant examples of AI algorithms used in biostatistical workflows;
- to evaluate AI practical applications in clinical trials, population health studies, and multi-omics research;
- to expose the strengths, limitations, and ethical issues of AI-assisted medical biostatistics.

MATERIALS AND METHOD

The present review was conducted through a comprehensive examination of specialized literature aiming to identify the most relevant and recent studies concerning AI-based applications in medical biostatistics and their impact on research quality and clinical utility.

DOI : 10.62610/RJOR.2025.1.17.82

Data Sources and Search Strategy

A systematic literature search was performed in three major scientific databases: PubMed, Scopus, and Web of Science. The search included articles published between 2014 and 2024, reflecting the rapid evolution and increasing implementation of AI technologies in medical research. The following keywords directed our research: "artificial intelligence", "machine learning in medicine", "AI in biostatistics", "predictive analytics", "medical data modeling", "clinical decision support", "neural networks in healthcare", "AI in public health", "big data in medicine", and "AI in epidemiology".

Eligibility – Inclusion and Exclusion Criteria

Studies were included if they:

- investigated the application of artificial intelligence in medical biostatistical analysis, such as predictive modeling, data classification, regression analysis, or outcome prediction;
- evaluated the accuracy, performance, or added value of AI compared to traditional statistical approaches;
- were published in peer-reviewed journals, indexed in international scientific databases, and were based on a clearly described methodology;
- were focused on human medical research, including clinical trials, epidemiological studies, or public health applications.

Studies were excluded if they:

- were case reports, editorials, conference abstracts, or reviews lacking primary data or generalizable results;
- did not clearly define their AI methodology or lacked description of biostatistical context;
- focused solely on technical algorithm development without medical or biostatistical relevance.

Data Selection and Analysis Methodology

Initial screening of titles and abstracts was performed to exclude studies that did not meet the inclusion criteria. The remaining articles underwent full-text review to ensure relevance and methodological rigor. Data were extracted and synthesized to assess the role of AI in various biostatistical processes.

RESULTS

AI PRACTICAL APPLICATIONS IN MEDICAL BIOSTATISTICS

AI-based predictive modeling of disease risk and prognostic:

Artificial intelligence (AI), and particularly machine learning (ML), are being applied across the medical domain to diagnose diseases and predict patient outcomes. Using large datasets, such as electronic health records, genetic profiles and imaging, AI models can identify patterns and risk factors that may not be captured by traditional approaches, specifically in cancer diagnosis [12-14]. Such models facilitate disease identification at an early stage, tailored treatment strategies, and efficient resource management. Also, AI-based models can predict disease's progression, drugs' efficacy, and prognosis [15, 16]. The main challenges to overcome include data quality, model interpretability, and ethical implications, as well as safety and efficacy for routine clinical usage.

AI-based data analysis in clinical trials and observational studies:

Current multilevel and integrated data sets allow new insights into medical research, by opening new perspectives and analysis directions. However, despite the large amount of available data, only a small fraction can be curated, integrated, understood, and properly analyzed. Considering AI abilities to learn from data and emulate human thought processes, AI-based data analysis applications

DOI : 10.62610/RJOR.2025.1.17.82

magnify learning and decision support system, having the potential to revolutionize the future of health care [17]. Cognitive programs influence both medical practice and biomedical research by using Natural Language Processing techniques, which allow us to identify quickly similar, but disparate electronic medical records or to check the scientific literature for relevant references. AI-based data analysis applications may trace out the optimal care of chronic disease, recommend precision therapies for complex diseases, flag medical errors, and enhance subjects' recruitment for clinical trials [18].

AI-based pattern recognition in omics and imaging data:

Deep learning is changing the face of omics (genomics, proteomics, metabolomics) and medical imaging pattern recognition, since AI-based algorithms can discover complex genomic patterns, biomarkers, and disease signatures through high-dimensional data analysis that sustains early diagnosis and precision medicine [19]. In omics, AI was used to detect gene-disease relations or to predict protein structures and pathways [20]. In medical imaging, AI enhances the identification of abnormalities in radiology or pathology, sometimes equaling or exceeding the specialist's efficiency [21,22]. Combination of omics and imaging with artificial intelligence provides a strong, comprehensive view for understanding the mechanisms of disease. The possible difficulties are related to data integration and models standardization, transparency and generalizability through various populations.

AI-based integration of multi-source biomedical data:

AI facilitates the optimal integration of various sources of biomedical data (clinical records, omics - genomics, proteomics, imaging, wearable device data, etc.) to provide a holistic depiction of health and disease. It harnesses the power of machine learning and deep learning to navigate the intricacy and diversity of such datasets, identify latent associations, and improving disease risk prediction, diagnosis, and individualized therapeutic approaches. Combining different multi-source biomedical data improves the quality of clinical decision-making and helps to create a holistic patient profile [23, 24]. Despite this potential, challenges remain; the most significant of them are represented by data standardization, missing data handling, interoperability as well as privacy and interpretability of AI models integrated into the clinic.

AI ALGORITHMS IN MEDICAL BIOSTATISTICAL RESEARCH

AI algorithms provide robust and advanced techniques to maximize data analysis, recognize complex data patterns, and improve predictive analytics and modeling [25]. While biostatistics usually follows a hypothesis-driven approach, AI and machine learning (ML) enables data-driven exploration across large and complex datasets. While traditional biostatistics usually requires complete databases and linear relationships between variables, neural networks provide optimization algorithms for variable selection, being familiar with missing data and nonlinear relationships between variables. AI algorithms allow prediction of risk for disease, survival analysis, modeling of treatment responses, and real-time monitoring. Common AI tools used in biostats workflows are deep learning, random forests as well as support vector machines [26].

Decision Trees:

Decision tree algorithms are popular in medical biostatistics because they are simple and easy to interpret, as well as effective for complex data. Such AI models work by

DOI : 10.62610/RJOR.2025.1.17.82

recursively segmenting the data based on the input features, forming a tree structure that ultimately reaches a set of predictions [27, 28]. Decision trees are used in medical research for identifying significant risk factors, risk stratification, disease outcome prediction, and clinical decision-making. Random Forests and Gradient Boosted Trees are variants that boost performances through overfitting reduction and increased predictive accuracy. For this class of prediction models (exhibiting so-called "bagging" behavior) correlations between predictions made by each model can reduce the effective number of predictions, leading to an increased stability [29, 30].

Support Vector Machines (SVM):

Support Vector Machines (SVM) is an AI technique widely applied in medical biostatistics for classification and regression problems. SVM is an emerging approach of machine learning that can robustly classify high-dimensional big data into a small number of data points called support vectors, allowing well-defined subgroup differentiation in a minimal time. They are resilient to overfitting, particularly in situations with relatively small sample sizes and a large number of variables. SVMs are able to represent complex, nonlinear relationships, making them very flexible for different medical applications, including novel drug discovery [31, 32], gastric pathologies [33], or neurological complications [34]. SVM have important clinical and public applications for high-throughput data analysis and contextualization, a relevant example in this regard being the diagnostic and treatment timetables of most aggressive cancers [35]. However, their “black-box” nature and parameter tuning sensitivity may lead to low interpretability, which can be an undesired feature in clinical research.

Clustering/PCA algorithms:

Clustering and PCA are unsupervised AI techniques that are frequently applied in medical biostatistics to analyze and visualize large and complex biomedical data. Combination of clustering and PCA techniques enables and facilitates data-driven discovery, pattern recognition, and hypothesis generation in medical research. Clustering algorithms (e.g., k-means, hierarchical clustering) categorize similar data according to features patterns, thus enabling researchers to discover hidden subgroups within samples of patients, disease phenotypes, or molecular profiles. It aids in patients’ stratification for specific therapies or in comprehending disease heterogeneity [36-39]. PCA (Principal Component Analysis) is a method of dimensionality reduction that converts a dataset with correlated features into fewer uncorrelated features (principal components) that preserve data variance. It makes analysis easier, filters noise, and facilitates eventually the complex data visualization (e.g. gene expression or imaging data) [40, 41].

Deep Learning Approaches in Medical Biostatistics:

Machine learning is a subdomain of artificial intelligence designed to identify complex patterns that can be used to predict or classify similarly new unseen data. Machine learning techniques can analyze multi-modal data from precision medicine; they provide broad capability for large scale genomic analysis, leading toward increased understanding of human health and disease [42]. Deep learning approaches, including Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Autoencoders, are also increasingly applied in medical biostatistics to analyze complex, high-dimensional data and uncover intricate patterns that traditional methods may miss.

DOI : 10.62610/RJOR.2025.1.17.82

CNNs are particularly effective for analyzing medical imaging data (e.g., MRI, CT, histopathology), from where they automatically extract spatial features to support diagnosis, classification, and segmentation tasks with high accuracy, especially in oncological pathology [43, 44].

RNNs are designed for sequential data, being widely used in analyzing time-series medical records, biosignals (e.g., ECG, EEG), and longitudinal patient data to predict outcomes, monitor disease progression, or model treatment effects over time [45, 46].

Autoencoders are unsupervised models used for data compression, denoising, and anomaly detection. In biostatistics, they help to reduce dimensionality in large omics datasets or to identify subtle data patterns for disease classification and patient stratification. A research group proposed recently a novel framework, MultiSC, for the analysis of MULTIomic Single-Cell data, by incorporation a single-cell hierarchical constraint autoencoder for merging the multi-omics data while performing clustering and matrix factorization based model (scMF) to predict transcription factors regulated target genes [47].

CHALLENGES AND LIMITATIONS

Challenges at the intersection of medical statistics with AI include population inference vs. prediction, data quality, model generalizability, reproducibility and interpretability of evidence, stability, as well as statistical guarantees [48-49]. Data quality and quantity are crucial for AI models. Data in the healthcare sector is very fragmented due to various reasons such as missing values, coding under different schemes, and storage in different infrastructures. These may hinder the generalizability and applicability of AI-driven biostatistical analyses. Many AI models are limited by data overfitting, meaning that good results are obtained on the training data while poor results are obtained on new test data. Thus, in medical biostatistics, predictions lose validity, limiting their utilization in the clinical space. More advanced AI models (deep learning networks) tend to be treated as "black boxes", which can lead to a lack of interpretability. In healthcare, these models do not allow explanation of the clinical decision-making, leading to lack of transparency as well as ethical and practical problems. Combination between AI-based and classic statistical approaches is challenging because the two systems arise from different methodologic traditions. They must be integrated properly through model selection, validation and interpretation to not lead to biased or misleading results. Other challenges related to AI-based application in healthcare are as follows [48-49].

- Biases: If training data reflect existing biases (e.g., certain populations are underrepresented), AI models could reinforce health disparities.
- Clinical adoption and trust: Clinicians can be reluctant in AI tools adoption due to concerns about reliability, usability or misalignment with clinical workflows; large adoption can be reinforced by validation, clear communication of limitations as well as demonstration of clinical utility.
- Computational and infrastructure demands: Need for cloud solutions, scalable computer power, high-performance data warehouses and databases can be challenging for medical institutions with limited technical capacity or funding.

Ethical, Legal, and Privacy Considerations

The implementation of AI algorithms in healthcare domain requires update of ethics issues and regulations due to concerns surrounding AI decision-making, data ownership and patient privacy. These issues

DOI : 10.62610/RJOR.2025.1.17.82

require proactive dialogues among regulators, healthcare providers, technologists and patients to ensure the considerations of ethical guidelines and regulations in patients' health data collection, processing and analysis healthcare [50]. Privacy and data security are critical issues that require the implementation of strong encryption and anonymization methods to protect patient information, since data collection practices based on decentralized data sharing can harm the patient privacy. Algorithmic bias is also a considerable challenge requiring constant monitoring for fairness [51].

To ensure accountability of manufacturers, healthcare institutions and professionals, clear responsibility frameworks for these stakeholders are required. The ethical guidelines, regularly updated and accessible to all stakeholders, can direct decision-making amidst this dynamic landscape.

While AI-based healthcare applications challenge the current codes of ethics and regulatory frameworks, most of the AI-specific ethics frameworks seem to be little used and most are not used as stand-alone ethics frameworks. Operationalizing them in healthcare and biomedical statistics is a complex operation involving multiple levels (ethics principles, design, technology, organizational, regulatory) [52]. A research group highlighted the need for a five level framework of ethical impacts (individual-relational, organizational, societal, global and historical) for the medical AI-systems ethics. A global approach of ethical issues must insert the local ethical regulations into a broader framework where ethical and legal considerations are taken into account to make sure AI-systems serve the individual needs of everyone, everywhere in the world [53]. Also, a more explicit ethical evaluation with benchmarks must be integrated into critical appraisal frameworks for each AI-based healthcare application [54].

FUTURE PERSPECTIVES

New AI technologies in medical biostatistics

New AI technologies in medicine and medical biostatistics include federated learning, generative AI, and explainable AI (XAI). Such technologies allow advanced analysis of distributed data without sacrificing privacy, facilitating more explainable prediction and decision-making models, as well as longitudinal biostatistical modeling. They address to specific issues such as data heterogeneity, high dimensionality and model interpretability for advanced biostatistical analysis [55].

Federated learning is based on machine learning algorithms and data collection at a centralized storage node to allow fast and accurate diagnosis and treatment. It was investigated in a healthcare surveillance system (Internet of Medical Things- IoMT) based on blockchain powered by federated learning [56].

Models of generative AI such as GANs, VAEs, etc., are used to learn the distribution of complex datasets and to generate synthetic biomedical data that resembles authentic samples [57]. Generative AI can generate realistic patient records and/or medical images which can help to prevent certain challenges such as limited sample sizes and imbalance of the clinical datasets. Generative AI also enables simulation studies, privacy-preserving data sharing, and hypothesis testing through the generation of controlled synthetic cohorts. In biostatistical modeling, these tools help make models more robust, enhance generalizability of predictive models, and improve information sharing across models with richer, multi-modal data [58].

Explainable AI (XAI) aims to provide humans with transparent, interpretable, and understandable AI models, one of the most

DOI : 10.62610/RJOR.2025.1.17.82

important requirements in medical biostatistics and clinical decision formulation. In biostatistics, XAI approaches such as SHAP or LIME explain how individual input variables impact model outcomes and correlate their effects, allowing researchers and clinicians to interpret predictions. Although there has been significant advancement in these areas, the challenge of balancing performance and interpretability in machine learning models, particularly in the context of high-dimensional or multi-modal biomedical data, remains an open question [59, 60].

AI-based personalized and precision medicine

The ability to capture the genetic makeup (mutations that drive individual tumors) has led to the concept of precision medicine. The association between artificial intelligence and precision medicine holds promising solutions to revolutionize health care. AI-based personalized medicine aims to account for patient-to-patient and between-population differences in disease susceptibilities and responses to health interventions such as drugs, nutrition, vaccines, and radiation [61]. Methods of precision medicine find phenotypes of patients who have more-uncommon responses to treatment or needs from health services. AI applies complex computation and deduction to provide deeper insight, facilitate reasoning and learning in the system, and enhance clinician decision making through advanced intelligence. Precision medicine matches each individual to the best treatment in a way that is tailored to his or her genetic uniqueness [62]. Translational research evaluating this convergence of nongenomic and genomic determinants, integrated with data on patient symptoms, clinical history and lifestyles will enable clinicians to address the biggest challenges facing precision medicine, through personalized diagnosis and prognosis [63].

An example of AI-based precision medicine are physiological genomic readouts in disease-related tissues providing an effective surrogate for gauging the effect of genetic and environmental factors and their interactions that shape disease development [64]. More than traditional statistical methods, AI biostatistical models can stratify patient risk, optimize treatment plans, and predict therapy response. With the scale-up of precision medicine initiatives, the role of AI will also be to improve patient phenotyping and provide real-time clinical decision support on an individual level.

Collaboration between data scientists, statisticians, and clinicians

Interdisciplinary collaboration between data scientists, statisticians, and clinicians is pivotal in the effective deployment of AI in biostatistics research. The clinicians usually do not have the biostatistical competencies to analyze large databases, while the data scientists and statisticians find difficult to deal with the issues regarding patient privacy and consent. The only solution to overcome such problems is therefore the multidisciplinary collaboration among clinicians, data scientists, machine-learning experts, biostatisticians and epidemiologists [65]. Suboptimal collaboration between clinicians and biostatisticians frequently leads to low quality research. Statisticians offer expertise on study design, curve fitting and inference; data scientists provide advanced algorithmic and computational skill; clinicians guarantee interpretability and clinical relevance. To bridge these domains, there is a need to establish a common language, to identify shared goals, and to co-design research workflows. Such collaboration is critical for translating AI models into practice, preserving methodological rigor, and aligning

DOI : 10.62610/RJOR.2025.1.17.82

technological innovation with patient-centered outcomes [66].

CONCLUSIONS

- AI techniques, such as deep learning, support vector machines, decision trees or clustering, expands upon traditional biostatistics to enable systematic processing of complex and high-dimensional unstructured biomedical datasets.
- AI has revolutionized the research methodologies employed in traditional medical biostatistics, providing new tools for data analysis, predictive modeling, and integration of heterogeneous data sources.
- AI promises to add value in many medical domains: disease risk prediction, clinical trial analysis, omics integration, and precision medicine, to name a few.
- Combined with traditional statistical techniques, AI provides more complex and comprehensive insights into the structure of medical information, having the potential to respond optimally to the expanding medical needs for swift, precise, and personalized treatments.

REFERENCES

1. Arias A, Peters OA, Broyles IL. New curricular design in biostatistics to prepare residents for an evidence-based practice and lifelong learning education: a pilot approach. *Int Endod J.* 2017 Oct;50(10):999-1010.
2. Ille TM, Muthukrishnan R, Ille ME. Integrating medical biostatistics into postgraduate medical curriculum. *Natl Med J India.* 2017 Jul-Aug;30(4):212-214.
3. Batmanabane G, Maiti R, Krishna MBN. Research Methodology and Biostatistics in the Postgraduate Curriculum: The AIIMS-Bhubaneswar Experience. *J Assoc Physicians India.* 2024 Nov;72(11):98-100.
4. Oster RA, Enders FT. The Importance of Statistical Competencies for Medical Research Learners. *J Stat Educ*. 2018;26(2):137-142.
5. Turner SL, Shaw T. Developing competence in biostatistics and research methodology during medical specialty training. *Med J Aust.* 2016 Feb 1;204(2):54-6e.1
6. Enders FT, Lindsell CJ, Welty LJ, Benn EKT, Perkins SM, Mayo MS, Rahbar MH, Kidwell KM, Thurston SW, Spratt H, Grambow SC, Larson J, Carter RE, Pollock BH, Oster RA. Statistical competencies for medical research learners: What is fundamental? *J Clin Transl Sci.* 2017 Jun;1(3):146-152.
7. Enders F. Evaluating mastery of biostatistics for medical researchers: need for a new assessment tool. *Clin Transl Sci.* 2011 Dec;4(6):448-54.
8. Dwivedi AK, Shukla R. Evidence-based statistical analysis and methods in biomedical research (SAMBR) checklists according to design features. *Cancer Rep (Hoboken).* 2020 Aug;3(4):e1211.
9. Hamet P, Tremblay J. Artificial intelligence in medicine. *Metabolism.* 2017 Apr;69S:S36-S40.
10. https://www.marketresearchfuture.com/reports/healthcare-artificial-intelligence-
11. Dwivedi AK. How to write statistical analysis section in medical research. J Investig Med. 2022 Dec;70(8):1759-1770.
12. Tan Y, Zhang WH, Huang Z, Tan QX, Zhang YM, Wei CY, Feng ZB. AI models predicting breast cancer distant metastasis using LightGBM with clinical blood markers and ultrasound maximum diameter. *Sci Rep.* 2024 Jul 6;14(1):15561. doi: 10.1038/s41598-024-66658-x.
13. Zhang WH, Tan Y, Huang Z, Tan QX, Zhang YM, Chen BJ, Wei CY. Development and validation of AI models using LR and LightGBM for predicting distant metastasis in breast cancer: a dual-center study. *Front Oncol.* 2024 Jun 14;14:1409273. doi: 10.3389/fonc.2024.1409273.
14. Kuijten RH, Zijlstra H, Groot OQ, Schwab JH. Artificial Intelligence and Predictive Modeling in Spinal Oncology: A Narrative Review. *Int J Spine Surg*. 2023 Jun;17(S1):S45-S56.

DOI : 10.62610/RJOR.2025.1.17.82

15. Paliwal A, Jain S, Kumar S, Wal P, Khandai M, Khandige PS, Sadananda V, Anwer MK, Gulati M, Behl T, Srivastava S. Predictive Modelling in pharmacokinetics: from in-silico simulations to personalized medicine. *Expert Opin Drug Metab Toxicol.* 2024 Apr;20(4):181-195.
16. Bassani D, Parrott NJ, Manevski N, Zhang JD. Another string to your bow: machine learning prediction of the pharmacokinetic properties of small molecules. *Expert Opin Drug Discov.* 2024 Jun;19(6):683-698.
17. Noorbakhsh-Sabet N, Zand R, Zhang Y, Abedi V. Artificial Intelligence Transforms the Future of Health Care. *Am J Med.* 2019 Jul;132(7):795-801.
18. Miller DD, Brown EW. Artificial Intelligence in Medical Practice: The Question to the Answer? *Am J Med.* 2018 Feb;131(2):129-133.
19. Liu Z, Elashoff D, Piantadosi S. Sparse support vector machines with L_0 approximation for ultra-high dimensional omics data. *Artif Intell Med*. 2019 May;96:134-141.
20. Cui H, Wang C, Maan H, Pang K, Luo F, Duan N, Wang B. scGPT: toward building a foundation model for single-cell multi-omics using generative AI. *Nat Methods*. 2024 Aug;21(8):1470-1480.
21. Pauwels R. A brief introduction to concepts and applications of artificial intelligence in dental imaging. *Oral Radiol*. 2021 Jan;37(1):153-160.
22. Mandal S, Greenblatt AB, An J. Imaging Intelligence: AI Is Transforming Medical Imaging Across the Imaging Spectrum. *IEEE Pulse.* 2018 Sep-Oct;9(5):16-24.
23. Su K, Xu L, Li G, Ruan X, Li X, Deng P, Li X, Li Q, Chen X, Xiong Y, Lu S, Qi L, Shen C, Tang W, Rong R, Hong B, Ning Y, Long D, Xu J, Shi X, Yang Z, Zhang Q, Zhuang Z, Zhang L, Xiao J, Li Y. Forecasting influenza activity using self-adaptive AI model and multi-source data in Chongqing, China. *EBioMedicine.* 2019 Sep;47:284-292.
24. Sidorenkov G, Stadhouders R, Jacobs C, Mohamed Hoesein FAA, Gietema HA, Nackaerts K, Saghir Z, Heuvelmans MA, Donker HC, Aerts JG, Vermeulen R, Uitterlinden A, Lenters V, van Rooij J, Schaefer-Prokop C, Groen HJM, de Jong PA, Cornelissen R, Prokop M, de Bock GH, Vliegenthart R. Multi-source data approach for personalized outcome prediction in lung cancer screening: update from the NELSON trial. *Eur J Epidemiol.* 2023 Apr;38(4):445-454.
25. Jackson G, Hu J; Section Editors for the IMIA Yearbook Section on Artificial Intelligence in Health. Artificial Intelligence in Health in 2018: New Opportunities, Challenges, and Practical Implications. *Yearb Med Inform*. 2019 Aug;28(1):52-54.
26. Ossowska A, Kusiak A, Świetlik D. Artificial Intelligence in Dentistry-Narrative Review. *Int J Environ Res Public Health.* 2022 Mar 15;19(6):3449. doi: 10.3390/ijerph19063449.
27. Rosenblatt WH, Yanez ND. A Decision Tree Approach to Airway Management Pathways in the 2022 Difficult Airway Algorithm of the American Society of Anesthesiologists. *Anesth Analg.* 2022 May 1;134(5):910-915.
28. Wei JX, Wang J, Zhu YX, Sun J, Xu HM, Li M. Traditional Chinese medicine pharmacovigilance in signal detection: decision tree-based data classification. *BMC Med Inform Decis Mak.* 2018 Mar 9;18(1):19. doi: 10.1186/s12911-018-0599-5.
29. Pallenberg R, Fleitmann M, Stroth AM, Gerlach J, Fürschke A, Barkhausen J, Bischof A, Handels H. Random Forest and Gradient Boosted Trees for Patient Individualized Contrast Agent Dose Reduction in CT Angiography. *Stud Health Technol Inform*. 2023 May 18;302:952-956.
30. Oguz BU, Shinohara RT, Yushkevich PA, Oguz I. Gradient Boosted Trees for Corrective Learning. *Mach Learn Med Imaging*. 2017 Sep;10541:203-211.
31. Heikamp K, Bajorath J. Support vector machines for drug discovery. *Expert Opin Drug Discov.* 2014 Jan;9(1):93-104.
32. Maltarollo VG, Kronenberger T, Espinoza GZ, Oliveira PR, Honorio KM. Advances with support vector machines for novel drug discovery. *Expert Opin Drug Discov*. 2019 Jan;14(1):23-33.
33. Silva RHLE, Machado AMC. Automatic measurement of pressure ulcers using Support Vector Machines and GrabCut. *Comput Methods Programs Biomed*. 2021 Mar;200:105867.
34. Siqueira Pinto M, Winzeck S, Kornaropoulos EN, Richter S, Paolella R, Correia MM, Glocker B, Williams G, Vik A, Posti JP, Haberg A, Stenberg J, Guns PJ, den Dekker AJ, Menon DK, Sijbers J, Van Dyck P, Newcombe VFJ. Use of Support Vector Machines Approach via ComBat Harmonized

DOI : 10.62610/RJOR.2025.1.17.82

Diffusion Tensor Imaging for the Diagnosis and Prognosis of Mild Traumatic Brain Injury: A CENTER-TBI Study. *J Neurotrauma*. 2023 Jul;40(13-14):1317-1338.

35. Ozer ME, Sarica PO, Arga KY. New Machine Learning Applications to Accelerate Personalized Medicine in Breast Cancer: Rise of the Support Vector Machines. *OMICS.* 2020;24(5):241-246.
36. Hennig C, What are the true clusters? *Pattern Recognit. Lett.* 2015; 64: 53-62.
37. Cornuéjols A, Wemmert C, Gançarski P, Bennani Y, Collaborative clustering: Why, when, what and how. *Information Fusion* 2018; 39: 81–95.
38. Kalita JK, Bhattacharyya DK, Roy S, Cluster analysis, in Fundamentals of Data Science: Theory and practice 2024; Academic Press, Elsevier, London, UK.
39. Han J, Pei J, Tong H, Cluster analysis: advanced methods, in Data Mining: Concepts and Techniques (4th edition) 2022; Morgan Kaufmann Publishers, Elsevier, Cambridge, US.
40. Jolliffe I, Principal Component Analysis, Springer Series in Statistics, Springer, 2002.
41. Zha H, He X, Ding C, Gu M, Simon HD, Spectral relaxation for k-means clustering, in: Dietterich TG, Becker S, Ghahramani Z (Eds.), Advances in Neural Information Processing Systems 14, MIT Press, 2002; 1057–1064.
42. MacEachern SJ, Forkert ND. Machine learning for precision medicine. *Genome.* 2021 Apr;64(4):416-425.
43. Kuntz S, Krieghoff-Henning E, Kather JN, Jutzi T, Höhn J, Kiehl L, Hekler A, Alwers E, von Kalle C, Fröhling S, Utikal JS, Brenner H, Hoffmeister M, Brinker TJ. Gastrointestinal cancer classification and prognostication from histology using deep learning: Systematic review. *Eur J Cancer.* 2021 Sep;155:200-215.
44. Haggenmüller S, Maron RC, Hekler A, Utikal JS, Barata C, Barnhill RL, Beltraminelli H, Berking C, Betz-Stablein B, Blum A, Braun SA, Carr R, Combalia M, Fernandez-Figueras MT, Ferrara G, Fraitag S, French LE, Gellrich FF, Ghoreschi K, Goebeler M, Guitera P, Haenssle HA, Haferkamp S, Heinzerling L, Heppt MV, Hilke FJ, Hobelsberger S, Krahl D, Kutzner H, Lallas A, Liopyris K, Llamas-Velasco M, Malvehy J, Meier F, Müller CSL, Navarini AA, Navarrete-Dechent C, Perasole A, Poch G, Podlipnik S, Requena L, Rotemberg VM, Saggini A, Sangueza OP, Santonja C, Schadendorf D, Schilling B, Schlaak M, Schlager JG, Sergon M, Sondermann W, Soyer HP, Starz H, Stolz W, Vale E, Weyers W, Zink A, Krieghoff-Henning E, Kather JN, von Kalle C, Lipka DB, Fröhling S, Hauschild A, Kittler H, Brinker TJ. Skin cancer classification via convolutional neural networks:systematic review of studies involving human experts.*Eur J Cancer*. 2021;156:202-216.
45. Laghari AA, Sun Y, Alhussein M, Aurangzeb K, Anwar MS, Rashid M. Deep residual-dense network based on bidirectional recurrent neural network for atrial fibrillation detection. *Sci Rep.* 2023 Sep 13;13(1):15109. doi: 10.1038/s41598-023-40343-x.
46. Lee S, Hussein R, Ward R, Jane Wang Z, McKeown MJ. A convolutional-recurrent neural network approach to resting-state EEG classification in Parkinson's disease. *J Neurosci Methods.* 2021 Sep 1;361:109282. doi: 10.1016/j.jneumeth.2021.109282.
47. Lin X, Jiang S, Gao L, Wei Z, Wang J. MultiSC: a deep learning pipeline for analyzing multiomics single-cell data. *Brief Bioinform*. 2024 Sep 23;25(6):bbae492.
48. Hunter DJ, Holmes C. Where Medical Statistics Meets Artificial Intelligence. *N Engl J Med.* 2023 Sep 28;389(13):1211-1219.
49. Koski E, Murphy J. AI in Healthcare. *Stud Health Technol Inform.* 2021;284:295-299.
50. Smallman M. Multi Scale Ethics-Why We Need to Consider the Ethics of AI in Healthcare at Different Scales. *Sci Eng Ethics*. 2022 Nov 28;28(6):63. doi: 10.1007/s11948-022-00396-z.
51. Elendu C, Amaechi DC, Elendu TC, Jingwa KA, Okoye OK, John Okah M, Ladele JA, Farah AH, Alimi HA. Ethical implications of AI and robotics in healthcare: A review. *Medicine (Baltimore).* 2023 Dec 15;102(50):e36671. doi: 10.1097/MD.0000000000036671.
52. Goirand M, Austin E, Clay-Williams R. Implementing Ethics in Healthcare AI-Based Applications: A Scoping Review. *Sci Eng Ethics.* 2021 Sep 3;27(5):61.
53. Vandemeulebroucke T. The ethics of artificial intelligence systems in healthcare and medicine: from a local to a global perspective, and back. *Pflugers Arch.* 2024 Jul 6. doi: 10.1007/s00424-024-02984-3.

DOI : 10.62610/RJOR.2025.1.17.82

54. Rogers WA, Draper H, Carter SM. Evaluation of artificial intelligence clinical applications: Detailed case analyses show value of healthcare ethics approach in identifying patient care issues. *Bioethics*. 2021 Sep;35(7):623-633.
55. Wang F, Preininger A. AI in Health: State of the Art, Challenges, and Future Directions. *Yearb Med Inform.* 2019 Aug;28(1):16-26.
56. Almalki J, Alshahrani SM, Khan NA. A comprehensive secure system enabling healthcare 5.0 using federated learning, intrusion detection and blockchain. *PeerJ Comput Sci.* 2024 Jan 10;10:e1778. doi: 10.7717/peerj-cs.1778.
57. Waqas A, Bui MM, Glassy EF, El Naqa I, Borkowski P, Borkowski AA, Rasool G. Revolutionizing Digital Pathology With the Power of Generative Artificial Intelligence and Foundation Models. *Lab Invest.* 2023 Nov;103(11):100255. doi: 10.1016/j.labinv.2023.100255.
58. Lu MY, Chen B, Williamson DFK, Chen RJ, Zhao M, Chow AK, Ikemura K, Kim A, Pouli D, Patel A, Soliman A, Chen C, Ding T, Wang JJ, Gerber G, Liang I, Le LP, Parwani AV, Weishaupt LL, Mahmood F. A multimodal generative AI copilot for human pathology. *Nature.* 2024 Oct;634(8033):466-473.
59. Tosun AB, Pullara F, Becich MJ, Taylor DL, Fine JL, Chennubhotla SC. Explainable AI (xAI) for Anatomic Pathology. *Adv Anat Pathol*. 2020 Jul;27(4):241-250.
60. Zhang Y, Weng Y, Lund J. Applications of Explainable Artificial Intelligence in Diagnosis and Surgery. *Diagnostics (Basel).* 2022 Jan 19;12(2):237. doi: 10.3390/diagnostics12020237.
61. Cong Y, Endo T. A Quadruple Revolution: Deciphering Biological Complexity with Artificial Intelligence, Multiomics, Precision Medicine, and Planetary Health. *OMICS.* 2024 Jun;28(6):257-260.
62. Hajirasouliha I, Elemento O. Precision medicine and artificial intelligence: overview and relevance to reproductive medicine. *Fertil Steril.* 2020 Nov;114(5):908-913.
63. Johnson KB, Wei WQ, Weeraratne D, Frisse ME, Misulis K, Rhee K, Zhao J, Snowdon JL. Precision Medicine, AI, and the Future of Personalized Health Care. *Clin Transl Sci.* 2021 Jan;14(1):86-93.
64. Williams AM, Liu Y, Regner KR, Jotterand F, Liu P, Liang M. Artificial intelligence, physiological genomics, and precision medicine. *Physiol Genomics*. 2018 Apr 1;50(4):237-243.
65. Organizing Committee of the Madrid 2017 Critical Care Datathon; Núñez Reiz A, Martínez Sagasti F, Álvarez González M, Blesa Malpica A, Martín Benítez JC, Nieto Cabrera M, Del Pino Ramírez Á, Gil Perdomo JM, Prada Alonso J, Celi LA, Armengol de la Hoz MÁ, Deliberato R, Paik K, Pollard T, Raffa J, Torres F, Mayol J, Chafer J, González Ferrer A, Rey Á, González Luengo H, Fico G, Lombroni I, Hernandez L, López L, Merino B, Cabrera MF, Arredondo MT, Bodí M, Gómez J, Rodríguez A, Sánchez García M. Big data and machine learning in critical care: Opportunities for collaborative research. *Med Intensiva (Engl Ed).* 2019 Jan-Feb;43(1):52-57.
66. Cheung YB, Tan SB, Khoo KS. The need for collaboration between clinicians and statisticians: some experience and examples. *Ann Acad Med Singap*. 2001 Sep;30(5):552-5.

DOI : 10.62610/RJOR.2025.1.17.82