Journal of Pioneering Medical Sciences
Received: August 18, 2025 | Accepted: October 11, 2025 | Published: January 05, 2026
Volume 14, Issue 12, Pages 38-42

DOI https://doi.org/10.47310/jpms2025141206

Artificially Intelligent (AI) Techniques in Student Assessment: Insight and Future Directions

Wajid Ali Chatha[1*], Syed Shahzad Hussain Shah[2], Asmara Syed[3], Zainiya Sherazi[4], Ekramy El Elmorsy[5], Abdul Ghaffar[6], Syed Sajid Hussain Shah[7], Mohamed Soliman[8], Nagla M Shalaby[9], Padmanabha BV[10] and Manahil Muhammad[11]

[1-3]Faculty of Medicine, Northern Border University, Arar, Kingdom of Saudi Arabia
[2]National University of Technology, Islamabad, Pakistan
[4]Department of Computer Science, COMSATS University, Islamabad, Pakistan
[5-7]College of Medicine, Northern Border University, Arar, Kingdom of Saudi Arabia
[8]Department of Parasitology, College of Medicine, Northern Border University, Arar, Kingdom of Saudi Arabia and Faculty of Medicine, Mansoura University, Egypt
[9]College of Medicine, Northern Border University, Arar, Kingdom of Saudi Arabia and Faculty of Medicine, Mansoura University, Egypt
[10]Assistant Professor of Physiology. College of Medicine, Northern Border University, Arar, Kingdom of Saudi Arabia
[11]BSCS Final year student, University of London, United Kingdom

Author Designation: [1]Associate Professor, [3,6]Lecturer, [4,11]Student, [5]Associate Professor, [7-8]Professor, [9,10]Assistant Professor

*Corresponding author: Wajid Ali Chatha (e-mail: chaudhary.chatha@nbu.edu.sa).

Abstract: Education is fundamental to development of any society anywhere and effective assessment is crucial for ensuring the quality and results of educational processes. With the swift progress in technology, Artificial Intelligence (AI) has become a revolutionary tool in education, especially in assessment techniques. This review examines the use, advantages, challenges, and future directions of AI-driven assessments in medical education. A systematic literature review was performed utilizing PubMed, Scopus, and Google Scholar, concentrating on studies that investigate the role of AI in medical education and evaluation. From an initial pool of 3,051 studies, 31 were selected based on strict inclusion criteria after a thorough multi-stage selection process. The findings suggest that AI can improve assessment accuracy, transparency, and efficiency through automated grading, adaptive testing, and prompt feedback. AI-driven assessments also have the potential to reduce human biases, lessen the workload for educators, and promote personalized learning experiences. However, challenges such as data privacy, algorithmic bias, and the necessity for AI literacy among educators continue to pose significant issues. In summary, the integration of AI into medical education assessments offers a chance to enhance evaluation standards, support competency-based education, and cultivate the development of more proficient healthcare professionals.

Key Words: Artificial Intelligence, Medical Education, Assessment. Computation, Technique

INTRODUCTION

Education, be it professional or traditional, has a fundamental role in the development and growth of the society. Assessment of this role of education depends on various indicators that need to defined and used for the same. The assessment is an important component of the educational process and it plays a vital role in the determination of quality of the education. The use of good assessment techniques in the educational program is a key factor in the prosperity of nation as it assures the provision of good skillful professionals to the community.

Artificial Intelligence (AI) and it's rising use in various computational technologies is revolutionizing various domains, including education and assessment. Its ability to mimic human intelligence and process vast amounts of data has positioned AI as a powerful tool for improving and automating educational assessments. This review examines the key applications, benefits, challenges, and future trends in AI-driven assessment.

The assessment could be performed by employing different methods. The summative assessment is quite commonly used while the formative assessment is also an important component of the academic program [1]. The summative assessment is done at the end of the session and it determines the level of knowledge and skill of the student and documents the marks obtained by the student. The formative assessment is carried out during the academic session to find out the deficiency in the knowledge and it provides a timely

feedback to the learner in order to bridge the knowledge gaps for better comprehension of the learning topics. The formative assessment helps the learner in obtaining high academic achievements [2]. The other types of assessments are norm-referenced assessment, criterion – referenced assessment, interim / benchmark assessments [3-5].

Learning from our experiences over the passage of time paves the pathway for the further improvements. These could then be achieved by bringing the modification and changes in the current practices. These updates are required in the all fields and would bring better results. Similarly updates in the assessment process are also required. The old techniques could contribute to the low performance in the current student population which is highly inclined to use of various artificially intelligent apps, softwares, etc. [6].

The replacement of old methods with the new technology could also improve the accuracy and transparency in the assessments and it is much more required in the all fields of life, but specially in health care, where it would have more effects on human wellbeing. The application of digital technology in the assessments could also improve the accuracy, transparency and reduce the laborious tasks of examination on the teachers. The use of digital technology in many fields including the health care, education, communication and in many businesses works has revealed encouraging results [7-10].

Artificial intelligence a rapidly emerging technique in the era of digital technology. The artificial intelligence deals with the development of intelligent software in which the system could learn and act autonomously that would make the computer to work like a part of human brain with the capability of cognitive function. One of the applications of artificial intelligence is machine learning in which the model learns and improves automatically [11].

The incorporation of artificial intelligence in education especially the medical education could be very useful for the application of the innovative assessment techniques [12]. The use of artificial intelligence could also make provision of immediate feedback to the learners [13]. There is a need for the update in the medical curriculum so that the physicians should learn the applications of artificial intelligence [14].

The main objective of this paper is to review the published literature regarding the use of artificial intelligence in the assessment of medical students and sketch the past and current trends and where and when to and not to select these AI programs.

Objectives of the Study

Briefly in this review, the authors wish to:

- Review the role of Artificial Intelligence (AI) in the assessment of medical students
- Evaluate the benefits and challenges of AI-driven assessment methods
- Explore the application of AI in formative and summative assessments within medical education
- Provide recommendations for integrating AI-based assessments to improve learning outcomes and professional competency

METHODS

A thorough literature review was undertaken to investigate the role of Artificial Intelligence (AI) in the realms of medical education and assessment. This review adhered to a structured, three-phase methodology: Identification, Screening, and Selection, all directed by established inclusion and exclusion criteria. The study complied with systematic review standards akin to PRISMA guidelines, ensuring both transparency and reproducibility. The review encompassed publications from January 2001 to August 2025. Most of the studies included dated after 2010 as in the beginning of the twenty first century AI was in a nascent stage.

Identification

A systematic search was executed across reputable scientific databases, including PubMed, Scopus, and Google Scholar, utilizing key phrases such as "artificial intelligence," "medical education," and "assessment." The objective of the search was to gather studies that examined both technological innovations and their applications in the evaluation of medical students and professionals. This phase resulted in an initial collection of 3,051 potentially pertinent articles.

Screening

The titles and abstracts of the collected articles were independently evaluated by three reviewers to determine their relevance to the research aims. Articles were excluded if they were:

- Not pertinent to AI applications in medical education
- Not published in English
- Conference abstracts, editorials, or commentaries lacking primary data, or
- Duplicates

After the screening process, 247 articles were selected for full-text evaluation.

Selection

The full texts of the remaining articles were independently reviewed by the same three reviewers for scientific rigor and relevance. The quality assessment focused on study design, methodology, clarity of results, and the contribution to the understanding of AI in medical education. Any disagreements were resolved through discussions among the reviewers. In the end, 31 studies were chosen for inclusion in the final analysis.

This structured methodology guaranteed a comprehensive, impartial, and transparent review of the existing literature regarding AI applications in medical education and assessment.

A PRISMA style flow diagram is attached as Figure I, explaining the literature selection for this paper.

Ethical Approval

No ethical approval was applied / required as per the institutional policy, as this article is using the already published data and not using any human tissue studies.

RESULTS

A large number of studies were found and this reflects the growing body of research at the intersection of these fields, demonstrating a broad interest in the application of AI within medical education systems.

The assessment of essays in the formative and summative assessment with the help of AI has been reported to have yielded very promising results. The automated essay scoring system has emerged as a very elegant technical tool for the summative assessment at academic programs which require to assess a large number of candidates. Time is an essence for communicating the results to students and AI driven scoring systems could also provide a rapid feed back to the students in the formative assessment.

Table 1: Overview of AI Techniques and Their Applications in Educational Assessment

AI Technique	Purpose / Application	Key Techniques Used	Examples / Software / Platforms
1. Machine Learning (ML)	Predicting student outcomes, automating grading, adaptive testing	Decision Trees, Random Forests, Support Vector Machines (SVM), Neural Networks, Reinforcement Learning	• Gradescope (Turnitin) – ML-assisted grading • Knewton Alta – adaptive learning • Smart Sparrow – adaptive quizzes • Civitas Learning – dropout prediction • Coursera ML Grading – automated scoring
2. Natural Language Processing (NLP)	Automated essay scoring, text analysis, plagiarism detection, feedback generation	Text classification (BERT, RoBERTa), semantic analysis, syntax parsing, sentiment analysis	• ETS e-rater (TOEFL/GRE) – automated essay scoring • Turnitin – plagiarism detection and text similarity analysis • Grammarly – grammar and writing feedback • OpenEssayist – essay structure and argument analysis • Project Essay Grade (PEG) – regression-based essay grading
3. Computer Vision (CV)	Handwriting recognition, diagram/object evaluation, engagement monitoring	Optical Character Recognition (OCR), image segmentation, facial recognition, emotion detection	• Gradescope – OCR-based grading for handwritten work • Proctorio – webcam-based remote proctoring • Google Cloud Vision API – handwriting and object recognition • Examus.ai – attention and engagement tracking
4. Intelligent Tutoring Systems (ITS)	Continuous formative assessment, personalized learning, mastery tracking	Bayesian Knowledge Tracing (BKT), Deep Knowledge Tracing (LSTM), Reinforcement Learning	• Carnegie Learning (MATHia) – math mastery tracking • ALEKS – adaptive assessment in math and science • Duolingo – reinforcement-based language learning • Cognii Virtual Tutor – conversational AI short-answer evaluation • DreamBox Learning – adaptive math instruction
5. Predictive Analytics	Forecasting academic success, identifying at-risk students, curriculum improvement	Regression models, Gradient Boosting (XGBoost), Neural Networks	• Civitas Learning – student retention prediction • Blackboard Analytics for Learn – engagement insights • Brightspace Insights (D2L) – outcome prediction • Panorama Education – social-emotional learning analytics
6. Learning Analytics and Educational Data Mining (EDM)	Tracking learning behavior, detecting misconceptions, performance visualization	Clustering (K-Means), Association Rule Mining, Time-Series Analysis	• Moodle Inspire – predictive learning analytics • Canvas Learning Analytics – performance dashboards • OpenLearning Analytics Framework – open-source analytics tools • Edmodo Insights – participation and progress analysis
7. Generative AI	Automatic question generation, rubric creation, adaptive assessment design	Large Language Models (LLMs), Question Generation (QG) models, Transformer architectures	• ChatGPT (OpenAI) – generating questions, rubrics, and feedback • Quizgecko – automatic quiz generation from text • Quillionz – question generation from study material • QuestionWell – AI-generated assessment creation • ExamSoft (AI extensions) – rubric automation
8. Explainable AI (XAI)	Ensuring fairness, transparency, and bias control in automated grading	SHAP (Shapley Additive Explanations), LIME, Bias Detection Frameworks	• IBM Watson OpenScale – fairness and model tracking • Microsoft Azure Responsible AI Toolkit – interpretability reports • Google What-If Tool – visualization and bias detection in AI models

Computerized adaptive testing is now an important AI tool for the summative assessment in many institutes. The studies showed that the automation of grading of the essay is possible with the help of the NLP algorithms. These algorithms revealed reliability which was at par with the evaluation grades done by humans. The I-MIB model is reported to have also revealed very encouraging results during the evaluation process of the undergraduates at the national level.

Table 1 depicts various AI driven modules in vogue.

DISCUSSION

The integration of advanced technologies, particularly artificial intelligence (AI), in medical education has become increasingly central for enhancing healthcare outcomes and developing competent healthcare professionals. Fully trained health professionals are essential for addressing societal health needs, and high-quality assessments play a critical role in evaluating their skills and guiding learning strategies. More accurate and transparent assessments can positively influence learning outcomes, while multiple-choice questions (MCQs) remain a widely accepted tool for assessing medical knowledge [15-17].

Learners exhibit different approaches to learning. Deep learners are intrinsically motivated to understand and engage with material, whereas surface learners are often driven by external factors and demonstrate lower engagement [18-19]. Effective assessment strategies can help differentiate between these approaches, enabling tailored educational support.

The assessments in medical education are quite labor intensive and time-consuming tasks. The assessments of assignments and essay questions may take quite significant time of teachers. The detection of plagiarized content in the assignments is very difficult for the teachers without the help of digital assessments tool.

AI applications offer promising solutions to these challenges. They can enhance formative assessment, automate essay scoring, and enable personalized evaluations, improving fairness, validity, and efficiency [20-23]. Adaptive testing, powered by AI, dynamically adjusts question difficulty based on performance, reducing test fatigue, minimizing guesswork, and enabling rapid identification of learning gaps. Internationally, computerized adaptive testing (CAT) models, such as those used in the GRE and GMAT, demonstrate the feasibility and reliability of AI-driven assessments.

Within competency-based medical education, AI has the potential to restructure assessment systems and improve learning outcomes [24-26]. Models such as the Intelligent Multilevel Item Bank (I-MIB) illustrate how digital tools can facilitate individualized and accurate evaluation [27]. Integrating AI into medical education can also enhance the effectiveness of educators and healthcare professionals, supporting more informed teaching and learning practices.

Despite these advantages, limitations and ethical considerations must be addressed. AI systems are dependent on the quality of their training data, and algorithmic biases may perpetuate inequities. Privacy, data security, and accountability in assessment decision-making require careful management. Furthermore, successful implementation necessitates faculty training to interpret AI-generated insights and integrate them effectively into curricula.

Looking forward, AI is likely to play an expanding role in medical education. Structured training programs are needed to equip both educators and students with the competencies required for responsible AI use. Addressing ethical, practical, and pedagogical challenges through research, policy development, and international collaboration will ensure that AI-driven assessments enhance educational outcomes while maintaining fairness and integrity.

CONCLUSION

The application of artificial intelligence (AI) in medical education is poised to significantly reshape assessment methodologies in the near future. By integrating AI algorithms into the assessment process, educators can leverage these advanced technologies as powerful adjunct tools to enhance both formative and summative evaluations. The instant feedback loop can accelerate the learning process and foster a deeper understanding of complex medical concepts. Medical education places a strong emphasis on academic integrity, and AI-powered smart tools can effectively identify instances of plagiarism. In addition, by comparing student submissions against extensive databases of published work, these tools can flag potential violations, enabling educators to uphold ethical standards and promote original thinking among students.

By harnessing AI's capabilities in grading, feedback, plagiarism detection, and data analysis, educators can create a more personalized and effective learning environment that prepares future healthcare professionals to excel in their careers. As AI technology continues to evolve, its potential to improve medical education will likely expand, driving innovation and excellence in the field of medicine.

Future Directions

AI is emerging and its future course is not foreseeable at the moment. Future research should focus on developing more transparent and explainable AI algorithms to strengthen trust in automated assessments. Further studies are needed to explore how AI-driven adaptive testing can be optimized for diverse medical specialties and learning contexts. Expanding structured training programs for faculty and students will be essential to ensure responsible and effective integration of AI tools. Local and International collaboration among various institutes can help establish standardized guidelines for ethical AI use in medical assessment systems.

Limitations of the Study

A major limitation is the risk of algorithmic bias, as AI outputs depend heavily on the quality and representativeness

of training data. Concerns regarding data privacy, security, and accountability may hinder full adoption without robust regulatory frameworks. Faculty may face challenges in interpreting AI-generated insights, highlighting the need for capacity-building and digital literacy. The high cost and technical complexity of AI systems may limit accessibility, particularly in resource-constrained medical institutions.

Acknowledgement

The authors are thankful to Dr. Saima Syed and Inshrah Syed for their assistance.

Conflict of interest

Authors report that there is no conflict of interest.

REFERENCES

[1] Ismail, S. *et al.* "Formative vs. summative assessment: impacts on academic motivation, attitude toward learning, test anxiety, and self-regulation skill." *Language Testing in Asia*, vol. 12, no. 1, 2022, pp. 40. doi:10.1186/s40468-022-00191-4.

[2] Ozan, C. and Kıncal, R.Y. "The effects of formative assessment on academic achievement, attitudes toward the lesson, and self-regulation skill." *Educational Sciences: Theory and Practice*, vol. 18, no. 1, 2018, pp. 85–118. doi:10.12738/estp.2018.1.0216.

[3] Hendricks, A.E. and Adlof, S.M. "Language assessment with children who speak nonmainstream dialects: examining the effects of scoring modifications in norm-referenced assessment." *Language, Speech, and Hearing Services in Schools*, vol. 48, no. 3, 2017, pp. 168–182. doi:10.1044/2017_LSHSS-16-0060.

[4] Ricketts, C. "A plea for the proper use of criterion-referenced tests in medical assessment." *Medical Education*, vol. 43, no. 12, 2009, pp. 1141–1146. doi:10.1111/j.1365-2923.2009.03541.x.

[5] Konstantopoulos, S. *et al.* "Effects of interim assessments across the achievement distribution: evidence from an experiment." *Educational and Psychological Measurement*, vol. 76, no. 4, 2016, pp. 587–608. doi:10.1177/0013164415606498.

[6] Ishaq, K. *et al.* "Exploring summative assessment and effects: primary to higher education." *Bulletin of Education and Research*, vol. 42, no. 3, 2020, pp. 23–50.

[7] Toivonen, J. *et al.* "Radiomics and machine learning of multisequence multiparametric prostate MRI: towards improved non-invasive prostate cancer characterization." *PLOS One*, vol. 14, no. 7, 2019, e0217702. doi:10.1371/journal.pone.0217702.

[8] Campanella, G. *et al.* "Clinical-grade computational pathology using weakly supervised deep learning on whole slide images." *Nature Medicine*, vol. 25, 2019, pp. 1301–1309. doi:10.1038/s41591-019-0508-1.

[9] Shah, S.S.H. *et al.* "The evaluation of artificial intelligence technology for the differentiation of fresh human blood cells from other species' blood in the investigation of crime scenes." *Cureus*, vol. 16, no. 4, 2024, e58496. doi:10.7759/cureus.58496.

[10] Humayun, M. *et al.* "The effectiveness of online learning for the cognitive domain of course learning outcomes from the students' point of view." *Pakistan Journal of Medical and Health Sciences*, vol. 15, no. 9, 2021, pp. 2186–2188. doi:10.53350/pjmhs211592186.

[11] Saini, N. "Research paper on artificial intelligence and its applications." *International Journal for Research Trends and Innovation*, vol. 8, no. 4, 2023, pp. 356–360.

[12] Abdellatif, H. *et al.* "Teaching, learning and assessing anatomy with artificial intelligence: the road to a better future." *International Journal of Environmental Research and Public Health*, vol. 19, no. 21, 2022, pp. 14209. doi:10.3390/ijerph192114209.

[13] Hamilton, B.C. *et al.* "Artificial intelligence based real-time video ergonomic assessment and training improves resident ergonomics." *American Journal of Surgery*, vol. 226, no. 5, 2023, pp. 741–746. doi:10.1016/j.amjsurg.2023.07.028.

[14] Civaner, M.M. *et al.* "Artificial intelligence in medical education: a cross-sectional needs assessment." *BMC Medical Education*, vol. 22, 2022, pp. 772. doi:10.1186/s12909-022-03852-3.

[15] Shah, S.S.H. "The evaluation of undergraduate medical students' perception regarding the summative assessment of pathology." *Pakistan Journal of Medical and Health Sciences*, vol. 13, no. 4, 2019, pp. 1030–1032.

[16] Coughlin, P.A. and Featherstone, C.R. "How to write a high quality multiple choice question (MCQ): a guide for clinicians." *European Journal of Vascular and Endovascular Surgery*, vol. 54, no. 5, 2017, pp. 654–658. doi:10.1016/j.ejvs.2017.07.012.

[17] Müller, S. *et al.* "A pilot survey of student perceptions on the benefit of the OSCE and MCQ modalities." *GMS Journal for Medical Education*, vol. 35, no. 4, 2018, Doc51. doi:10.3205/zma001197.

[18] Snelgrove, S. and Slater, J. "Approaches to learning: psychometric testing of a study process questionnaire." *Journal of Advanced Nursing*, vol. 43, no. 5, 2003, pp. 496–505. doi:10.1046/j.1365-2648.2003.02747.x.

[19] Zeegers, P. "Approaches to learning in science: a longitudinal study." *British Journal of Educational Psychology*, vol. 71, 2001, pp. 115–132. doi:10.1348/000709901158424.

[20] Gardner, J. *et al.* "Artificial intelligence in educational assessment: 'Breakthrough? Or buncombe and ballyhoo?'" *Journal of Computer Assisted Learning*, vol. 37, no. 5, 2021, pp. 1207–1216. doi:10.1111/jcal.12577.

[21] Kakungulu, S.J. "The role of artificial intelligence in educational assessment." *Eurasian Experiment Journal of Scientific and Applied Research*, vol. 5, no. 2, 2024, pp. 44–48.

[22] Gierl, M.J. *et al.* "Automated essay scoring and the future of educational assessment in medical education." *Medical Education*, vol. 48, no. 10, 2014, pp. 950–962. doi:10.1111/medu.12517.

[23] Wang, Z. "Computer-assisted EFL writing and evaluations based on artificial intelligence: a case from a college reading and writing course." *Library Hi Tech*, vol. 40, no. 1, 2022, pp. 80–97. doi:10.1108/LHT-05-2020-0113.

[24] Lentz, A. *et al.* "AI-ssessment: towards assessment as a sociotechnical system for learning." *Academic Medicine*, vol. 96, no. 7S, 2021, pp. S87–S88. doi:10.1097/ACM.0000000000004104.

[25] Kolade, O. *et al.* "Is AI changing learning and assessment as we know it? Evidence from a ChatGPT experiment and a conceptual framework." *Heliyon*, vol. 10, no. 4, 2024, e25953. doi:10.1016/j.heliyon.2024.e25953.

[26] Janpla, S. and Piriyasurawong, P. "The development of an intelligent multilevel item bank model for the national evaluation of undergraduates." *Universal Journal of Educational Research*, vol. 8, no. 9, 2020, pp. 4163–4172. doi:10.13189/ujer.2020.080942.

[27] Sridharan, K. and Sequeira, R.P. "Artificial intelligence and medical education: application in classroom instruction and student assessment using a pharmacology and therapeutics case study." *BMC Medical Education*, vol. 24, no. 1, 2024, pp. 431. doi:10.1186/s12909-024-05365-7.