Journal of Contemporary Educational Research, 2026, Volume 10, Issue 9
https://bbwpublisher.com/journal/JCER
Online ISSN: 2208-8474
Print ISSN: 2208-8466

Current Status of the Application of Big Data Models in Standardized Residency Training of Orthopedic Residents

Yuanguo Luo[†]*, Feng Wu[†], Yuxi Li, Guojian Chen, Guodong Wu, Dongdong Shi

The 923rd Hospital of the Joint Logistics Support Force of the Chinese People's Liberation Army, Nanning 530021, Guangxi, China
[†]These authors contributed equally to the work.

Author to whom correspondence should be addressed.

Abstract: Standardized residency training is an essential pathway for cultivating competent orthopedic surgeons, while big-data-based models are reshaping medical education. Based on domestic and international literature published in the past decade, this review summarizes the current application of big data models in standardized residency training of orthopedic residents from five aspects: data-driven process management and individualized teaching, deep-learning-based radiograph interpretation training, virtual and augmented reality simulation, large language model-assisted learning and assessment, and predictive analytics for training outcomes. Major challenges include data silos and quality problems, algorithmic opacity and bias, risks of overreliance and hallucination, as well as the lag in evaluation systems and faculty digital literacy. Future efforts should focus on standardized data governance, integration of artificial intelligence literacy into the curriculum, and rigorous educational research to evaluate the real teaching effectiveness, so as to achieve the deep integration of technology empowerment and competency-based training.

Keywords: Big data; Artificial intelligence; Standardized residency training; Orthopedics; Medical education

Online publication: September 17, 2026

1. Introduction

Since 2013, China has fully implemented the standardized residency training (hereinafter referred to as "residency training") system. In 2017, the "Opinions of the General Office of the State Council on Deepening the Collaboration between Medical Education and Further Promoting the Reform and Development of Medical Education" (Guo Ban Fa [2017] No. 63) further emphasized improving the training system and enhancing training quality [1]. In 2020, the "Guiding Opinions of the General Office of the State Council on Accelerating the Innovative Development of Medical Education" (Guo Ban Fa [2020] No. 34) proposed strengthening the deep integration of modern information technology with medical education and teaching, and exploring new forms of intelligent medical education [2]. Against this background, how to utilize educational big data to improve the quality of residency training has become a hot topic in medical education research.

Orthopedics is a typical highly operative and imaging-dependent discipline. Residents need to master core competencies such as clinical decision-making, image interpretation, and complex surgical operations within a limited training period. Traditional training is mainly based on bedside teaching, intraoperative mentoring, and apprenticeship-style instruction, which has limitations such as difficulty in retaining process data, uneven distribution of operative opportunities, and delayed and subjective feedback. Meanwhile, the scale of medical data is growing rapidly, and big data models—namely predictive and generative models based on massive multi-source data and using machine learning, especially deep learning and large language models, as core methods—have shown performance approaching or even reaching expert level in disease diagnosis, treatment decision-making, and prognosis prediction [3–5]. The medical education community has accordingly proposed that the talent cultivation paradigm should move from the "information age" to the "artificial intelligence age" [6].

At present, intelligent tools such as preoperative planning, imaging measurement, and prosthesis selection have gradually entered the orthopedic clinical workflow, and residents need to learn to collaborate with them during training. Systematically reviewing the current status of big data model applications in orthopedic residency training not only responds to the practical pressure of technological development but also represents a proactive layout for the connotative development of residency training.

Based on domestic and international literature from the past decade (2016–2026), this paper searched PubMed, Web of Science, CNKI, and Wanfang databases using combinations of subject terms and free words such as "big data," "artificial intelligence," "machine learning," "large language model," "virtual reality," "standardized residency training," and "orthopedics," and manually traced references cited in important reviews. A total of 26 representative articles were ultimately included (including 2 randomized controlled trials and 4 systematic or scoping reviews) to review the current status and challenges of big data model applications in orthopedic residency training.

2. Main pathways of big data models empowering orthopedic residency training

2.1. Data-driven training process management and individualized teaching

The essence of residency training is a long-cycle, multi-rotation, continuous cultivation process, and its quality depends on the ability to collect, analyze, and provide feedback on training process data. In recent years, domestic residency training bases have relied on information management platforms to gradually achieve structured collection of process data such as orientation education, disease registration, number of operations, stage assessments, and teaching activities. Wang Xinbei et al. collected daily assessment and base performance data of 556 residents from a tertiary hospital from 2018 to 2021, and used logistic regression to construct a prediction model for professional theoretical examination scores in the completion assessment, confirming that process data have quantifiable predictive value for training outcomes [7]. In terms of assessment and evaluation, competency-oriented formative evaluation tools such as Mini-CEX and DOPS have shown good results in orthopedic residency training, and the structured assessment data they generate provide material for correlation analysis between "process data and training outcomes" [8]. The above practices have shifted teaching management from experience-driven to evidence-driven, compensating for the shortcomings of traditional summative evaluation that "emphasizes results and neglects process."

Learning analytics provides methodological support for the above practices. A scoping review by Bojic et al. pointed out that the measurement, collection, analysis, and reporting of data on learners and their learning contexts can provide data-driven insights and individualized support for medical education, especially with potential in predicting learning difficulties and optimizing intervention timing [9]. In the orthopedic residency training scenario, this can be refined into: personalized recommendation of rotation prescriptions based on disease and operation registration data, diagnosis of weak areas based on assessment knowledge points, and characterization of skill growth curves based on operation logs. Wartman et al. pointed out that the core competitiveness of medical education in the artificial intelligence era will shift from knowledge memorization to critical application of knowledge produced by data and algorithms, which requires that process management data ultimately serve the development of clinical thinking and post competency, rather than mere report-based assessment [6].

2.2. Intelligent imaging models and fracture interpretation training

Image interpretation is one of the core competencies of orthopedic residents and is also the field where deep learning first achieved breakthroughs. Kim et al. used convolutional neural networks to detect wrist fractures on X-ray films, showing interpretation ability approaching that of orthopedic physicians and demonstrating the feasibility of transfer learning in small-sample medical imaging tasks [10]. Domestic reviews also show that deep learning has achieved near-expert performance in orthopedic imaging tasks such as X-ray fracture detection [11]. Oakden-Rayner et al. validated the detection performance of a deep learning system for proximal femoral fractures in the emergency department in a diagnostic accuracy study, and through algorithmic auditing revealed characteristic missed detection patterns of the model in specific patient subgroups and imaging situations [12]. The implications of such research for teaching are bidirectional: intelligent imaging systems can serve as "sparring partners" for residents' self-study, providing immediate interpretation feedback; introducing algorithmic audit results into film-reading teaching can help cultivate residents' critical awareness of the limitations of artificial intelligence.

It should be emphasized that artificial intelligence in fracture interpretation training is a "double-edged sword." When a model prioritizes high sensitivity, it inevitably accompanies false positives; when it prioritizes high specificity, it may miss diagnoses. If residents lack understanding of basic concepts such as sensitivity, specificity, and receiver operating characteristic curves, they may be at a loss when "human and machine results are inconsistent." Therefore, when embedding intelligent imaging models into teaching, their performance boundaries and failure modes should be taught simultaneously, so that residents develop an interpretation habit of "human-machine collaboration, with humans as the primary decision-makers."

In terms of instructional design, intelligent imaging models can be integrated into residency training in a "human first, machine second" mode: residents first independently interpret and record diagnostic confidence, then compare with model output, focusing on reviewing cases of human-machine disagreement and converting disagreement points into teaching cases; model visualization heatmaps can be used to explain the spatial distribution of fracture signs, but it must be clarified that heatmaps only indicate the model's areas of attention rather than diagnostic evidence; performance concepts such as sensitivity, specificity, and positive predictive value can also be incorporated into end-of-rotation assessment questions. Existing research on human-machine collaborative interpretation teaching mostly targets radiology residents, and controlled studies in the orthopedic specialty environment are still lacking, urgently requiring localized exploration.

2.3. Virtual reality and augmented reality simulation training

Surgical operation skills training is the top priority of orthopedic residency training and is also the scenario where big data models and immersive technologies are most closely integrated. A randomized controlled trial by Lohre et al. showed that subjects receiving immersive virtual reality (VR) training were significantly superior to the control group in acquiring complex skills [13]; their randomized controlled trial targeting senior surgical residents further confirmed that VR training brought significant improvements in both knowledge and skill outcomes [14]. A systematic review by Clarke considered that VR simulation is expected to become an important component of the orthopedic training system, but called for more rigorous long-term efficacy studies [15].

From the perspective of big data, the value of VR/AR simulation training lies not only in being "repeatable, zero-risk, and free from time constraints," but also in the fine-grained operational data it naturally generates: instrument trajectories, force curves, viewpoint switching, operation duration, and error events can all be recorded in real time, thereby training machine learning models to automatically score and predict operation quality, making "data-driven formative assessment of surgical skills" possible. Domestic teaching teams have applied augmented reality (AR) and mixed reality (MR) technologies to orthopedic residency training and reported positive results [16]. VR technology and VR-based micro-lectures have also been used in surgical simulation teaching and theoretical teaching [17–18]. 3D-printed personalized fracture models have likewise been used in acetabular fracture teaching with positive results [19]. However, overall, such studies are still mainly single-center, small-sample designs.

2.4. Large language models assisting learning and assessment

Large language models (LLMs), represented by ChatGPT, are reshaping the forms of medical knowledge acquisition and assessment evaluation with powerful natural language understanding and generation capabilities. Kung et al. reported that ChatGPT reached a passing level on United States Medical Licensing Examination (USMLE)-style questions and systematically discussed its potential to assist medical education [20]. A scoping review by Gordon et al. (BEME Guide No. 84) included 278 articles, mapping the application landscape of artificial intelligence in medical education from admissions, teaching, and assessment to clinical reasoning training, and pointed out that evidence in this field is growing rapidly while rigorously designed studies remain few, urgently requiring ethical guidelines and standardized reporting frameworks [21]. In the assessment and evaluation link, Zhou Yuankai et al. introduced generative artificial intelligence into residency training theory and case analysis assessment, preliminarily verifying the feasibility of LLM-assisted question generation and evaluation [22].

At the orthopedic specialty level, Kung et al. evaluated ChatGPT's performance on the 2020–2022 actual questions (360 text-only questions without images) of the American Orthopedic In-Training Examination (OITE): GPT-3.5 answered 54.3% correctly, equivalent to the average level of first-year residents (PGY-1) nationwide; GPT-4 answered 73.6% correctly, reaching the average level of fifth-year (PGY-5) residents and exceeding the 67% passing line of Part I of the American Board of Orthopedic Surgery examination;

the study did not include image-dependent questions [23]. This not only indicates that LLMs already possess considerable knowledge reserves and can serve as auxiliary tools for after-class self-testing, knowledge point explanation, and case discussion, but also delineates their capability boundaries—strong in textual knowledge and weak in spatial imaging—so their positioning should be a learning partner rather than an answer source. Domestic scholars have reflected on how ChatGPT changes medical education models, arguing that it can be used for personalized teaching scenario design, clinical problem-solving ability training, and teaching research efficiency improvement, while vigilance is needed against risks of academic integrity and ability compensation [24]. Wang Di et al. introduced generative artificial intelligence into orthopedic clinical teaching, and a controlled study showed its potential in improving learning outcomes and teaching satisfaction [25].

For orthopedic residency training, potential application pathways of LLMs include: (1) individualized question banks and prompt-driven case deduction; (2) medical English and literature reading/writing training; (3) generation of drafts for lesson plans, question items, and comments to reduce faculty clerical burden; (4) serving as an "always-online" Q&A tutor after combining retrieval augmentation and specialized knowledge bases. The common prerequisite for the above applications is that residents possess sufficient prompt literacy and the ability to discriminate against model "hallucination" outputs.

2.5. Predictive analytics and training outcome management

Predictive applications of big data models in orthopedic clinical practice have become relatively mature. Lisacek-Kiosoglous et al. systematically reviewed the application of artificial intelligence in orthopedic surgery, showing its broad potential in imaging assessment, surgical planning, and outcome prediction [26]. Transferring this paradigm to the training field means modeling past training data to predict risks in residents' completion assessment performance and even future practice competency, achieving "early identification and early intervention." The study by Wang Xinbei et al. is precisely an initial practice in this direction in residency training management, and systematic reviews in the field of learning analytics also provide a framework for reference [7, 9]. At present, specialized prediction research on orthopedic residency training outcomes remains a blank, and its prerequisite is the construction of a multi-center, long-cycle training process database.

3. Existing challenges

3.1. Data level: Silos, quality, and privacy

The premise of big data models is high-quality, standardized, and shareable data. At present, residency training data management in China still exhibits "siloed" characteristics: platforms across different training bases are not interconnected, and data barriers exist among the National Residency Training Information Platform, hospital information systems, and picture archiving and communication systems (PACS). The quality of process data reporting varies considerably, and the authenticity of disease type and procedure registration is difficult to verify, with risks of retrospective data falsification that directly undermine the reliability of downstream modeling and analysis. Most platforms record only outcome indicators such as assessment scores, lacking process data that reflect learning behaviors and clinical decision-making processes, which places learning analytics models in a predicament of "insufficient features." In addition, orthopedic teaching involves large volumes of patient imaging and medical record data, and secondary use for teaching purposes must comply with medical ethics and personal information protection requirements, with high costs for de-identification and authorization. If trainee training profiles are improperly used, they may be distorted into monitoring and rating tools, bringing new ethical risks.

3.2. Algorithm level: Black box, bias, and hallucination

The "black box" nature of deep learning models makes it difficult to explain when and why errors occur, while medical and educational contexts are highly sensitive to errors. Algorithm audits by Oakden-Rayner et al. demonstrated that even fracture detection models with excellent overall performance can exhibit systematic missed detections in specific subgroups [12]. If such models are directly used for teaching demonstrations without disclosing failure modes, they may convey false confidence to resident physicians. Population bias in training data further amplifies unfairness: scoring models trained on data from senior physicians may systematically underestimate the performance of trainees with different learning styles. The "hallucination" problem of large language models (LLMs) poses a direct risk to knowledge-based teaching—the models can fluently generate seemingly authoritative but factually inaccurate explanations, with junior residents lacking critical appraisal skills being particularly vulnerable [20–21].

3.3. Educational level: Dependency risks and lagging evaluation systems

Technological embedding may inversely weaken the essence of education. First, over-reliance on intelligent tools may lead to "competency compensation": automatic image annotation may undermine independent interpretation training, and instant answers from LLMs may compress the processes of deep thinking and knowledge construction [24]. Second, the current residency training evaluation system has not yet incorporated artificial intelligence literacy into graduation requirements, and a significant gap remains between residents' ability to "use" AI and their ability to "use it well and critically." Third, digital literacy among faculty varies considerably, and most supervising physicians lack an understanding of the basic principles of machine learning, making it difficult for them to correctly demonstrate human–AI collaboration. Gordon et al. particularly emphasized that current research on AI in medical education urgently needs ethical guidelines and standardized reporting frameworks [21]. This is highly relevant to the goal of "homogenized training" in residency programs: improper use of intelligent tools may instead exacerbate quality disparities across training bases.

3.4. Resources and accessibility

The procurement and maintenance costs of immersive simulation equipment and intelligent teaching platforms are high. The deployment of VR devices, computing resources, and specialty datasets remains difficult to popularize in most prefecture-level training bases, potentially creating a new "digital divide." How to share high-quality digital teaching resources with county-level and grassroots collaborative residency training units at lower cost is a practical issue that policy design must address.

4. Countermeasures and prospects

First, strengthen top-level design and data governance. Relying on the national residency training management platform, promote the unification of data standards and interoperability, and establish teaching data usage specifications of "minimum necessity, tiered authorization, and full-process traceability." Encourage the construction of ethically approved multicenter teaching datasets for orthopedic residency training.

Second, incorporate artificial intelligence literacy into the curriculum system. In orientation education, include an overview of the principles of intelligent tools, their performance boundaries, and the teaching of failure cases, and take "critical use of artificial intelligence" as an observation point for formative assessment. Conduct tiered training for faculty to cultivate teaching staff who understand both orthopedics and data.

Third, adhere to the application positioning of human-machine collaboration and competency-based orientation. Intelligent imaging, VR scoring, and LLM question answering should all be positioned as tools for formative assessment and individualized feedback, serving core competencies such as clinical reasoning, surgical operation, and doctor-patient communication. Irreplaceable components such as bedside teaching, intraoperative instruction, and humanistic cultivation should not be diluted by technology.

Fourth, support dissemination decisions with rigorous research evidence. Current evidence is mainly based on single-center studies and short-term outcomes. Multicenter randomized controlled trials and long-term follow-up should be conducted, focusing on answering whether the integration of intelligent tools can improve graduation assessment and independent practice capability, which subgroups of trainees benefit or are harmed, and what the cost-effectiveness is.

Fifth, promote industry-university-research collaboration and inter-base assistance. Encourage residency training bases to jointly build teaching application scenarios with universities and enterprises, share VR courses, intelligent question banks, and teaching datasets through regional residency training alliances, and have national-level bases provide counterpart support to prefecture-level and municipal bases to narrow the digital divide.

5. Conclusion

Big data models are permeating the entire process of orthopedic residency training through pathways such as process management, imaging teaching, simulation training, knowledge question answering, and outcome prediction, demonstrating the potential to improve the homogeneity and individualization of training. Their implementation, however, remains constrained by multiple factors, including data governance, algorithm interpretability, evaluation systems, and resource accessibility. The future direction is not to let technology replace teachers, but to build a new orthopedic residency training ecosystem of "teacher-led, data-supported, and human-machine collaborative" through standardized data, transparent algorithms, and literacy-oriented curricula, ultimately serving the competency-oriented goal of cultivating orthopedic physicians.

Disclosure statement

The authors declare no conflict of interest.

References

[1] General Office of the State Council, 2017, Opinions of the General Office of the State Council on Deepening the Collaboration between Medicine and Education to Further Promote the Reform and Development of Medical Education: Guo Ban Fa [2017] No. 63. (2017-07-11). Accessed on September 8, 2026. http://www.moe.gov.cn/jyb_xxgk/moe_1777/moe_1778/201707/t20170711_309175.html

[2] General Office of the State Council, 2020, Guiding Opinions of the General Office of the State Council on Accelerating the Innovative Development of Medical Education: Guo Ban Fa [2020] No. 34 (2020-09-23). Accessed on September 8, 2026. https://www.gov.cn/zhengce/content/2020-09/23/content_5546104.htm
[3] Beam AL, Kohane IS, 2018, Big Data and Machine Learning in Health Care. JAMA, 319(13): 1317–1318. https://doi.org/10.1001/jama.2017.18391.
[4] Rajkomar A, Dean J, Kohane I, 2019, Machine Learning in Medicine. The New England Journal of Medicine, 380(14): 1347–1358. https://doi.org/10.1056/NEJMra1814259
[5] Topol EJ. High-performance Medicine: The Convergence of Human and Artificial Intelligence. Nature Medicine, 25(1): 44–56. https://doi.org/10.1038/s41591-018-0300-7
[6] Wartman SA, Combs CD, 2018, Medical Education Must Move from the Information Age to the Age of Artificial Intelligence. Academic Medicine, 93(8): 1107–1109. https://doi.org/10.1097/ACM.0000000000002044
[7] Wang XB, Zhou A, Yan XL, et al., 2024, Analysis of Influencing Factors on the Results of the Professional Theory Examination in the Graduation Assessment of Standardized Residency Training. Chinese Journal of Medical Education Research, 23(6): 835–840. https://doi.org/10.3760/cma.j.cn116021-20221229-01620
[8] Hua WB, Du YX, Deng FW, et al., 2024, Application of Competency-oriented Mini-CEX and DOPS in Standardized Residency Training of Orthopedics. Orthopedics, 15(3): 253–257. https://doi.org/10.3969/j.issn.1674-8573.2024.03.011
[9] Bojic I, Mammadova M, Ang CS, et al., 2023, Empowering Health Care Education through Learning Analytics: In-depth Scoping Review. Journal of Medical Internet Research, 2023(25): e41671. https://doi.org/10.2196/41671
[10] Kim DH, MacKinnon T, 2018, Artificial Intelligence in Fracture Detection: Transfer Learning from Deep Convolutional Neural Networks. Clinical Radiology, 73(5): 439–445. https://doi.org/10.1016/j.crad.2017.11.015
[11] Li S, Li KN, 2024, Research Progress of Artificial Intelligence in Orthopedic Diagnostic Techniques. Chinese Journal of Geriatric Orthopedics and Rehabilitation (Electronic Edition), 10(1): 46–50. https://doi.org/10.3877/cma.j.issn.2096-0263.2024.01.008
[12] Oakden-Rayner L, Gale W, Bonham TA, et al., 2022, Validation and Algorithmic Audit of a Deep Learning System for the Detection of Proximal Femoral Fractures in Patients in the Emergency Department: A Diagnostic Accuracy Study. The Lancet Digital Health, 4(5): e351–e358. https://doi.org/10.1016/S2589-7500(22)00004-8
[13] Lohre R, Bois AJ, Athwal GS, et al., 2020, Improved Complex Skill Acquisition by Immersive Virtual Reality Training: A Randomized Controlled Trial. The Journal of Bone and Joint Surgery, 102(6): e26. https://doi.org/10.2106/JBJS.19.00982
[14] Lohre R, Bois AJ, Athwal GS, et al., 2020, Effectiveness of Immersive Virtual Reality on Orthopedic Surgical Skills and Knowledge Acquisition among Senior Surgical Residents: A Randomized Clinical Trial. JAMA Network Open, 3(12): e2031217. https://doi.org/10.1001/jamanetworkopen.2020.31217
[15] Clarke E, 2021, Virtual Reality Simulation: The Future of Orthopedic Training? A Systematic Review and Narrative Analysis. Advances in Simulation (London, England), 6(1): 2. https://doi.org/10.1186/s41077-020-00153-x
[16] Zhu ZL, Gao XB, Guan JZ, 2022, Application of Augmented Reality and Mixed Reality Technology Based on “Internet+” in Standardized Residency Training of Orthopedics. Chinese Journal of General Practice, 20(6): 1052–1055. https://doi.org/10.16766/j.cnki.issn.1674-4152.002522
[17] Wang ZZ, Sun JC, Wang Y, et al., 2020, Application of Virtual Reality Technology in Standardized Residency Training of Orthopedics. Chinese Journal of Medical Education, 40(7): 557–560. https://doi.org/10.3760/cma.

j.cn115259-20191216-01087
[18] Guo SQ, Qiao B, Shui W, 2022, Application of Micro-lectures Based on Virtual Reality Technology in Standardized Residency Training of Orthopedics. Chinese Journal of Medical Education Research, 21(5): 592–595.
[19] Wang XP, Liu XC, Wu M, et al., 2023, Application of 3D Printing Technology in Clinical Teaching of Acetabular Fractures. Chinese Journal of General Practice, 21(10): 1770–1773. https://doi.org/10.16766/j.cnki.issn.1674-4152.003221
[20] Kung TH, Cheatham M, Medenilla A, et al., 2023, Performance of ChatGPT on USMLE: Potential for AI-assisted Medical Education Using Large Language Models. PLOS Digital Health, 2(2): e0000198. https://doi.org/10.1371/journal.pdig.0000198
[21] Gordon M, Daniel M, Ajiboye A, et al., 2024, A Scoping Review of Artificial Intelligence in Medical Education: BEME Guide No. 84. Medical Teacher, 46(4): 446–470. https://doi.org/10.1080/0142159X.2024.2314198
[22] Zhou YK, Liu P, Liu SJ, et al., 2026, Exploration of the Application of Generative Artificial Intelligence in the Assessment of Standardized Residency Training in Critical Care Medicine. Medical Journal of Peking Union Medical College Hospital, 17(1): 286–291. https://doi.org/10.12290/xhyxzz.2024-0739
[23] Kung JE, Marshall C, Gauthier C, et al., 2023, Evaluating ChatGPT Performance on the Orthopedic In-Training Examination. JB JS Open Access, 8(3): e23.00056. https://doi.org/10.2106/JBJS.OA.23.00056
[24] Qu X, Yang JM, Chen T, et al., 2023, Reflections on the Changes in Medical Education Models Brought by ChatGPT. Journal of Sichuan University (Medical Sciences), 54(5): 937–940. https://doi.org/10.12182/20231360302
[25] Wang D, Li DL, Luo ZJ, et al., 2025, Research on the Application of Generative Artificial Intelligence in Orthopedic Clinical Teaching. Orthopedics, 16(5): 431–435. https://doi.org/10.3969/j.issn.1674-8573.2025.05.008
[26] Lisacek-Kiosoglous AB, Mason AN, Dowling CN, et al., 2023, Artificial Intelligence in Orthopedic Surgery. Bone & Joint Research, 12(7): 447–454. https://doi.org/10.1302/2046-3758.127.BJR-2023-0111.R1