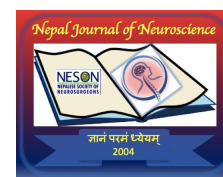


# Large Language Models in Spine Surgery: A Clinical Decision-Making Framework for the Next Decade with Emphasis on Degenerative Spine Care and LMIC Applications

Swaminathan Ganesh<sup>1</sup>, Jeena Joseph<sup>1</sup>

<sup>1</sup>Department of Neurosurgery, Christian Medical College and Hospital, Vellore, Tamil Nadu, India



Date of Submission: 10<sup>th</sup> May 2026

Date of Acceptance: 3<sup>rd</sup> June 2026

Date of Publication: 15<sup>th</sup> June 2026

## Abstract

Degenerative spine disorders are a leading cause of disability worldwide and impose a growing burden on health systems, particularly in low- and middle-income countries. Large language models are emerging as powerful tools capable of supporting clinical decision-making, synthesising complex evidence, generating patient-specific explanations, and improving clinical documentation. Their ability to process free-text clinical narratives and integrate multiple sources of information makes them particularly suited to degenerative spine care, where decision-making requires the integration of symptoms, neurological findings, imaging, and patient preferences.

This narrative review examines the evolving role of artificial intelligence in spine surgery, with particular emphasis on large language models in degenerative conditions of the cervical and lumbar spine. We describe current and emerging applications across clinical triage, radiological interpretation, guideline synthesis, patient communication, and workflow optimisation. Building on these insights, we propose a structured six-level clinical decision-making framework spanning initial patient contact to postoperative care.

We also discuss key ethical, medico-legal, and governance considerations relevant to the safe implementation of these technologies, particularly in resource-constrained environments. Large language models are unlikely to replace clinical judgment; however, when integrated within structured workflows and appropriate safety systems, they have the potential to enhance the quality, efficiency, and equity of degenerative spine care globally over the coming decade.

## Introduction

Low back pain remains the leading cause of years lived with disability worldwide, with a growing burden driven by degenerative spine conditions such as lumbar disc herniation, lumbar canal stenosis, degenerative spondylolisthesis, and cervical spondylotic myelopathy<sup>1,2</sup>. Clinical decision-making in these disorders is complex, requiring integration of symptoms, neurological findings, imaging, comorbidities, and patient goals, often under time constraints.

Artificial intelligence (AI) and machine learning (ML) approaches are increasingly explored in spine surgery,

particularly for outcome prediction, risk stratification, and imaging analysis<sup>3-6</sup>. However, these models are typically task-specific, dependent on structured inputs, and limited in their ability to process the free-text clinical narratives that dominate routine spine practice.

Large language models (LLM) represent a significant evolution in artificial intelligence, with the ability to process natural language, integrate diverse sources of information, and perform complex reasoning tasks<sup>7-10</sup>. Early studies suggest potential applications in clinical decision support, patient education, and workflow optimisation, but guidance on their integration into degenerative spine care—particularly in resource-constrained settings—remains limited<sup>11-16</sup>. This narrative review examines the role of LLMs in degenerative spine care, current and emerging applications, a clinically oriented decision-making framework, and key ethical and governance considerations<sup>17-27</sup>.

## From Traditional Machine Learning to Large Language Models in Spine Care

Early AI work in spine surgery was dominated by conventional machine learning (ML) approaches. Using curated datasets with predefined variables (e.g., age, comorbidities, radiographic parameters, baseline ODI or VAS), models such as logistic regression, random forests, gradient boosting, and neural networks have been applied to predict postoperative improvement, estimate complications such as surgical site infection or non-union, and classify imaging findings in scoliosis

### Access this article online

Website: <https://www.nepjol.info/index.php/NJN>

DOI: <https://doi.org/10.3126/njn.v23i2.91054>



### HOW TO CITE

Ganesh, S., & Joseph, J. Large Language Models in Spine Surgery: A Clinical Decision-Making Framework for the Next Decade with Emphasis on Degenerative Spine Care and Applications in Low- and Middle-Income Countries. *NJNS*. 2026;23(2):4-9

### Address for correspondence:

Dr. Swaminathan Ganesh  
Department of Neurosurgery Christian Medical College and Hospital  
Vellore, Tamil Nadu – 632517, India  
Email: glioma1105@gmail.com

Copyright © 2023 Nepalese Society of Neurosurgeons (NESON)  
ISSN: 1813-1948 (Print), 1813-1956 (Online)



This work is licensed under a Creative Commons Attribution-Non Commercial 4.0 International License.

and degenerative disease<sup>3-6</sup>.

Lopez et al. found that most AI applications in spine surgery focused on patient selection and outcome prediction after fusion, decompression, or tumour surgery<sup>3</sup>. Subsequent reviews have reported good discrimination for functional outcomes and complication risks but highlighted limitations in generalisability, external validation, and interpretability<sup>4</sup>. Prognostic models have also been developed to predict pain and disability trajectories following lumbar disc herniation surgery<sup>6</sup>.

Despite their promise, conventional ML models generally require structured, labelled datasets, are designed for narrow tasks, and cannot directly process the free-text narratives common in clinic notes, radiology reports, and patient communication.

By contrast, large language models (LLMs) are transformer-based systems trained on large text corpora, including biomedical literature and clinical-like documents<sup>7-10</sup>. Rather than structured prediction alone, LLMs are designed to process language, enabling tasks such as clinical triage, explanation, summarisation, and shared decision-making support. In spine care, an LLM may integrate symptoms, imaging reports, and guidelines to generate a structured explanation of diagnoses and management options, while traditional ML remains better suited for numeric prediction tasks such as outcome estimation. Table 1.

**Table 1.** Traditional machine learning (ML) versus large language models (LLMs) in degenerative spine

Feature	Traditional ML models	Large language models
Typical inputs	Structured numerical or categorical variables (age, ODI, comorbidities, radiographic angles)	Free-text symptoms, examination notes, radiology reports, guidelines, patient questions
Output type	Probabilities, classifications, risk scores	Narrative explanations, recommendations, summaries, counselling text
Task scope	Narrow, task-specific (e.g. predict ODI improvement)	Broad, multi-task (triage, education, documentation)
Data requirements	Curated, labelled datasets	Large text corpora; less need for task-specific labels
Suitability for degenerative spine	Strong for outcome prediction and risk stratification	Strong for triage, explanation, shared decision-making
Suitability in LMICs	Limited by need for high-quality labelled datasets	Potentially high, as models can be used with minimal local structured data but require governance and adaptation

### Current and Emerging Applications of LLMs in Degenerative Spine Care

Clinical decision support in common degenerative scenarios  
Early studies have evaluated LLM performance in spine-related clinical decision-making. Almekkawi et al. found that Chat GPT and Claude could approximate spine surgeons' decisions in selected surgical and radiological scenarios, albeit with

variability and occasional clinically important errors<sup>14</sup>. Lang et al. suggested that Chat GPT may support diagnostic and treatment decisions in lumbar spinal stenosis, particularly in summarising conservative options and surgical indications<sup>15</sup>. Similar studies have reported promising performance while emphasising the need for expert oversight<sup>16</sup>.

In routine degenerative spine care, LLMs may assist by structuring and cross-checking decision pathways in conditions such as lumbar disc herniation, lumbar stenosis, degenerative spondylolisthesis, and cervical myelopathy. They can summarise red flags, correlate symptoms with imaging findings, and generate guideline-linked management narratives to support—rather than replace—clinical judgment, particularly under time pressure or in training environments.

#### Literature synthesis and guideline interpretation

LLMs can summarise biomedical literature, assist in evidence synthesis, and generate responses to clinical queries<sup>9,13,21</sup>. In degenerative spine care, this may support rapid appraisal of evidence for procedures such as decompression versus fusion, minimally invasive approaches, or non-operative therapies. Retrieval-augmented generation (RAG) in which the model accesses curated guidelines or literature repositories (e.g., NASS or AO Spine recommendations), is particularly valuable because it reduces hallucinations and improves traceability of recommendations<sup>7,9,17</sup>.

#### Patient education and shared decision-making

Negrini et al. reported that Chat GPT-generated patient education material for idiopathic scoliosis was generally clear and accessible but variable in scientific accuracy, supporting its use as an adjunct rather than a replacement for expert-reviewed material<sup>11,13</sup>. In degenerative spine care, LLMs may generate explanations of lumbar stenosis, disc herniation, or cervical myelopathy at different literacy levels and in local languages, potentially improving informed consent and patient understanding in high-volume LMIC settings.

#### Documentation and workflow optimisation

LLMs are increasingly used to support clinical documentation, including consultation summaries, operative notes, discharge summaries, referral letters, and follow-up templates<sup>20</sup>. Reviews consistently identify administrative support as a high-yield early application<sup>12,13,15,28</sup>. In LMIC centres, such efficiencies may reduce administrative burden and allow more time for complex decision-making and teaching.

### Technical Foundations Relevant to Spine Surgeons

Spine surgeons do not need detailed expertise in deep learning to use LLMs safely, but familiarity with key concepts is useful. LLMs are transformer-based systems trained on large text corpora and subsequently refined to better follow user instructions<sup>7,10</sup>. Models are pretrained on large corpora and then often “instruction-tuned” on curated datasets to better follow user prompts.

A clinically important concept is retrieval-augmented generation (RAG), in which LLMs access external, up-to-date sources such as guidelines or institutional protocols to supplement internal training knowledge<sup>7,9,17</sup>. This is particularly relevant in degenerative spine care, where evolving evidence may influence indications for fusion or minimally invasive

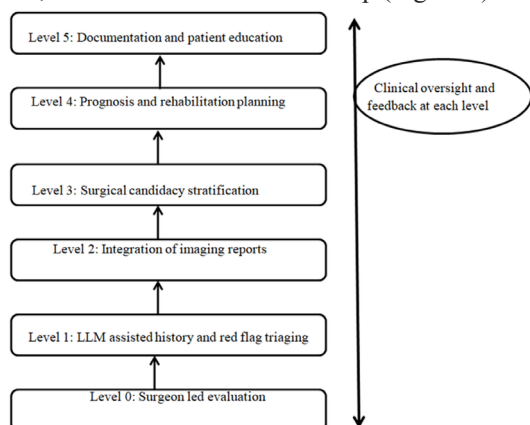
procedures. Clinicians must also be aware of typical LLM failure modes. These include hallucinations (confident but incorrect statements), inconsistency across similar prompts, sensitivity to superficial prompt changes and embedding of biases present in training data<sup>7,10,17,18</sup>. Recent studies have also highlighted how non-clinical features of a patient’s message (e.g., style, grammar, emotional tone) can inappropriately influence LLM outputs, potentially leading to unsafe advice if not supervised. These behaviours underscore the importance of using LLM outputs as decision support, not as autonomous decision-makers.

**Table 2.** Key risks of LLM use in spine care and potential mitigation strategies

Risk	Description	Potential mitigation
Hallucination	Confident, incorrect statements not supported by evidence	Use RAG with curated guideline repositories; mandate human verification
Automation bias	Over-reliance by clinicians, especially trainees	Training on AI literacy; clear institutional policies on LLM role
Bias and inequity	Reduced performance in under-represented populations	Ongoing evaluation in LMIC cohorts; fine-tuning with local data where feasible
Privacy breaches	Inadequate de-identification or insecure data transfer	Use on-premise or regionally compliant models; strict data governance
Regulatory uncertainty	Lack of clear liability and approval pathways	Align with emerging WHO and national guidance; integrate into hospital AI governance structures

### An LLM-Enhanced Clinical Decision-Making Framework for Degenerative Spine Care

To structure the safe integration of LLMs into spine practice, we propose a six-level framework centred on degenerative conditions, from first contact to follow-up (Figure 1).



**Figure 1.** LLM-enhanced degenerative spine care pathway. The framework illustrates six levels of integration, from surgeon-led evaluation (Level 0) to LLM-assisted history and red-flag triage (Level 1), integration of imaging reports (Level 2), surgical candidacy stratification with evidence summaries (Level 3), prognosis and rehabilitation planning (Level 4), and documentation and patient education (Level 5). Clinician oversight and feedback operate at all levels.

### Level 0 – Surgeon-led baseline

Clinical judgment remains the foundation of spine care. All LLM use is optional, adjunctive and subject to clinician oversight.

### Level 1 – History triage and red-flag screening

At first contact, particularly in primary care or emergency triage settings, LLMs can help structure patient narratives. A clinician or trained assistant enters the patient’s free-text symptoms, which the LLM converts into a structured summary: onset and duration of back or neck pain, radicular distribution, aggravating and relieving factors, red-flag features (recent trauma, fever, weight loss, sphincter dysfunction), and myelopathic symptoms. In LMIC settings where non-specialists often perform initial assessments, such structured outputs may reduce missed red flags and support timely referral<sup>17-20</sup>.

### Level 2 – Integration of imaging reports

Once imaging is available, LLMs can parse MRI or CT reports, extracting key descriptors such as the level and severity of central canal or foraminal stenosis, presence of disc herniation, translation or angulation in spondylolisthesis, and degree of cord compression or signal change. The model can then align these findings with the clinical picture summarised at Level 1. For example, it may highlight that severe L4–L5 stenosis with classic neurogenic claudication symptoms is concordant, whereas multi-level degenerative changes with predominantly non-specific low back pain suggests a more conservative trajectory. This level is particularly useful in high-volume clinics, where rapidly synthesising multiple reports is challenging.

### Level 3 – Surgical candidacy stratification

At this stage, the surgeon remains the decision-maker but can request an LLM-generated evidence summary tailored to the patient’s profile. For a 60-year-old with single-level degenerative spondylolisthesis, predominant leg pain and failure of a six-month conservative trial, the system might summarise key trials comparing decompression alone versus decompression with fusion, outline expected outcomes and risk profiles, and highlight factors that tilt the balance towards one option or another. The LLM thus acts as an on-demand, context-aware literature assistant, particularly helpful in settings where ready access to full-text articles is limited.

### Level 4–Prognosis, rehabilitation and return-to-work counselling

For patients undergoing surgery or long-term conservative care, LLMs can generate recovery timelines, highlight risk factors that may delay return to work or baseline function, and provide a first draft of a rehabilitation plan that physiotherapist can refine. For example, after lumbar decompression without fusion, the model can suggest typical milestones for pain reduction, walking tolerance and lifting restrictions, adapted to the patient’s occupation. In LMIC contexts where rehabilitation services may be fragmented, such templated plans, suitably adapted by local clinicians, may support more consistent post-operative counselling.

### Level 5 – Documentation, education and audit

Finally, LLMs can assist with generating operative notes, discharge summaries, follow-up letters and patient education leaflets, based on structured inputs from the surgeon. Over

time, these outputs can be standardised and audited for quality and consistency. In teaching hospitals, LLMs may help junior trainees draft notes that senior surgeons can quickly validate, improving both efficiency and training.

### LMIC-Specific Opportunities and Challenges

Spine care in LMICs is characterised by late presentation, limited access to imaging and specialist surgeons, and constrained operating capacity<sup>17-19</sup>. Recent narratives on global spine surgery emphasise bottlenecks in “time to spine” pathways, where administrative and logistical delays compound the effects of scarce resources<sup>20</sup>. Endoscopic and minimally invasive spine techniques are particularly challenging to implement, given shortages in equipment and training<sup>18</sup>.

Within this context, LLMs offer several specific advantages. First, they can support non-specialist clinicians in primary and secondary care by structuring back pain assessments, identifying red flags and suggesting when to refer, thereby improving triage quality without requiring immediate specialist involvement. Second, they can help prioritise which patients most urgently need imaging or surgical consultation, by combining symptom profiles and basic examination findings into risk-focused narratives. Third, they can produce patient education materials in local languages and at appropriate literacy levels, potentially improving adherence to conservative measures and post-operative instructions.

However, LMIC deployment also raises particular challenges. Internet bandwidth may be unreliable, making reliance on large cloud-hosted models impractical. Data protection frameworks may be less mature, heightening concerns about transmitting sensitive clinical information to external servers. Locally hosted or regionally federated LLMs, potentially fine-tuned on anonymised local data, may therefore be preferable for long-term sustainability.

### Ethics, Governance and Regulatory Considerations

Ethical concerns regarding AI in health—including transparency, accountability, equity, privacy and the need for human oversight—have been extensively discussed in the general literature<sup>23-27</sup>. The World Health Organization and other bodies have issued guidance for the use of AI and, more recently, large multimodal models in healthcare, emphasising that these systems must not undermine human autonomy, must be developed inclusively, and must be subject to rigorous evaluation before clinical deployment<sup>23,24</sup>. National frameworks, including those from agencies such as the Indian Council of Medical Research, echo these principles and propose governance mechanisms for AI tools used in clinical care<sup>27</sup>.

In the context of LLMs for degenerative spine care, several specific issues warrant consideration:

1. Over-reliance and automation bias. There is a risk that clinicians, particularly trainees, may over-trust LLM outputs. Institutional policies should explicitly state that LLMs provide decision support, not definitive clinical decisions.
2. Data protection and consent. Institutions need clear policies on what patient data can be processed by external AI services, how it is de-identified, and whether explicit patient consent is required.

3. Transparency and auditability. Any AI system that influences clinical care should provide logs of inputs and outputs, enabling retrospective review and quality assurance.
4. Equity. LLMs trained primarily on data from high-income settings may embed biases that reduce performance in LMIC populations. Ongoing evaluation in diverse contexts is mandatory.
5. Liability. Regulatory frameworks will need to address whether responsibility for AI-related errors rests with clinicians, institutions or vendors, and under what conditions.

Adhering to emerging international guidelines on responsible AI, and integrating LLM deployment into existing clinical governance structures, will be essential to harness benefits while minimising harm.

### Future Directions

Looking ahead, several developments are likely to shape the role of LLMs in degenerative spine surgery. Multimodal models capable of directly ingesting imaging data alongside text may allow more nuanced integration of MRI/CT findings with clinical narratives. Digital “spine twins” that combine biomechanical simulation with LLM-based reasoning could support personalised planning for complex deformity or revision surgery. On the operational side, hospital-tuned LLMs with embedded local guidelines and EMR integration may provide real-time decision support at the point of care.

At the same time, rigorous prospective evaluation is needed. Recent systematic reviews of LLMs in medicine highlight the heterogeneity of evaluation methods and the paucity of high-quality, task-specific validation studies<sup>12,17</sup>. Spine-specific benchmarks for scenarios such as lumbar stenosis triage, myelopathy detection, or postoperative complication counselling would allow more meaningful comparison of models and safer deployment.

## Conclusion

LLMs represent a major advance in AI capabilities, particularly suited to the language-intensive tasks that permeate degenerative spine care. They can assist with triage, radiology report interpretation, guideline synthesis, patient education and documentation, and their potential impact may be greatest in high-volume, resource-constrained LMIC settings. However, these tools are not a replacement for clinical judgment and must be embedded within carefully designed workflows, supported by robust ethical and governance frameworks. The six-level clinical decision-making framework proposed here offers one possible structure for staged implementation. As evidence accumulates and regulatory guidance matures, spine surgeons should be active partners in shaping how LLMs are used—ensuring that their adoption enhances, rather than undermines, the quality and equity of spine care worldwide.

### Acknowledgement

None.

### Funding

None.

## Disclosures

None.

## Conflict of interest

The authors declare no conflict of interest.

## References

1. Ferreira ML, De Luca K, Haile LM, et al. Global, regional, and national burden of low back pain, 1990–2020, its attributable risk factors, and projections to 2050: a systematic analysis of the Global Burden of Disease Study 2021. *The Lancet Rheumatology*. 2023;5(6):e316–e329. doi:10.1016/S2665-9913(23)00098-X
2. Li Y, Zou C, Guo W, et al. Global burden of low back pain and its attributable risk factors from 1990 to 2021: a comprehensive analysis from the global burden of disease study 2021. *Front Public Health*. 2024;12:1480779. doi:10.3389/fpubh.2024.1480779
3. Lopez CD, Boddapati V, Lombardi JM, et al. Artificial Learning and Machine Learning Applications in Spine Surgery: A Systematic Review. *Global Spine Journal*. 2022;12(7):1561-1572. doi:10.1177/21925682211049164
4. Han H, Li R, Fu D, et al. Revolutionizing spinal interventions: a systematic review of artificial intelligence technology applications in contemporary surgery. *BMC Surg*. 2024;24(1):345. doi:10.1186/s12893-024-02646-2
5. Benítez-Andrades JA, Prada-García C, Ordás-Reyes N, Blanco ME, Merayo A, Serrano-García A. Enhanced prediction of spine surgery outcomes using advanced machine learning techniques and oversampling methods. *Health Inf Sci Syst*. 2025;13(1):24. doi:10.1007/s13755-025-00343-9
6. Berg B, Gorosito MA, Fjeld O, et al. Machine Learning Models for Predicting Disability and Pain Following Lumbar Disc Herniation Surgery. *JAMA Netw Open*. 2024;7(2):e2355024. doi:10.1001/jamanetworkopen.2023.55024
7. Thirunavukarasu AJ, Ting DSJ, Elangovan K, Gutierrez L, Tan TF, Ting DSW. Large language models in medicine. *Nat Med*. 2023;29(8):1930-1940. doi:10.1038/s41591-023-02448-8
8. Meng X, Yan X, Zhang K, et al. The application of large language models in medicine: A scoping review. *iScience*. 2024;27(5):109713. doi:10.1016/j.isci.2024.109713
9. Nazi ZA, Peng W. Large Language Models in Healthcare and Medical Domain: A Review. *Informatics*. 2024;11(3):57. doi:10.3390/informatics11030057
10. Omiye JA, Gui H, Rezaei SJ, Zou J, Daneshjou R. Large Language Models in Medicine: The Potentials and Pitfalls: A Narrative Review. *Ann Intern Med*. 2024;177(2):210-220. doi:10.7326/M23-2772
11. Negrini F, Malfitano C, Ferriero G, et al. Evaluating ChatGPT-4.0's accuracy and potential in idiopathic scoliosis conservative treatment: a preliminary study on clarity, validity, and expert perceptions. *Eur Spine J*. Published online July 21, 2025. doi:10.1007/s00586-025-09166-4
12. Shool S, Adimi S, Saboori Amleshi R, Bitaraf E, Golpira R, Tara M. A systematic review of large language model (LLM) evaluations in clinical medicine. *BMC Med Inform Decis Mak*. 2025;25(1):117. doi:10.1186/s12911-025-02954-4
13. Yapar A, Yapar D, Tokur Sonuvar E. Letter to the Editor concerning “Evaluating ChatGPT-4.0's accuracy and potential in idiopathic scoliosis conservative treatment: a preliminary study on clarity, validity, and expert perceptions” by negrini F, et al. (*Eur Spine J* [2025]: <https://doi.org/10.1007/s00586-025-09166-4>). *Eur Spine J*. Published online September 11, 2025. doi:10.1007/s00586-025-09347-1
14. Almekki AK, Caruso JP, Anand S, et al. Comparative Analysis of Large Language Models and Spine Surgeons in Surgical Decision-Making and Radiological Assessment for Spine Pathologies. *World Neurosurgery*. 2025;194:123531. doi:10.1016/j.wneu.2024.11.114
15. Lang SP, Yoseph ET, Gonzalez-Suarez AD, et al. Analyzing Large Language Models' Responses to Common Lumbar Spine Fusion Surgery Questions: A Comparison Between ChatGPT and Bard. *Neurospine*. 2024;21(2):633-641. doi:10.14245/ns.2448098.049
16. Li H, Zeng C, Miao L, Wang Y, Xia J, He D. Evaluation of the performance of large language models in endoscopic lumbar surgery: a comparative analysis. *Annals of Medicine & Surgery*. 2025;87(8):4835-4840. doi:10.1097/MS9.00000000000003519
17. Chikani MC, Mesi M, Okwunodulu O, et al. Changing Trends and Challenges of Spine Surgery in a Developing Country. *World Neurosurgery*. 2019;130:e815-e821. doi:10.1016/j.wneu.2019.06.229
18. Khalafallah AM, Diez S, Di L, et al. Are There Unique Barriers and Opportunities for Access to Endoscopic Spine Surgery in Low-Income Countries? A Narrative Review. *JCM*. 2025;14(11):3876. doi:10.3390/jcm14113876
19. Avetisian H, Karakash W, Wang JC. The next decade in spine surgery: Opportunities, challenges, and global synergy. *Spine Research*. 2025;1(1):3-6. doi:10.1097/br9.0000000000000006
20. Pedro KM, Alvi MA, Fehlings MG. Obstacles in “Time to Spine”: Challenges for the Timely Delivery of Acute Surgical Care for Patients with Traumatic and Non-Traumatic Spinal Cord Injury. *Healthcare*. 2024;12(22):2222. doi:10.3390/healthcare12222222
21. Wu J, Ma Y, Wang J, Xiao M. The Application of ChatGPT in Medicine: A Scoping Review and Bibliometric Analysis. *JMDH*. 2024;Volume 17:1681-1692. doi:10.2147/JMDH.

22. Zhang K, Meng X, Yan X, et al. Revolutionizing Health Care: The Transformative Impact of Large Language Models in Medicine. *J Med Internet Res.* 2025;27:e59069. doi:10.2196/59069
23. World Health Organization. Ethics and governance of artificial intelligence for health [Internet]. Geneva: WHO; 2021 [cited 2026 Mar 28]. Available from: <https://www.who.int/publications/i/item/9789240029200>.
24. World Health Organization. Ethics and governance of artificial intelligence for health: large multi-modal models [Internet]. Geneva: WHO; 2024 [cited 2026 Mar 28]. Available from: <https://www.who.int/publications/i/item/9789240084759>
25. Saenz AD, Mass General Brigham AI Governance Committee, McCoy T, et al. Establishing responsible use of AI guidelines: a comprehensive case study for healthcare institutions. *npj Digit Med.* 2024;7(1):348. doi:10.1038/s41746-024-01300-8
26. Ganbaatar U. Do Ethics in AI Still Matter? A Review of the 2021 UNESCO Recommendation on the Ethics of AI. *The Review of Faith & International Affairs.* 2025;23(3):26-33.
27. Indian Council of Medical Research. Ethical guidelines for application of artificial intelligence in biomedical research and healthcare [Internet]. New Delhi: ICMR; 2017 [cited 2026 Mar 28]. Available from: [https://main.icmr.nic.in/sites/default/files/guidelines/ICMR\\_AI\\_Guidelines\\_2017.pdf](https://main.icmr.nic.in/sites/default/files/guidelines/ICMR_AI_Guidelines_2017.pdf)
28. Zhang K, Meng X, Yan X, et al. Revolutionizing Health Care: The Transformative Impact of Large Language Models in Medicine. *J Med Internet Res.* 2025;27:e59069. doi:10.2196/59069