Journal of Institute of Medicine Nepal
Institute of Medicine, Kathmandu, Nepal

Review Article

JIOM Nepal. 2025 Dec;47(3):1-5.

Artificial Intelligence and Big Data in Medicine: A Nepali Medical Institute Perspective

Mohan Raj Sharma

Author(s) affiliation

Department of Neurosurgery, Maharajgunj Medical Campus, Tribhuvan University Teaching Hospital, Institute of Medicine, Kathmandu, Nepal

Corresponding author

Mohan Raj Sharma, MBBS, MS
mohanrajsharma@iom.edu.np

DOI

10.59779/jiomnepal.1456

Submitted

Dec 6, 2025

Accepted

Dec 21, 2025

ABSTRACT

The integration of big data and artificial intelligence (AI) has been changing the three facets of medicine namely, medical care, medical education and research especially in the last 10 years. By offering the modeling and rapid diagnostic and evaluating facilities and offering personalized approaches they have been reshaping the medical care. These technologies strengthen biomedical research by providing better predictive modelling and optimizing clinical trials. This, in turn, helps in the evidence-based approach to medicine. Similarly, they have the potential to transform medical education as to how the course is designed, how the contents are delivered, and how we assess the competency of the trainees. However, a major concern is the responsible use of these technologies and data privacy. This article briefly discusses the application of these technologies and guidelines regarding their responsible use.

Keywords

Artificial intelligence; big data; medical research; medical education; Nepal

INTRODUCTION

Clinical care, medical education and biomedical research are undergoing a profound transformation in the 21st century. Now there is unprecedented availability of large volumes of data, better data analytics, and personalized medical care. These are mainly due to the arrival of two technologies Big Data and Artificial Intelligence (AI).

The concept of big data arose in late 1990s and refers to datasets that are too large to be processed by currently available data processing software.[1] Sources of big data in healthcare include:

- Electronic health records (EHRs)
- Medical imaging
- Pathology and laboratory works
- Genomic and omics such as whole genome sequencing, proteomics, and metabolomics
- Wearables and sensors such as those for heart rate, glucose, activity monitoring, etc.
- Public health and social data such as disease registries, population health surveys, social media, etc.

Simultaneously, AI has emerged as a major force, capable of providing a human-like adjunct in the scientific community with a strong application in the medical field. It not only helps in making highly accurate predictions in healthcare, but also of generating synthetic data especially useful in rare diseases research supporting informed human decision-making. Together, big data and AI, work synergistically transforming the three pillars of medical science, viz. health care, medical education and research.[2] For medical institutes, particularly in low- and middle-income countries, these tools offer opportunities to accelerate the growth of medicine often at a low cost.[3–5] However, several pertinent issues particularly reliability, data privacy, and ethical use remain contentious. All stakeholders must embrace these changes and make the transition carefully, making sure that these advancements enhance patient care, encourage high-end medical research, and improve educational quality, taking into consideration the challenges associated with their implementation.

BIG DATA

As described earlier, big data refers to datasets that are too large and complex, and rapidly generated so that traditional data processing software is unable to capture, manage, and analyze them. This is best described as "5 Vs"[1,6]:

1. **Volume**: This refers to the very large volume of data. These datasets often range from terabytes to petabytes, making traditional storage and processing methods just inadequate. Managing such massive volumes requires advanced infrastructure like cloud computing and distributed databases.
2. **Velocity**: This indicates the speed at which new data is produced, collected, and analyzed. In healthcare, this includes real-time patient monitoring data available through devices in the intensive care unit settings, streaming data from wearable sensors, and rapid updates of public health surveillance systems.
3. **Variety**: This refers to the variety of data available in healthcare settings :
   - Structured data that fits neatly into databases such as laboratory tests, billing codes, and demographic information.
   - Semi-structured data such as genomic sequences and biosensor outputs that do not fit rigid formats.
   - Unstructured data like physicians' notes, radiology and pathology images, and audio or video recordings from clinical encounters.

   This diversity necessitates sophisticated tools such as natural language processing (NLP) and machine learning algorithms for effective analysis and interpretation.
4. **Veracity**: This refers to the trustworthiness and quality of the data. Healthcare data needs to be of highest quality as low-quality data leads to suboptimal care but they often contain inaccuracies due to human error in documentation, missing values, inconsistencies across institutions, or biases in data collection. Ensuring veracity involves applying rigorous data-cleaning techniques, validating algorithms, and addressing AI hallucinations in AI-driven decision support systems.
5. **Value**: This refers to the extent to which the data and its analysis support decision-making.

Use Of Big Data

Big data has enabled several advancements in medical research, clinical practice, and public health.

- **Precision medicine**. This refers to personalized care of individual patients both preventive and curative. By combining genomic, proteomic, imaging and clinical data, big data helps identify patient subgroups defined by specific mutations, biomarker profiles, or molecular signatures. Those subgroups let clinicians select targeted treatments, tailor drug dosing based on pharmacogenomics, and stratify patients for clinical trials so that therapies are tested in the people most likely to benefit.[7,8]
- **Epidemiology and surveillance**. Real-time and near-real-time data streams such as EHRs, lab reports, syndromic surveillance, mobile/wearable telemetry let public-health teams detect outbreaks, track geographic spread, and monitor disease severity faster than traditional paper-based reporting. This supports early warnings, more efficient testing, and judicious allocation of healthcare resources when resource availability is an issue.[9]
- **Health services research**. Large linked datasets (claims, EHRs, registries, administrative and operational data) make it possible to study patterns of care, utilization, cost indicators, and outcomes across populations. This allows the researchers evaluate policy interventions with much larger sample sizes that leads greater generalizability and robust recomendations.[10]
- **Drug discovery**. Analyzing molecular databases and high-throughput screening results helps identify new drug targets and biomarkers, predict how drugs interact with their targets, and find opportunities to repurpose existing drugs. Machine learning can help prioritize compounds, combine genetic and clinical data, and reduce late-stage failures by predicting safety and effectiveness earlier in development process.[8]

ARTIFICIAL INTELLIGENCE

Computational systems that can perform tasks typically requiring human-like intelligence are referred to as AI. Subfields of AI relevant to the medical field include:

- **Machine learning (ML)**: Algorithms that learn from data. Three subtypes are included here:
  - **Supervised Learning**: Disease diagnosis (e.g., diabetes, cancer risk prediction), outcome prediction (mortality, readmission risk), etc.
  - **Unsupervised Learning**: Patient clustering and phenotyping, discovery of disease subtypes, etc.
  - **Reinforcement Learning**: Adaptive decision-making in dynamic environments. The application is mainly in treatment optimization (radiotherapy dosing, ICU decision support, etc.) and personalized therapy pathways
- **Deep learning (DL)**: Neural networks that excel at image and signal recognition
- **Natural language processing (NLP)**: Extraction of meaning from clinical notes and literature

Applications in Clinical Medicine and Research

This has wide-ranging applications in clinical medicine and research.[8,11,12] In the field of diagnostics, deep learning–based approaches can detect neoplastic lesions on medical imaging with performance comparable to, and occasionally surpassing, that of experienced clinicians. In prognostics, machine learning models are used to predict disease progression and responses to treatment. Within clinical trials, AI can optimize patient recruitment and continuously monitor safety outcomes. Additionally, in drug discovery, AI accelerates molecular screening processes and facilitates the repurposing of existing drugs.

Applications in Medical Education

There are multiple applications of AI in medical education.[13–15] Adaptive learning platforms personalize educational content according to individual learner's learning pace and needs. Automated assessment systems analyze learners' clinical reasoning, communication abilities, and procedural skills. Virtual patients and simulation-based tools have allowed learners to practice in realistic environments minimizing risks to the patients. All these help in implementing the competency based medical education in the medical schools effectively. In addition, predictive analytics can identify learners at risk of underperformance, enabling timely academic support and remediation.

Six Principles Recommended by World Health Organization

To ensure that AI serves the public interest and minimizes risks and misuse, World Health Organization (WHO) outlines six guiding principles regarding the design and use of AI in health[16]:

1. **Protect human autonomy**: This principle states that humans should be in control of health decisions regarding privacy and confidentiality of data to ensure valid informed consent.
2. **Promote well-being and safety**: Regulatory standards for safety, accuracy, and efficacy, must be met while using AI with quality control mechanisms in practice.
3. **Ensure transparency and explainability**: All stakeholders are recommended to make AI design and deployment information accessible for public consultation and understanding.
4. **Ensure responsibility and accountability**: Stakeholders remain accountable for proper use, with mechanisms in place for remedial measures if harm occurs.
5. **Ensure inclusiveness and equity**: AI should promote equitable access to all regardless of demographics or social factors.
6. **Promote responsiveness and sustainability**: Designers, developers and users should continuously evaluate AI performance, minimize environmental impact, and prepare health workers for workplace changes.

These principles are indispensable for countries to adopt AI for use in healthcare settings.

Synergistic Effects of the Combination of Big Data and Artificial Intelligence

The combination of big data and AI creates a powerful synergy in which each strengthens the other. Big data provides large, diverse, and real-world datasets, while AI supplies advanced analytical tools to derive meaningful patterns and conclusions from these data. Together, they enable more accurate predictions, faster decision-making, and automation of complex tasks.[17] For example, AI applied to large genomic datasets has identified novel disease pathways in cancer and neurodegeneration.[18] Integration of omics, imaging, and clinical data offers a hitherto understanding of mechanisms of disease causation. Clinical trials are costly and time-consuming. Big data and AI improve trial efficiency by enabling better patient recruitment and stratification. Especially in rare diseases, generative AI can create synthetic control arms using real-world data in order to provide robust conclusion.[19]

In education and policy-making, they enhance planning, monitoring, and evaluation of projects.

The Role of Medical Institutes

Medical institutes play a central role in capitalizing the potential of AI and big data.[15,20] Academic leadership is essential in providing a strategic vision by adopting these technologies in the policy especially reforming the institutional research agendas and educational curricula. They should help in the capacity building of the institutes by providing the systematic training of faculty and students to improve digital competencies in the changing scenario. Institutes must also invest in appropriate infrastructure, including secure data platforms and modern simulation centers, skill labs and interactive online platforms. In addition, they should try to build partnerships with technology companies, government bodies, and international organizations to facilitate innovation and resource sharing. Finally, medical institutes should have a strong ethical regulation in place especially regarding the use of AI to ensure that both trainees and trainers are not just the passive consumer of technology taking away the critical thinking and logic. In addition, regulatory requirements should be implemented so that AI is used responsibly to meet professional standards and core medical values.[15] In this way, big data and AI will definitely change the future of medicine without undermining the benefits of human interaction and personal touch that underpins the fiduciary relation between medical institute and society at large.

Challenges

While the positive impact of big data and AI in the contemporary medicine is undisputed, they are not without limitations.[15,21] Inaccuracies and bias remain a major concern, as AI systems trained on biased and inaccurate datasets will inevitably amplify the existing inaccuracies. To date, many complex predictive models still have limited interpretability. Another area of improvement is the lack of clear regulatory guidelines and mechanisms ensuring data privacy and security for safe and responsible AI use.[22]

Global and Local Relevance

These technologies hold significant promise to provide standard training materials and improve service delivery in countries such as Nepal where infrastructure development and human resources are expanding. For example, AI-driven online platforms and digital tools can extend quality educational content beyond urban centers to remote learners who otherwise lack access to traditional resources.

In developed countries, big data analytics is already being used to optimize healthcare operations and financial management in hospitals. Common uses include predicting demand and supply, tracking medical inventory, and optimizing budgeting so that limited funds in low-resource settings are used more effectively. Nepal's healthcare system can utilize this aspect to improve administrative performance and service delivery.

In Nepal, telemedicine, mobile health, and AI-assisted diagnostics are being increasingly explored to expand care to rural areas, illustrating how digital health initiatives can bridge geographical gaps in health care access.[23] Nepal's internet penetration is reasonably good and expanding.[24] However, there is a need to provide reliable internet and digital infrastructure beyond urban areas. This should be addressed at the federal and provincial levels to bridge the digital divide to offer the rural communities the benefits of these technologies in healthcare.[25]

Nepal's National AI Policy 2025 emphasizes inclusive adoption for all sections of the society and capacity building, with plans to support marginalized groups through training and infrastructure development while strengthening data governance and equitable access.[26] The policy seeks to expand the pool of skilled human resources and digital infrastructure to benefit all.

CONCLUSION

In Nepal, these technologies offer opportunities to conduct clinical trials, competency-based education, and healthcare delivery. However, addressing challenges such as availability of good quality data, data privacy, interpretability, and ethical use is essential. Concentric effort of all stakeholders is required for investment in digital infrastructure, capacity building, and governance of these technologies to benefit the society at large.

REFERENCES

1. What is Big Data? | IBM. Accessed December 2, 2025. https://www.ibm.com/think/topics/big-data
2. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25(1):44-56. doi:10.1038/s41591-018-0300-7
3. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. Health Inf Sci Syst. 2014;2:3. doi:10.1186/2047-2501-2-3
4. Obermeyer Z, Emanuel EJ. Predicting the Future — Big Data, Machine Learning, and Clinical Medicine. N Engl J Med. 2016;375(13):1216-1219. doi:10.1056/NEJMp1606181
5. Beam AL, Kohane IS. Big Data and Machine Learning in Health Care. JAMA. 2018;319(13):1317-1318. doi:10.1001/jama.2017.18391
6. Big Data: The Management Revolution. Accessed December 2, 2025. https://hbr.org/2012/10/big-data-the-management-revolution
7. Wu PY, Cheng CW, Kaddi CD, Venugopalan J, Hoffman R, Wang MD. -Omic and Electronic Health Records Big Data Analytics for Precision Medicine. IEEE Trans Biomed Eng. 2017;64(2):263-273. doi:10.1109/TBME.2016.2573285
8. Nayarisseri A, Khandelwal R, Tanwar P, et al. Artificial Intelligence, Big Data and Machine Learning Approaches in Precision Medicine & Drug Discovery. Curr Drug Targets. 2021;22(6):631-655. doi:10.2174/1389450122999210104205732
9. Benke K, Benke G. Artificial Intelligence and Big Data in Public Health. Int J Environ Res Public Health. 2018;15(12):2796. doi:10.3390/ijerph15122796

10. Janjua HM, Rogers M, Read M, Grimsley EA, Kuo PC. A of analytics and B of big data in healthcare research: Telling the tale of health outcomes research from the eyes of data. Am J Surg. 2024;230:105-107. doi:10.1016/j.amjsurg.2023.11.026
11. Palmer J. How to harness AI's potential in research - responsibly and ethically. Nature. 2024;632(8027):1181-1183. doi:10.1038/d41586-024-02762-2
12. Shah NH, Milstein A, Bagley PhD SC. Making Machine Learning Models Clinically Useful. JAMA. 2019;322(14):1351-1352. doi:10.1001/jama.2019.10306
13. Wartman SA, Combs CD. Medical Education Must Move From the Information Age to the Age of Artificial Intelligence. Acad Med J Assoc Am Med Coll. 2018;93(8):1107-1109. doi:10.1097/ACM.0000000000002044
14. Nagi F, Salih R, Alzubaidi M, et al. Applications of Artificial Intelligence (AI) in Medical Education: A Scoping Review. Stud Health Technol Inform. 2023;305:648-651. doi:10.3233/SHTI230581
15. Chan KS, Zary N. Applications and Challenges of Implementing Artificial Intelligence in Medical Education: Integrative Review. JMIR Med Educ. 2019;5(1):e13930. doi:10.2196/13930
16. WHO. WHO issues first global report on Artificial Intelligence (AI) in health and six guiding principles for its design and use. Accessed December 2, 2025. https://www.who.int/news/item/28-06-2021-who-issues-first-global-report-on-ai-in-health-and-six-guiding-principles-for-its-design-and-use
17. Ngiam KY, Khor IW. Big data and machine learning algorithms for health-care delivery. Lancet Oncol. 2019;20(5):e262-e273. doi:10.1016/S1470-2045(19)30149-4
18. O'Connor O, McVeigh TP. Increasing use of artificial intelligence in genomic medicine for cancer care- the promise and potential pitfalls. BJC Rep. 2025;3:20. doi:10.1038/s44276-025-00135-4
19. Deivayanai VC, Swaminaathanan P, Vickram AS, et al. Transforming healthcare: the impact of artificial intelligence on diagnostics, pharmaceuticals, and ethical considerations – a comprehensive review. Int J Surg. 2025;111(7):4666. doi:10.1097/JS9.0000000000002481
20. Xu Y, Jiang Z, Ting DSW, et al. Medical education and physician training in the era of artificial intelligence. Singapore Med J. 2024;65(3):159-166. doi:10.4103/singaporemedj.SMJ-2023-203
21. Yu KH, Kohane IS. Framing the challenges of artificial doi:10.1136/bmjqs-2018-008551
22. Cajueiro DO, Celestino VRR. A comprehensive review of Artificial Intelligence regulation: Weighing ethical principles and innovation. J Econ Technol. 2026;4:77-91. doi:10.1016/j.ject.2025.07.001
23. Adhikari LM. Emerging Trends of using Digital Tools including Artificial Intelligence in Health sector in Nepal: What Next? NPRC J Multidiscip Res. 2025;2(1):12-24. doi:10.3126/nprcjmr.v2i1.74654
24. Online THT. Internet user density reaches 144.23% in Nepal. The Himalayan Times. May 29, 2025. Accessed December 25, 2025. https://thehimalayantimes.com/nepal/internet-user-density-reaches-14423-in-nepal
25. AI Inequality: UN Warns of a Growing Divide. Accessed December 2, 2025. https://nepalniceties.com/ai-inequality-global-nepal-digital-gap/?utm_source=chatgpt.com
26. National Artificial Intelligence (A.I.) Policy 2025. Accessed December 2, 2025. https://aiassociationnepal.org/national-artificial-intelligence-ai-policy-2082/