International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Navigating Opportunities and Ethical Challenges of Artificial Intelligence in Healthcare

SRINIKA, KRISHIV SIMHA, DHINESH SAAI, BS-MS

VISHWA VISHWANI INSTITUTE OF SYSTEMS AND MANAGEMENT
| THUMKUNTA, HYDERABAD, TELANGANA, INDIA |
| EMAIL: krishivsimha1@gmail.com| Mobile: 7207823322

ABSTRACT:

Artificial intelligence (AI) is emerging as a transformative force in healthcare, with the potential to improve diagnostic accuracy, operational efficiency, accessibility, and evidence-based clinical decision-making. This paper examines the ethical, operational, and social implications of healthcare AI, emphasizing patient rights, digital inclusion, and responsible innovation. AI-driven diagnostic tools and clinical decision-support systems can assist with early disease detection, treatment planning, and outcome prediction in specialties such as oncology, neurology, and cardiology. AI-enabled telemedicine may also extend specialist services to rural and underserved populations. However, these benefits are accompanied by concerns about algorithmic bias, limited transparency, patient-data privacy, unequal digital access, and medico-legal accountability. Using qualitative content analysis of selected secondary sources, including published studies, institutional reports, and government documents, the study examines these challenges in the context of India's Digital Personal Data Protection (DPDP) Act, 2023, and the DPDP Rules, 2025. The analysis suggests that technical performance alone cannot establish the ethical acceptability of healthcare AI. The paper proposes a four-layer governance framework encompassing legal, institutional, clinical, and technical safeguards, supported by digital inclusion and accessible redress mechanisms. It argues that responsible AI integration requires clinical validation, regular auditing, transparency, and meaningful human oversight to protect patient welfare, equity, dignity, and public trust while ensuring that AI supports rather than replaces clinical judgment.

KEYWORDS: Artificial Intelligence, Healthcare, Algorithmic Bias, AI Ethics, Digital Inclusion, Human Rights, Medico-Legal Accountability, Telemedicine.

1.INTRODUCTION:

In 2019, a widely used US healthcare algorithm was found to underestimate the healthcare needs of Black patients because it used past healthcare spending as a proxy for health needs *(*Obermeyer et al., 2019*)*. The lesson is not simply that AI can fail, but that it can perform as designed while producing ethically problematic outcomes. A diagnostic system may improve accuracy while remaining difficult to explain; telemedicine may extend specialist care while excluding patients without reliable connectivity. These tensions raise the central question of this paper: in India's evolving data-protection and healthcare landscape, when does a clinically useful AI system become an ethically acceptable one?

The promise, however, arrives with a shadow. The World Health Organization (WHO, 2023) has cautioned that the rapid adoption of untested AI systems in healthcare may lead to errors by health workers, harm patients, and erode public trust. It has also highlighted concerns about biased training data that may undermine equity and inclusiveness. Systems trained on historical data can reproduce existing inequalities, as demonstrated by Obermeyer et al. (2019), while opaque models can make decisions difficult to scrutinize, and unequal access to digital services can exclude vulnerable populations. These concerns highlight that technical performance alone is insufficient to establish the ethical acceptability of healthcare AI. Responsible healthcare AI must protect patients' rights, make decisions open to scrutiny, clarify responsibility for errors, and ensure that different communities can access its benefits.

The growing adoption of artificial intelligence in healthcare makes responsible implementation an important consideration for healthcare institutions, technology developers, and policymakers. However, the value of this transformation should not be measured by technological progress alone. AI systems must demonstrate clinical usefulness while protecting patient rights, promoting equitable access, and maintaining clear lines of accountability. Their successful integration therefore depends on safeguards that ensure technological innovation contributes meaningfully to patient welfare and public trust.

This paper argues that AI should be understood as an augmentative and human-centered technology. It should strengthen clinical judgment, expand access to care, and support evidence-based decision-making—but it must not replace human responsibility, empathy, or ethical reasoning. The study concludes that ethical acceptability requires safeguards extending

beyond data protection to address clinical reliability, fairness, transparency, accountability, and equitable access.

2. REVIEW OF LITERATURE:

- **Topol (2019)** examined the convergence of human and artificial intelligence in medicine and found that AI can improve diagnostic accuracy and efficiency but works best when clinicians remain actively involved.
- **Esteva et al. (2017)** showed that deep neural networks can classify skin cancer at a level comparable to dermatologists. However, such studies mainly measure technical performance and say little about fairness or accountability in real-world settings.
- **Mittelstadt et al. (2016)** studied the ethical issues raised by algorithms and identified opaque evidence, unfair outcomes, privacy risks, and difficulty in assigning responsibility as major concerns.
- **Obermeyer et al. (2019)** gave these concerns empirical support by showing that a widely used US healthcare algorithm underestimated the needs of Black patients because it used past healthcare spending as a proxy for need. However, most such evidence comes from high-income settings and may not transfer directly to India's rural and digitally excluded populations.
- **WHO (2021), UNESCO (2021)**, and **Floridi et al. (2018)** emphasize ethical principles for responsible AI, including human autonomy, transparency, accountability, and inclusiveness.
- **NITI Aayog (2021)** discusses responsible-AI principles in the Indian context. However, principle-based frameworks require appropriate implementation and enforcement mechanisms to translate their recommendations into practice.
- **Price (2015)** examined legal and regulatory challenges associated with opaque medical algorithms. This work highlights the importance of examining how algorithmic systems are evaluated and governed. In the Indian healthcare context, further analysis is needed to clarify how existing legal and professional responsibilities apply when AI contributes to clinical decisions.
- **ICMR (2023)** issued ethical guidelines for AI in biomedical research and healthcare in India, covering principles such as accountability, privacy and oversight, while the **Digital Personal Data Protection Act, 2023 (MeitY, 2023)** provides a legal basis for protecting personal data. However, data-protection law does not itself address algorithmic bias, clinical validation, or explainability.

3. RESEARCH GAP:

Although artificial intelligence (AI) is increasingly used in diagnostics, clinical decision support, and telemedicine, existing scholarship often examines its clinical benefits, ethical risks, and data-protection implications separately, with substantial evidence drawn from high-income settings. To the authors' knowledge, limited integrated analysis has examined how India's evolving data-protection framework intersects with the specific challenges of healthcare AI, including algorithmic bias, limited transparency, digital exclusion, and medico-legal accountability.

This gap has particular significance following the notification of the Digital Personal Data Protection (DPDP) Rules, 2025, which provide for phased implementation of the DPDP Act, 2023. As organizations prepare to meet their obligations under the evolving framework, it is important to distinguish the protection of patients' personal data from the broader safeguards needed to address potentially harmful or inequitable AI-driven clinical decisions. Data protection alone does not resolve questions of clinical reliability, fairness, transparency, or professional responsibility.

This study addresses the gap through a rights-based qualitative content analysis of relevant academic literature, policy documents, and regulatory frameworks. Guided by the question, when does a clinically useful AI system become an ethically acceptable one? it proposes a layered governance framework that connects patient rights and data protection with equitable access, institutional oversight, and professional accountability. The study treats AI as a tool to support, rather than replace, clinical judgment.

4. OBJECTIVES OF THE STUDY:

The study pursues five objectives, each contributing to its central question: When does a clinically useful AI system become an ethically acceptable one?

1. To critically assess the clinical applications of artificial intelligence in diagnostics, clinical decision support systems (CDSS), and telemedicine, examining their potential benefits and limitations.
2. To examine the ethical and socio-digital challenges of healthcare AI, with particular attention to algorithmic bias, transparency, patient-data privacy, and digital exclusion.
3. To evaluate the scope and implications of India's Digital Personal Data Protection Act, 2023, and the DPDP Rules, 2025, for healthcare AI governance, identifying areas that require complementary safeguards beyond data protection.

4. To analyze medico-legal accountability in AI-assisted clinical decision-making, focusing on professional responsibility, liability, and meaningful human oversight.
5. To propose a layered governance framework integrating legal, institutional, clinical, and technical safeguards to protect patient rights, promote equitable access, and ensure that AI supports rather than replaces clinical judgment.

5. RESEARCH METHODOLOGY:

Research Design and Data Sources: The study adopts a descriptive and analytical research design based entirely on secondary data, including peer-reviewed research, institutional reports, policy documents, and relevant legal and regulatory materials. The sources were considered for their relevance to healthcare AI, credibility, and contribution to the study's research objectives. The analysis is a qualitative synthesis of selected secondary sources rather than a systematic review of all available literature.

Analytical Approach: The study employs rights-based qualitative content analysis to examine the relationship between ethical principles, legal frameworks, and clinical practice. The material is analyzed thematically across six areas: clinical effectiveness, patient privacy, algorithmic fairness, transparency, equitable access, and professional accountability. These themes are examined alongside India's Digital Personal Data Protection Act, 2023, and the applicable DPDP Rules, 2025, to identify governance challenges and areas requiring complementary safeguards.

Scope and Intended Contribution: The study focuses on AI applications in diagnostics, clinical decision support, and telemedicine, particularly in the Indian context. It proposes a layered governance framework integrating legal, institutional, clinical, and technical safeguards while preserving meaningful human oversight.

Limitations: The study relies on secondary sources and includes no primary clinical or patient data. Findings from other healthcare systems may not fully apply to India. Furthermore, the implications of the DPDP framework must be assessed considering its phased commencement and evolving implementation.

6. CONCEPTUAL BACKGROUND:

1. Artificial Intelligence in Healthcare: Artificial Intelligence (AI) refers to computer systems capable of performing tasks such as learning, pattern recognition, reasoning, and decision support. In healthcare, AI analyses medical images, clinical records, and laboratory results to assist with diagnosis, treatment planning, and risk prediction.

2. Machine Learning and Deep Learning: Machine learning enables systems to identify patterns in data, while deep learning uses multilayered neural networks to process complex information. These techniques support medical-image analysis, disease prediction, and clinical risk assessment (Rajkomar et al., 2019).

3. Clinical Decision-Support Systems (CDSS): CDSS are software tools that assist healthcare professionals by analyzing clinical information and providing alerts, recommendations, or treatment suggestions. Their outputs require clinical evaluation and should not replace professional judgment.

4. Telemedicine: Telemedicine enables healthcare services to be delivered remotely through digital communication technologies. AI-enabled applications may support virtual consultations, preliminary triage, and remote monitoring. Their effectiveness depends on clinical suitability, connectivity, accessibility, and digital literacy.

5. Algorithmic Bias and Fairness: Algorithmic bias occurs when an AI system produces systematically unfair outcomes for particular groups. It may arise from unrepresentative training data, flawed assumptions, or historical inequalities. Obermeyer et al. (2019) demonstrated how an inappropriate proxy for healthcare needs could produce racial disparities in algorithmic decisions. Fairness therefore requires evaluating outcomes across relevant patient populations.

6. Transparency and Explainability: Transparency concerns the availability of meaningful information about an AI system's design, capabilities, limitations, and use. Explainability refers to the ability to understand and communicate factors influencing its outputs. These concepts are important for clinical examination, informed evaluation, and accountability (Mittelstadt et al., 2016).

7. Patient Rights, Privacy, and Data Protection: Patient rights include autonomy and participation in healthcare decisions. Privacy concerns the protection of personal information, while data protection involves legal and organizational safeguards governing its collection, use, storage, and sharing. AI-enabled healthcare therefore requires appropriate information, consent where legally required, confidentiality, and responsible data handling.

8. Digital Exclusion and Health Equity: Digital exclusion occurs when people face barriers to accessing or using digital services because of limited connectivity, affordability, accessibility, language support, or digital literacy. Health equity requires fair opportunities to benefit from healthcare. Inclusive AI deployment should address these barriers and consider the needs of underserved populations.

9. Human Oversight and Medico-Legal Accountability: Human oversight requires qualified professionals to evaluate AI outputs critically and intervene when necessary. It also helps address automation bias, the tendency to accept algorithmic recommendations without sufficient analysis. Medico-legal accountability concerns the legal and professional responsibilities associated with healthcare decisions and harm. In AI-assisted care, responsibility depends on the circumstances, applicable law, and the respective duties of clinicians, healthcare institutions, developers, and other relevant parties.

10. Responsible AI and Digital Health Governance: Responsible AI involves developing and using AI in ways that promote safety, fairness, privacy, transparency, and accountability. Digital health governance encompasses the laws, policies, institutional processes, and technical safeguards for guiding implementation and monitoring. The World Health Organization (2021) emphasizes the importance of human rights, ethical safeguards, and accountability in AI for health. Effective governance should include appropriate clinical validation, risk assessment, human oversight, and mechanisms for addressing adverse outcomes.

11. India's Digital Personal Data Protection (DPDP) Framework: India's DPDP framework comprises the Digital Personal Data Protection Act, 2023, and the Digital Personal Data Protection Rules, 2025. It establishes a framework for processing digital personal data, subject to applicable provisions and exemptions. The framework is relevant to protecting health-related personal data but must be considered alongside other safeguards addressing clinical safety, algorithmic fairness, transparency, and professional accountability.

7. ANALYSIS AND DISCUSSION:

The analysis reveals that artificial intelligence (AI) can improve clinical decision-making and healthcare access, but its benefits depend on reliable performance, equitable implementation, and effective governance. Clinical usefulness alone does not establish that an AI system is ethically acceptable.

7.1 Clinical Utility and Decision Quality

AI can assist with medical-image interpretation, disease identification, risk prediction, and clinical decision-making (Esteva et al., 2017; Topol, 2019). However, strong performance in selected tasks does not guarantee reliability across hospitals and patient populations. Adoption should therefore depend on clinical validation, demonstrated practical benefits, and continuous performance monitoring. AI recommendations must remain subject to professional examination and patient-specific clinical judgment.

7.2 Telemedicine and Equitable Access

India's e-Sanjeevani service and the Telemedicine Practice Guidelines, 2020, demonstrate the potential of digital healthcare to expand remote consultations (Ministry of Health and Family Welfare, n.d.; Board of Governors in supersession of the Medical Council of India, 2020). However, limited connectivity, device affordability, language barriers, and digital literacy may restrict access to vulnerable groups. Equitable implementation requires multilingual interfaces, low-bandwidth options, accessible service design, reasonable cost for device, and non-digital alternatives where necessary.

7.3 Data Protection and Patient Rights

India's Digital Personal Data Protection (DPDP) Act, 2023, and the Digital Personal Data Protection Rules, 2025, establish a legal framework for processing and protecting digital personal data. Their commencement follows separate schedules. Under the Act's commencement notification dated 13 November 2025, specified provisions took effect upon publication, certain provisions are scheduled to commence after one year, and further provisions after eighteen months, corresponding to May 2027. The Rules, notified on 14 November 2025, have a separate schedule under which some provisions took effect upon publication and others are deferred. Accordingly, the Act and Rules should not be described as though all their provisions became effective on a single date (Ministry of Electronics and Information Technology, 2025a, 2025b). Although this framework addresses important aspects of personal data protection, compliance alone does not ensure that healthcare AI systems are clinically reliable, fair, or sufficiently transparent. Complementary safeguards are needed to support clinical validation, meaningful human oversight, and accountability for AI-assisted decisions.

7.4 Algorithmic Bias and Transparency

AI systems may reinforce existing inequalities when training data or predictive methods inadequately represent particular populations (Ghassemi et al. ,2021). Obermeyer et al. (2019) demonstrated that a US healthcare algorithm using expenditure as a proxy for health needs underestimated the needs of Black patients. Although this finding cannot automatically be generalized to India, it illustrates how seemingly reasonable design choices can produce unequal outcomes. Healthcare institutions should evaluate performance across relevant patient groups and conduct regular bias audits. Transparency regarding intended use, limitations, and reliability is equally important, so clinicians can critically assess AI-generated recommendations (Mittelstadt et al., 2016).

7.5 Medico-Legal Accountability and Human Oversight

Consider a hypothetical case in which a rural clinician accepts an AI-generated chest X-ray report indicating no abnormality, but tuberculosis is subsequently confirmed. This scenario raises questions about the adequacy of clinical review, system reliability, institutional procedures, and the respective duties of the clinician and developer. Legal responsibility depends on the facts, applicable law, and each party's role; it should not be assigned automatically to a single party. Meaningful human oversight requires clinicians to have the competence, information, and authority to question AI recommendations and recognize automation bias. Institutions should establish documented responsibilities, escalation procedures, and mechanisms for investigating adverse outcomes (Char et al., 2018).

7.6 Towards a Layered Governance Framework

The analysis supports a four-level framework for responsible healthcare AI:

1. **Legal level:** Compliance with applicable data protection and healthcare laws, clearer accountability arrangements, and appropriate standards for clinical validation and incident reporting.
2. **Institutional level:** Responsible procurement, staff training, documented oversight procedures, adverse-event reporting, and periodic review.
3. **Clinical level:** Evidence-based use, clinician authority over patient-care decisions, and appropriate communication with patients about AI-assisted care.
4. **Technical level:** Validation across relevant patient populations, assessment of system limitations, regular bias and performance audits, and monitoring changes in accuracy over time.

Digital inclusion and accessible complaint and redress mechanisms should underpin all four levels.

These safeguards must function together. Data protection cannot substitute for clinical validation, while human oversight is ineffective without adequate training, reliable information, and institutional support. A layered approach enables healthcare organizations to assess not only whether AI performs a clinical task, but whether its use is safe, fair, accountable, and appropriate for the patients affected.

8. MAJOR FINDINGS:

Based on the secondary sources reviewed, the study identifies six major findings concerning the clinical value of artificial intelligence (AI), its ethical and social risks, and the governance challenges associated with its use in healthcare.

8.1. Clinical Benefits and the Need for Reliable Oversight

The reviewed literature indicates that AI can support medical diagnosis, image interpretation, and clinical decision-making across specialties such as radiology, oncology, and cardiology. However, strong performance in specific tasks does not guarantee reliable results in every clinical setting. Limited transparency in some AI systems may also make their recommendations difficult to interpret, highlighting the need for clinical validation, clear communication of limitations, and professional review.

8.2. Human Oversight and Automation Bias

AI should support rather than replace clinical judgment. Overreliance on algorithmic recommendations may lead healthcare professionals to overlook relevant symptoms or accept incorrect outputs without sufficient . Effective human oversight therefore requires appropriate training, independent clinical assessment, and the authority to question or override AI recommendations, particularly in high-risk situations.

8.3. Algorithmic Bias and Health Inequality

The reviewed literature shows that biased data and unsuitable predictive methods can produce unequal outcomes for different patient groups. Obermeyer et al. (2019), for example, demonstrated how an algorithmic design choice disadvantaged Black patients in a US healthcare setting. Although this finding cannot automatically be generalized to India, it illustrates how technical decisions can reinforce existing inequalities. Unequal access to digital devices, internet connectivity, affordable services, and local-language platforms may further prevent vulnerable communities from benefiting from AI-enabled healthcare. Fair implementation requires both bias assessment and measures to improve digital inclusion.

8.4. Data Protection and the Scope of India's DPDP Framework

Healthcare AI creates challenges involving personal data security, unauthorized disclosure, and secondary data use. India's Digital Personal Data Protection (DPDP) Act, 2023, and the DPDP Rules, 2025, provide a legal framework for digital personal data processing, subject to their applicable provisions and phased commencement. Although the Act does not establish a separate statutory category for sensitive health data, health data may still receive protection under the applicable framework. The scope of lawful processing and exemptions must be assessed according to the relevant provisions and circumstances. As implementation is phased, with commencement extending into 2027, the framework's practical implications continue to develop. Furthermore, data-protection compliance alone does not establish that an AI system is clinically reliable, fair, or sufficiently transparent. Additional clinical, technical, and institutional safeguards are necessary.

8.5. Medico-Legal Accountability

AI-assisted clinical decisions can make responsibility for patient harm more complex, particularly when clinicians, healthcare institutions, and technology developers contribute to decision-making. The applicable duties and allocation of liability depend on the circumstances, the roles of the parties involved, and relevant law. Clear institutional procedures, appropriate professional oversight, and effective complaints and redress mechanisms are essential to address these challenges.

8.6. The Need for a Layered Governance Framework

The study's central finding is that responsible healthcare AI cannot be achieved through a single legal or technical safeguard. It requires a coordinated framework comprising four layers:

- **Legal:** Compliance with applicable data protection and healthcare requirements, clearer accountability arrangements, and appropriate standards for clinical validation and incident reporting.
- **Institutional:** Staff training, documented responsibilities, incident review, and effective oversight procedures.
- **Clinical:** Evidence-based use, meaningful human oversight, and appropriate patient information and consent where required.
- **Technical:** Security measures, bias assessment, transparency of system limitations, and monitoring changes in performance.

Digital inclusion and accessible complaint and redress mechanisms should underpin all four layers.

Taken together, these findings suggest that responsible healthcare AI depends not only on technological capability but also on patient safety, fairness, accountability, and equitable access. The proposed layered governance framework offers a practical approach to integrating these requirements while ensuring that AI supports, rather than replaces, professional clinical judgment.

9. SUGGESTIONS / RECOMMENDATIONS:

The recommendations follow the four-layer governance framework developed for assigning actions to the stakeholders best positioned to implement them.

9.1. Legal Layer: Government and Regulators

The Immediate priority is to Develop healthcare-specific guidance and clearer AI accountability standards while India's DPDP framework undergoes phased implementation.

- **Issue healthcare-specific data guidance** for AI training, validation, and deployment, clarifying the application of the DPDP Act, 2023, and DPDP Rules, 2025, including relevant requirements for data use, security, breach response, and grievance redressal.
- **Develop AI-specific medico-legal standards** that clarify responsibilities according to the roles and duties of developers, healthcare institutions, clinicians, and relevant regulatory authorities.
- **Establish a national AI-in-healthcare ethics and safety framework** integrating clinical validation, data protection, cybersecurity, accountability, and digital inclusion, building on existing ICMR guidance (Indian Council of Medical Research, 2023; NITI Aayog, 2021).

9.2. Institutional Layer: Hospitals and Health-Tech Companies

- Establish multidisciplinary AI governance committees comprising clinicians, data scientists, legal experts, ethicists, healthcare administrators, technology developers, and patient representatives to oversee the responsible implementation of healthcare AI.
- Promote interdisciplinary collaboration by assigning clear responsibilities according to expertise, including technical performance, clinical suitability, legal compliance, ethical considerations, and patient welfare. Encourage joint decision-making to establish practical standards for safe, transparent, and accountable AI use.
- Provide regular training for healthcare professionals and institutional staff on AI limitations, responsible use, data protection, and appropriate human oversight.

9.3. Clinical Layer: Clinicians and Care Teams

- **Preserve meaningful clinical oversight:** Clinicians should retain the authority to question, modify, or reject AI recommendations based on the patient's condition, individual needs, and available clinical evidence.
- **Promote informed patient participation:** Patients should receive understandable information about AI's role in their care, relevant data practices, potential risks, and applicable rights. Consent should be obtained where required, and patients should have appropriate opportunities to raise concerns and exercise available choices.

9.4. Technical Layer: Developers and Evaluators

- **Require appropriate clinical validation before deployment**, including assessment in relevant real-world settings and across representative patient groups, rather than relying exclusively on laboratory tests or retrospective data.

- **Conduct regular performance and safety audits** to identify bias, model drift, cybersecurity vulnerabilities, error rates, and unequal outcomes. Systems presenting unacceptable risks should be corrected, restricted, or withdrawn as appropriate.
- **Improve transparency and explainability** by communicating a system's intended purpose, relevant training-data characteristics, limitations, and uncertainty in an understandable form. Explanations should help clinicians assess recommendations, particularly in high-risk areas such as cancer diagnosis, emergency care, and mental health.
- **Evaluate fairness across relevant population groups**, including differences related to gender, age, disability, geography, language, and socioeconomic circumstances. Research should assess AI performance among Indian populations, including rural, tribal, and other underserved communities.

9.5. Cross-Cutting Priorities: Digital Inclusion and Public Participation

- Conduct community consultations and public awareness programs to explain how healthcare AI works, its potential benefits and limitations, and its implications for patient rights and personal data.
- Improve digital inclusion by investing in affordable connectivity and low-cost devices, providing information in local languages, promoting digital literacy, and involving community health workers to help underserved populations access AI-enabled healthcare.
- Establish accessible feedback and grievance-redress mechanisms that allow patients to raise concerns about AI-assisted decisions, seek clarification, and report possible harm. Encourage public participation in developing and evaluating healthcare AI policies.

9.6. Limitations and Future Research

These recommendations are based on secondary sources and require further empirical evaluation. Future research should incorporate the perspectives of clinicians, patients, technology developers, and regulators to assess feasibility, implementation barriers, and real-world outcomes in the Indian healthcare context.

India has an opportunity to strengthen these safeguards proactively, reducing preventable harm while ensuring that technological progress serves patients and the wider public.

10. CONCLUSION:

This paper began with a central question: When does a clinically useful AI system become an ethically acceptable one? The analysis suggests that usefulness is the starting point, not the final measure. An AI system may perform well yet produce unfair outcomes, as demonstrated

by Obermeyer et al. (2019), or extend specialist care while excluding people who lack reliable digital access.

Three findings stand out. First, data protection is not decision protection. India's Digital Personal Data Protection Act, 2023, and DPDP Rules, 2025, establish a framework for digital personal data processing, but compliance alone does not address every concern involving algorithmic bias, explainability, clinical validation, or liability. With implementation proceeding in phases through 2027, complementary safeguards remain important as AI-enabled healthcare expands. Second, oversight must be meaningful. Human review is effective only when clinicians have the competence, information, and authority to question AI recommendations and guard against automation bias. Third, accountability needs greater clarity. When AI contributes to patient harm, responsibility may involve clinicians, healthcare institutions, and developers, depending on their respective roles, duties, and applicable law.

No single safeguard can address these challenges. The study therefore proposes four connected layers of governance: legal safeguards for data protection and AI-related standards; institutional mechanisms for governance, audits, and training; clinical safeguards for professional oversight and appropriate patient participation; and **technical** measures for validation, explainability, and bias and performance monitoring. Digital inclusion and accessible complaint and redress mechanisms should support all four layers.

As this study relies on secondary sources, its conclusions and proposed framework should be examined through future empirical research involving clinicians, patients, healthcare institutions, and regulators. Nevertheless, its central argument remains clear: an ethically acceptable healthcare AI system must be appropriately validated, sufficiently transparent, fair, meaningfully overseen, accountable, and accessible.

The future of healthcare AI should be measured not merely by what technology can achieve, but by whom it benefits, whom it might harm, and how responsibly it is governed. India has an opportunity to strengthen these safeguards proactively, helping ensure that innovation advances patient welfare, dignity, and public trust.

11.REFERENCES:

Board of Governors in supersession of the Medical Council of India. (2020). *Telemedicine practice guidelines: Enabling registered medical practitioners to provide healthcare using telemedicine.*
https://esanjeevani.mohfw.gov.in/assets/guidelines/Telemedicine_Practice_Guidelines.pdf

Char, D. S., Shah, N. H., & Magnus, D. (2018). *Implementing machine learning in health care: Addressing ethical challenges. The New England Journal of Medicine, 378(11), 981–983.* https://doi.org/10.1056/NEJMp1714229

Ekeland, A. G., Bowes, A., & Flottorp, S. (2010). *Effectiveness of telemedicine: A systematic review of reviews. International Journal of Medical Informatics, 79(11), 736–771.* https://doi.org/10.1016/j.ijmedinf.2010.08.006

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). *Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115–118.* https://doi.org/10.1038/nature21056

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). *AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Minds and Machines, 28(4), 689–707.* https://doi.org/10.1007/s11023-018-9482-5

Ghassemi, Price, W. N., II. (2015). Black-box medicine. *Harvard Journal of Law & Technology, 28(2), 419–467.*

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

Topol, E. J. (2019). 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

United Nations Educational, Scientific and Cultural Organization. (2021). *Recommendation on the ethics of artificial intelligence.* https://www.unesco.org/en/legal-affairs/recommendation-ethics-artificial-intelligence

World Health Organization. (2021). *Ethics and governance of artificial intelligence for health: WHO guidance*. https://www.who.int/publications/i/item/9789240029200

World Health Organization. (2023, May 16). *WHO calls for safe and ethical AI for health*. https://www.who.int/news/item/16-05-2023-who-calls-for-safe-and-ethical-ai-for-health