AI, PRIVACY AND MEDICAL NEGLIGENCE: ACCOUNTABILITY IN THE ERA OF ROBOTICS

Saransh Kumar, BALLB, CHRIST (Deemed to be University), Delhi NCR

ABSTRACT

Artificial intelligence (AI) is now the essential as well as crucial part of all the innovative field and that includes the field of healthcare, that has experienced a rapid transformation from traditional practices to advance robotic-assisted surgery, Telemedicine and the complex infrastructure for the development of biomedical research. However, with the ease of AI it instigates the challenges in the medical field relating to the question of accountability, liability and ethical governance. Unlike traditional medical negligence, where it was easy to identified the liability, whereas Artificial Intelligence make it complex on whom to hold accountable in the case of fault, whether it is the doctor, hospital, manufacturer, or the developer of the Robotic-Assisted Surgery machinery? This study also examines the extend till which the Artificial intelligence can be allowed to access the data of the patient without violating the right of privacy of the Individual, and highlights the crucial aspect of black box problem that questions the complex AI driven decision making algorithm and its biasness in the medical field. It also critically examines the current legal frame work governing AI in healthcare and compares it with the legislation of leading AI nation such as USA: AI Bill of Rights and UK: AI Regulation White Paper. This study argues for a hybrid liability model that has oversight over the algorithmic transparency, strict regulation of medical AI system and mandatory auditing mechanism. Ultimately, the paper highlights the urgent need to reorient medico- legal responsibility in an era where the Machines increasingly mediate clinical decision-making in the healthcare.

Keywords: Artificial Intelligence, Robotic-assisted surgery, Telemedicine, Medical research, medical negligence.

1. Introduction

Since the invention of Computers, there is an attempt of replacing the human efforts and expertise with that of the help of technological advancements and the efforts that are best known by the name of Artificial Intelligence.[1] All the human efforts that include problem solving, language learning, solving mathematical equation, complex decision making, retain memory and communicate.[2] The science tries to develop a system that act like as a human in the place actual human and they termed it as Artificial Intelligence in the world of science.[3]

The first person who think of using the machine to do work like the human was Alan Turing in 1950. He developed the concept of programming machines to do tasks like humans automatically. The first person to term these human alike machine as Artificial intelligence was McCarthy in 1956 as the science of making machine that are intelligent. Since it was firstly come in the knowledge of human beings that something this extraordinary can be created that this close to replicate human behaviour, it has gone through various stages of development and advancement in various fields specially in the field of medicines.

The use of AI in medical field has seen the rapid growth in last five decades due to the introduction of new technology and deep algorithm learning in the field of medical science and now AI is said to be the Ubiquitous in the medical field. It become the integral part of all medical technologies from disease diagnosis such as cancer diagnoses to personalize treatment, medical data analysis, better healthcare and many more areas of offering healthcare to patients. These advances and the better algorithms and programs used in the medical field led to a higher trust in these technologies from all interested parties such as doctors, nurses, health administrators, as well as patients.[4]

The aim of this paper is to identify and examine the role of the artificial intelligence plays in the field of medical along with addressing the legal frame work in AI-induced medical harm, especially in contexts involving robotic-assisted surgery, telemedicine, and algorithm-driven clinical decision-making.[5] By analysing the liability attribute, data governance of Black box theory, this paper also take stance on the comparative study of the regulatory approach by India,

[1] Russell, S., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.
[2] Nilsson, N. J. (2010). *The quest for artificial intelligence*. Cambridge University Press.
[3] McCarthy, J. (1956). *A proposal for the Dartmouth summer research project on artificial intelligence*.
[4] Reddy, S., et al. (2019). Trust in AI systems. *BMJ Health & Care Informatics*, 26(1).
[5] Price, W. N., & Cohen, I. G. (2019). Privacy in AI healthcare. *Harvard Law Review*, 132, 1127.

The USA, The United Kingdom.[6] The paper aims to propose a hybrid liability and regulatory framework that ensures accountability, transparency, and effective patient protection while supporting responsible innovation in AI-driven healthcare.[7]

2. Conceptual Framework: Artificial Intelligence in Healthcare

2.1 Meaning and Scope of Artificial Intelligence

Artificial Intelligence (AI) is technology enabling computers to perform tasks needing human intelligence, like learning, problem-solving, understanding language, and recognizing patterns[8], by using data and algorithms to learn, reason, and make decisions, powering everything. In the field of healthcare, AI has a wide hold that being the robots performing surgery or the algorithm that help human counterpart in medicine research. In contrast to traditional software systems that are programmed to follow the rules strictly made by law and work in according of the algorithm put by the user, AI systems are capable of learning from experience, either its are others who have been already be part of that work and changing their actions gradually based on what they have learned.

Machine learning, the most important subset of AI, allows machines to detect trends in huge amounts of data and to provide predictions or recommendations based on statistical relationships. Deep learning is a more sophisticated version of machine learning that utilizes multi-layered neural networks which are able to process intricate medical data like imaging scans, genetic data, and electronic health records.[9] These techs have been very successful in diagnosing cancer, interpreting radiology, and predicting diseases for better management. Nevertheless, the same characteristics that make AI great and capable to be used by human in all field – autonomy, versatility, and complexity- are also the ones that cause legal ambiguity. AI systems are not like human professionals as they lack consciousness, do not have the capacity to intend, and their outputs do not reflect the essence of being human that is moral responsibility.[10] Yet, there output can significantly influence clinical outcomes. This creates

[6] European Commission. (2020). *White paper on artificial intelligence*.
[7]
[8] Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms. *Big Data & Society, 3*(2), 1–21.
[9] Price, W. N. (2017). Black-box medicine. *Harvard Journal of Law & Technology, 28*(2), 419–471.
[10] Pasquale, F. (2019). *New laws of robotics: Defending human expertise in the age of AI.* Harvard University Press.

tension within legal framework that are predicted on human agency and fault.

2.2 Evolution of AI in Medical Practice

The employment of the computational tool in the field of medicine is not something new. Expert on late twentieth century has developed and relied upon a rule based logic system and explicit medical knowledge.[11] Such system worked as a decision- support tools, assisting clinicians while remining transparent and traceable and hence there was no disruption to the legal doctrine that were in existence of that time as it follows basic rule that generally follows the format "IF condition THEN action." For example, in an expert system for medical diagnosis, a rule might be "IF patient has fever [AND cough THEN consider flu.]" But the current wave of AI in healthcare sector is completely turnaround from the earlier system. Current Ai system is capable to learn from the gigantic data sets which provide the better and superior performance from that of the computational tool used earlier but at the same time, give up the feature of being explainable.[12] Robotic-assisted surgery is considered as the apex of the evolution in medical field as it combine the two distinct as well as important aspect of medical practices – that is the physical precision of machines and the smartness of algorithms that the only posses by the specimen of human beings. The surgeons always in control of robotic system while having supervision over it which result in human interaction into very accurate and sometimes AI- led recommendation about the surgical routes and risk of assessments. Along with that telemedicine has directly involved the AI in healthcare. The adoption of AI in the medical field led to the easy accessibility of the healthcare in the area that are neglected otherwise. However, due to Ai the physical as well as relational distance has increase between the healthcare provider and the patients has brought a about complication in the traditional concepts of duty of care and professional supervision

2.3 Robotic-Assisted Surgery and Degrees of Autonomy

Robotic- assisted surgery occupies a crucial position in the world of AI- healthcare due to it having the nature of surgical intervention that are of high risk. Human supervision is still required as AI didn't reach the level to be able to let operated on its own, but still the degree of autonomy varies.[13] As there are system that support surgeons by acting as an assistant while

[11]Burrell, J. (2016). How the machine "thinks". *Big Data & Society, 3*(1), 1–12.
[12] Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why a right to explanation does not exist in the GDPR. *International Data Privacy Law, 7*(2), 76–99.
[13] Topol, E. (2019). *Deep medicine.* Basic Books.

performing their duties while other give real- time analytics or have whole process automated. As in the increase of the autonomy in surgical robots, the control of the human operators decreases thus raising a concern about the accountability and the quality of the surgeries. The Cruz of the matter isn't that Ai will take place of the doctors or the other medical actor but rather the point of whether it has meaningful influence over clinical decision.[14] Wherever Ai system shape diagnosis or treatment pathways, they become a crucial to the decision making process. This fundamentally raise the question of how the liability of the human and non-human works in the healthcare ecosystem should be distributed.

2.4 Legal Significance of AI as a Decision-Influencing Actor

The challenge presented by AI as tool in healthcare emerge from its unclear legal status. AI is neither a tool to be used but it's an independent legal entity. Still, the clinical decision made by AI is significant in the operating as well as medical room. Considering Ai as a neutral instrument overlooks its ability to influence outcomes, on the other hand, giving it a legal entity introduces complicated ethical and legal issues. Therefore, the law must adopt an approach that focus on control, foreseeability, capacity to prevent harm and to held AI accountable. Understanding AI as decision- influencing entity provided a foundation to rethink the liability and regulatory oversight without resorting legal fiction.[15]

3. Traditional Medical Negligence: Doctrinal Foundations and Legal Assumptions

3.1 Core Elements of Medical Negligence

Medical negligence is specific area of tort law that deals with liability due to healthcare practitioners unprofessional act. The traditional framework of medical negligence requires the claimant or the plaintiff to establish four essential elements: (i) the establishment of a legal relationship of care, (ii) breach of that duty, (iii) the injury being a result of the overlap of breach and patient suffering, and (iv) actual damage.[16] The above law is kept to be in line of the general negligence law but is specially made to fit the highly specialized and technical medical practice. In most jurisdiction, including India, law have recognised that the medical professional owe a duty of care towards their patient once a relationship of doctor-patient is

[14] Abbott, R. (2020). The reasonable computer. *George Washington Law Review, 86*(1), 1–44.
[15] Calo, R. (2015). Robotics and the lessons of cyberlaw. *California Law Review, 103*(3), 513–563.
[16] Indian Medical Association v. V. P. Shantha, (1995) 6 SCC 651.

established. This obligation includes not only clinical diagnosis and therapy but also surgical procedure and post-operative care. The question of whether there is a breach of duty is answered in terms that if the doctor of same specialization and experience would have provided in the same manner and under the same conditions. The standard is not one of perfection but of reasonable skill and diligence. Causation requires proof that the breech of duty was the proximate case of the injury, It is the legal principle determining if an action is sufficiently linked to an injury to establish liability, meaning the harm was a reasonably foreseeable outcome, not just a random consequence in a long chain of events. This element is particularly significant in medical cases, where multiple factors may contribute to adverse outcomes[17]. Finally, damage refers to demonstrable harm, it can be physical, psychological, or economic, suffered by the patient as result of negligent conduct by doctor. The doctrine framework provides a relationship between professional conduct of healthcare professional and patient harm. Its assume clinical decision made by the professional whose reasoning can be examined, explained, and evaluated by courts. The above assumptions do not hold when artificial intelligence is integrated into the medical decision process.

3.2 The Bolam Principle and Professional Deference

Legal medical negligence, one of the areas of law where judicial deferential literally comes to the fore, is very much based on professional judgment. The 1957 English case Bolam v. Friern Hospital Management Committee laid the foundation for this concept that a medical professional acting according to the standards that a responsible body of medical opinion would recognize as right, is not in fact, negligent.[18] Bolam test is a test that is widely used in courts of common law jurisdictions, including but not limited to India. But the Idea is very much relevant in Indian jurisdiction. The Supreme Court in Jacob Mathew v. State of Punjab strongly opined that for the negligence to be considered a crime, it has to be very grave or reckless[19] and at the same time there is a need for medical professionals to be protected from unwarranted harassment which is sometimes the case, when, through the courts, they become liable for their mistakes, as it were, letting the medical practice itself be the judge. Likewise, in Kusum Sharma v. Batra Hospital, the Supreme Court set out guidelines that highlighted the need for judicial restraint and the role of expert opinion[20] in medical negligence cases. The Bolam principle

[17] Maneka Gandhi v. Union of India, AIR 1978 SC 597.
[18] Bolam v. Friern Hospital Management Committee, [1957] 1 WLR 582 (UK)
[19] Jacob Mathew v. State of Punjab, (2005) 6 SCC 1.
[20] Kusum Sharma v. Batra Hospital & Medical Research Centre, (2010) 3 SCC 480.

reflects an implicit trust in human expertise and professional consensus. It assumes that medical knowledge evolves through collective human experience and that peer opinion provides a reliable benchmark for evaluating conduct. However, this principle becomes problematic in the context of AI-driven healthcare. Algorithmic systems do not participate in professional communities, nor are their "opinions" shaped through ethical deliberation or experiential learning in the human sense. When the healthcare practitioners rely on AI formed recommendations, determining whether such reliance conforms to a responsible body of medical opinion" becomes unclear. If most of the hospitals adopt particular AI system following same algorithm, does adherence to its recommendations automatically satisfy the Bolam standard? Conversely, can deviation from AI device be considered negligent? These questions reveal the doctrinal strain imposed by AI on traditional negligence principles.

3.3 Human-Centric Assumptions in Negligence Law

The law of medical negligence is built its foundation on the assumption that is created by law. It presumes that decision-makers possess intent, judgment, and moral responsibility. By recreating the thought process that caused the injury, the courts are able to judge negligence and to determine if a reasonable professional would have acted differently considering the same situation. The use of artificial intelligence challenges the legal framework by introducing non-human decision-making agents that lack consciousness, ethical reasoning, moral and intent. AI system operates on statistical correlations instead of normative judgment. Although there output can be accurate, but the reasoning they have to reach that output are often opaque and inaccessible. As a result, Court finds out it difficult when the breach of duty occurred due to the decision taken by AI results in harm. Furthermore, foreseeability is a major factor in negligence law. A practitioner is required to foresee the risks that are reasonably foreseeable and subsequently take appropriate measures. In the case of self-learning AI systems, it becomes particularly hard to make any predictions about future behaviour. The algorithms can be updated with new data which might lead to the generation of outcomes that were not predictable at the time of installation. This unpredictability serves to weaken the application of the traditional fault-based liability mode.

4. AI-Induced Medical Harm and the Crisis of Accountability

4.1 Diffusion of Responsibility in AI-Driven Healthcare

One of the most significant legal challenges posed by AI in healthcare is the diffusion of responsibility In contrast to the past when the doctor or the hospital was solely responsible for the patient's care, the new AI-powered medical systems rely on several actors who are responsible for different steps of the process starting from the design and going all the way to the actual use.

During the diagnosis and treatment process, the doctors are in contact with the AI systems and usually depend on algorithm-based suggestions to make their clinical choices. Health care institutions acquire AI systems, implement them, and keep them running as a part of their daily operations. Robotics companies do the hardware part, while the software developers go the extra mile to ensure the algorithms are working properly, even permitting update installations remotely after going live. A shared power dynamic exists among the different players in the system; however, total control is still not available to any one of them.

It is difficult to pinpoint the legally responsible person when a patient suffers injury during an AI-aided surgery. The physician might contend that the use of AI was tantamount to following the highest technological standards. The hospital might argue that it only offered a base for the operations. The manufacturers might claim the equipment worked properly, while the developers might indicate problems with the data that were not under their control. Such a scattering of accountability leads to a legal vacuum which obstructs patient's access to justice.[21]

4.2 The Problem of Causation in Algorithmic Decision-Making

Causation is a significant factor in negligence law and has the burden of proof that the accused person’s conduct was the primary reason for the patient’s suffering. In healthcare application of AI, it is a challenge to demonstrate causation since the AI systems work in intricate and frequently non-transparent manners, thus complicating the process of identifying the specific decision that caused the harm.

Black box algorithm[22] on which AI works on do not provide clear explanation for their outputs, making it difficult to find the pathway from input to outcome. In the case where an AI system suggests a certain surgical method or a diagnostic result, and subsequently the patient suffers damage, the courts might not be able to tell if the damage was caused by erroneous algorithm

[21] Calo, R. (2015). Robotics and the lessons of cyberlaw. *California Law Review, 103*(3), 513–563.
[22] Burrell, J. (2016). How the machine “thinks”. *Big Data & Society, 3*(1), 1–12.

reasoning, a mistake made by a doctor, or a combination of both factors.

Increasing this uncertainty is the result of automation bias, which is a phenomenon in which human decision-makers trust the automated systems too much.[23] Medical professionals may even go with AI's recommendations when their gut feeling tells them the opposite. Even though such dependence may lead to efficiency, it still questions how far the physicians can be responsible for the choices affected by the non-transparent algorithms.

4.3 Absence of a "Reasonable AI" Standard

Traditional negligence law evaluates conduct against the standard of a "reasonable professional. But there is no equivalent standards exists for artificial intelligence As the courts who decided the reasonability on traditional negligence lack criteria for assessing whether AI system acted reasonably or not, specially when its decision-making process is inaccessible.

Some scholars argue for the development of a "reasonable AI" standard based on industry best practices, transparency, and safety benchmarks.[5] However, without regulatory guidance, such standards remain speculative. The absence of clear benchmarks exacerbates legal uncertainty and discourages consistent adjudication of AI-related medical negligence cases.

Classic negligence law evaluates human behaviour by comparing it with that of a "reasonable professional" as per the standard. There is no parallel principle for AI. The courts lack a basis for the evaluation of an AI system's actions when the rationale is kept secret. Some scholars are proposing a "reasonable AI" standard[24] that is derived from factors such as the best practices of the industry, transparency, and safety benchmarks.[5] However, these standards will continue to exist only in theory without the enforcement of regulations. The vagueness of the rules adds to the already complicated legal situation and also results in the inconsistent judgments of medical negligence cases involving AI.

4.4 Ethical Governance and Trust Deficit

The ethical governance and public trust issues of AI-driven healthcare revolve around accountability gaps. The live and the personal data entrusted by the patients to healthcare system are their responsibility to kept secure. When harm occurs and the healthcare system

[23] Citron, D. K., & Pasquale, F. (2014). The scored society. *Washington Law Review, 89*, 1–33.
[24] Abbott, R. (2020). The reasonable computer. *George Washington Law Review, 86*(1), 1–44.

fails to clearly impose the responsibility that has causes the harm, results in people loosing trust on medical institution. Legal accountability serves not only a compensatory function but also a normative one. It makes the consumer of the healthcare service to have expectation regarding the acceptable conduct of the institution which helps in reinforcing ethical standards. The failure to frame a proper liability frame work to AI- mediated healthcare risk normalising irresponsibility and weakening safeguards against harm.

5. Product Liability and Institutional Responsibility

Accountability and liability concerns are inevitable as artificial intelligence systems have a greater impact on medical decision-making. Because current legal frameworks were primarily created for human decision-makers and static medical devices, while AI promises efficiency and accuracy in healthcare delivery, it also creates a "accountability gap" when errors occur. This section highlights the shortcomings of current liability models while examining product liability and institutional responsibility as two important strategies for addressing medical harm caused by AI.

5.1 AI as a Product for Medicine

Treating AI systems used in healthcare as medical products subject to product liability regimes is one way to close accountability gaps. Regardless of fault, this model could hold developers and manufacturers strictly accountable for flaws in AI systems. Because strict liability lessens the burden of proving negligence, the main benefit of this strategy is patient protection. Such a model can greatly increase access to compensation given the technical complexity of AI systems and the challenges patients face in comprehending algorithmic decision-making.

However, there are significant conceptual difficulties when applying conventional product liability principles to AI systems. AI systems are dynamic and adaptable, in contrast to traditional medical devices. As they continue to learn from fresh data following deployment, they might eventually alter their behavior. Determining whether a defect existed at the time of manufacture thus becomes challenging, especially when damage results from software updates or post-deployment learning. This calls into question the fundamental tenet of product liability law, which emphasizes flaws found at the point of sale.

Furthermore, biased or insufficient training data may be the source of errors in AI systems

rather than design defects. For example, an AI diagnostic tool may generate incorrect results for patients if it is trained on data that underrepresents specific demographic groups. In these situations, it is difficult to attribute liability solely to a manufacturing defect. As a result, assigning accountability necessitates a careful evaluation of data governance procedures, model training procedures, and the deployment context of the AI system. These elements show that although product liability offers a helpful framework, the complexity of AI-driven medical harm cannot be adequately addressed by it alone.

5.2 Liability for Hospitals as Institutions

Hospitals play a critical role in healthcare through the use of artificial intelligence, and they cannot escape liability for any harm caused to a patient by the use of an artificial intelligence application simply by claiming that the device was the cause of the harm. As an institution, hospitals are required to provide care for patients, including using artificial intelligence systems, and therefore, hospitals have a responsibility to perform due diligence to ensure that healthcare providers receive appropriate training to use the artificial intelligence technology, as well as ongoing performance evaluation of the system in order to identify and address errors or bias in the application of the technology.

When an institution does not have any oversight or governance over AI systems used within the organisation, this could be considered as "institutional negligence." An example of this could be a hospital relying solely on its AI System to help make clinical decisions without any human involvement or supervision or ignoring the limitations of the technology. In those situations, it could be determined that the organisation is liable for any injuries or damages caused by the AI System. Hospitals are also in a better position than an individual clinician to manage and distribute risk through insurance or other compliance programme options. The way hospitals are structured puts them into a unique position to absorb any losses and be able to take corrective actions much more efficiently.

Hospitals can also directly benefit from AI usage through improved efficiency, reduced costs, and better management of patients. Because hospitals are able to gain these advantages from AI usage, it is reasonable to assume that the organisation will also be liable for any associated risks. The rationale for supporting enterprise or apportioned liability models is that the responsibility for the risk will be determined according to the institutions control, benefit and ability to mitigate any potential harm.

The inability of either negligence law or product liability standards alone to sufficiently deal with AI-driven medical malpractice is documented. Negligence laws focus on accountability and assume that responsibility for harming patients stems from an identifiable breach of duty by an individual. That is not true in the case of AI-assisted healthcare where decision-making is shared between practitioners (clinicians) and machines/algorithms. As a result, determining who is at fault and how they inflicted harm is very challenging due to algorithmic opacity and limits to the explainability of many AI models.

In the case of product liability standards/systems, the way they are currently set up is to simplify the complexities surrounding AI as an inanimate or static product. Product liability sets a standard of identifying defects at the time of manufacturing; however, the modelling of AI systems is dynamic in that they continue to learn and evolve and, therefore, do not conform to this standard. The limitations of current liability models demonstrate the need for a hybrid liability structure where those liable for AI-related issues share responsibility based on control over the AI system, severity of harm suffered, and benefits received from the system. By applying this kind of model, all stakeholders (healthcare providers, patients, innovators) will have a more realistic and effective response to AI-related issues and be accountable for their respective roles in enhancing patient safety, providing equitable compensation, and promoting responsible innovation.

6: Privacy, Consent and Patient Data Governance in AI-Driven Healthcare

6.1: The Placard of Medical AI on Patient Data

Although there is an increasing trend among many health AI developers to create AI tools and technologies that originate from their own data source, many existing health AI tools were originally based on existing data sources. Unlike traditional clinical decision support, in which data is collected episodic and for specific purposes, medical AI models require the continuous availability of diverse and large numbers of patient data to be effectively functional; therefore, it is essential that medical AI models use real-time patient data, including but not limited to, demographic, social determinants of health, laboratory, radiologic, and other clinical information.

Representing a major paradigm shift in how the relationship between patients and healthcare providers regarding patient information and data, the introduction of large electronic health

record data sets to the development of an AI model makes it possible to utilize patient data for secondary and future applications, such as information for training algorithms, optimizing models, and generating future products. Although the transformation of patients from being a rights-bearing individual to being a data source may enhance medical innovation and support efficient healthsystems, the patient's loss of control over personal data creates significant issues regarding the deterioration of patient autonomy and agency.

There are certain types of medical records that display evidence of the psychological disposition(s) and the physical condition(s) of the individual(s). As a result, Medical Records have a heightened degree of sensitivity when compared to typical personal records. Historically, legal systems have recognized that unauthorized access to or misuse by third-parties of Medical Records could result in harm to or discrimination against an individual or group based upon their medical diagnosis.

The introduction of AI technologies has escalated the potential for misuse and discrimination against individuals who have medical disabilities due to the high volume of medical-related information (i.e. Big Data) being processed through automated, unaccountable methods.

6.2 Privacy and the Right to Determine the Use of Personal Information

Privacy has been viewed as a founding principle of the modern Data Protection Framework; Accordingly, Privacy is recognized in the USA, as well as in many jurisdictions; however, the Supreme Court of India has recently recognized Privacy as a Fundamental Right, based upon Justice K. S. Puttaswamy v. Union of India[25] and the inherent nature of Privacy to Human Dignity and Personal Autonomy. Central to this issue of Privacy is the "Right to Determine the Collection, Use and Disclosure" of Personal information, specifically with respect to sensitive Medical Records and other similar sensitive Personal Records.

The application of advanced technology in the healthcare system creates a tremendous strain on the right of informational self-determination. Many advanced healthcare technologies rely on the collection of a large volume of continuous data that may not be necessary to carry out a medical intervention. Predictive analytics, automated patient profiling, and monitoring patients in real time converge in a manner that blurs the lines of whether the data is being used for

[25] Puttaswamy (Justice K. S.) v. Union of India, (2017) 10 SCC 1.

legitimate clinical purposes or if the data is being misappropriated inappropriately.

As a result, when a person does not understand how their data will be used and/or how its use may be changed in the future, the right of informational self-determination is significantly hindered.

In order for a healthcare data processing system to comply with constitutional standards, the data must be collected in a legal, necessary, and proportionate manner. It is important to clearly articulate the lawful purpose of collecting data and to ensure that data is used according to pre-established guidelines. Without these safeguards, healthcare AI systems may violate a person's right to privacy.

Concerns about the protection of the right of privacy via advanced healthcare technology are heightened within the context of publicly funded healthcare systems. In situations of healthcare systems where there are significant structural power imbalances, patients have limited ability to refuse their data from being processed without jeopardising their ability to receive necessary medical services. In these cases, consent will likely be used as a mere formality rather than a legitimate exercise of patient choice.

6.3 Informed Consent Within the Context of Algorithmic Medicine

The core principle of medical law and ethics, for many practitioners and scholars alike, has always been respect for patient autonomy and control over their own body through informed consent. Informed consent is based on traditional models, which presume patients can be adequately educated regarding the purpose and nature of a medical intervention, its associated risks and benefits as well as the alternative treatment options available.

The introduction of Artificial Intelligence (AI)-supported decision-making complicates this basic presumption. First, algorithmic “black boxes” raise the question of how transparent (or non-transparent) AI systems will provide output but do not provide any meaningful information about the reasoning behind those outputs; therefore, healthcare practitioners may also not have access to sufficient knowledge on how an AI recommendation was generated. As a result, when patients consent to a course of treatment, it is possible they do not understand what role (if any) AI played in that decision or the potential for algorithmic bias or errors resulting from the automated process. In such instances, this type of consent is both ethically and legally flawed.

In addition, the dynamic nature of AI systems makes providing informed consent at a single point in time insufficient. Patient data could be used multiple times for future training of AI systems, upgraded versions of AI systems, or secondary research purposes not envisioned when the patient provided consent. As a result, these issues create a need to develop an alternative conceptualisation of informed consent as a continuous, evolving process, rather than a standardised procedure (i.e., "one-time") only. Regulatory systems need to ensure that they not only provide the means for formal disclosures, but also utilize a model of continual patient engagement and substantive transparency.

6.4 Regulatory Framework for Protecting Patient Information

The Regulation for Protecting Digital Personal Data in India (2023)[26] is a response to the increasing concerns surrounding the protection and security of personal data. The Regulation specifies critical principles for how patient data should be handled, such as ensuring that patients provide consent for their data being used and the purpose of using data, minimizing the amount of data collected, and implementing security measures to ensure its safety.

However, while the Regulation provides a general framework for protecting digital personal data, it does not take into account the specific issues that arise from the use of AI, including issues related to predictive algorithms learning from the data used for training, the use of predictive algorithms for automated decision-making, and the ability of predictive algorithms to provide comprehensive logic and explanations.

By comparison, the European Union's General Data Protection Regulation (GDPR) has specific provisions relating to predictive algorithmic decision-making, including the right for a user of an algorithm to receive meaningful information regarding the logic and methodology employed to create and implement a predictive algorithm. Although the GDPR is not directly applicable to the Republic of India, it provides an important normative framework for developing the regulations required for AI predictive algorithms used in medical settings.

The lack of tailored regulation governing medical AI results in a number of protection gaps. AI predictive algorithm systems could operate without specific requirements regarding audit, audit trail, and governance of the lifecycle of that data, which could lead to how they undermine individuals' privacy rights, autonomy, and contribute to erosion of public trust. Addressing

[26] Digital Personal Data Protection Act, 2023 (India).

these lacks in protection is essential for establishing and maintaining patient rights and ensuring the long-term legitimacy and ethical sustainability of AI technology in healthcare.

7. The Black Box Algorithm, and Its Application to Clinical Care: A Lack of Transparency

7.1 The Definition of the Black Box Problem and A Brief History

The "Black Box" issue describes a situation where the workings of specific AI technologies are not readily accessible or understood by humans. The latest generation of machine-learning (or "ML") software, on the other hand, is based on very complex, multiple layers of calculation, making it impossible for most people to reasonably explain how an ML program is generating its results. These technologies tend to produce highly accurate predictive results, although they do not offer a rationale that a human could easily understand, challenge, or question.

Concerns over the lack of transparency in the algorithms being used in AI-powered systems have risen in importance as these technologies move into areas with significant impact on people's lives, such as grounding laws, finance, and healthcare. Scholars refer to this phenomenon as the epistemic opacity[27], where the developers themselves may not have a complete understanding of how their algorithm will produce its final output (i.e., its predictions) and consequently create a lack of understanding for the end-user. In particular, in the context of healthcare, the lack of transparency surrounding the logic behind medical decisions has severe consequences because of the gravity of many of the medical-related decisions; therefore, with the lack of understanding surrounding ML technologies, there is no accountability of the medical professional nor is there a foundation to provide patient-centred, ethical medical care.

7.2 Constitutional Concerns Pertaining to Articles 14 and 21

The Black Box Algorithms create disturbing implications for two parts of the Constitution - Article 14 and Article 21. Article 21, as interpreted broadly by the Supreme Court, protects an individual’s right to life as well as rights associated with their dignity, physical autonomy and ability to make decisions relating to them with the benefit of being informed. In the landmark judgement of Justice K.S. Puttaswamy vs Union of India, the Supreme Court reiterated that

[27] Pasquale, F. (2015). *The black box society.* Harvard University Press.

decisional autonomy and the right to determine the information they are going to have to base their decisions on is part of an individual's dignity. Black box algorithms that create medical decisions about patients do so without patients having the relevant knowledge and subsequently the control over how those decisions were made. Consequently, the dignity component of Article 21 cannot be met when patients do not have a meaningful understanding or level of control over how medical decisions are made about them.

Black box algorithms also contravene Article 14 which establishes equality before the law and prohibits arbitrary action by the State. The Supreme Court has historically established that arbitrary action by the State is inconsistent with equality. When medical decisions are made by way of an AI, and those decisions cannot be explained, by definition those decisions have the potential to be made arbitrarily. This is particularly troublesome when an AI policy is implemented in a public healthcare system which would also constitute "State Action". If two patients with the same medical conditions are treated in different ways, that would violate the constitutional protections against arbitrary State Action and equal protection, which is a cornerstone of our Constitution.

7.3 Algorithm Bias and Equality Issues

Algorithm Bias creates additional layers of confusion on the existence of Bias within medical decision-making processes. Algorithm Bias occurs when AI algorithms create systemically unequal outcomes based on the training data that has been compromised due to Biased Data, Bad Proxies or Not Aligning with the True Context of the Patient. In Healthcare, an example of this could be the creation of incorrect/incorrectly diagnosed Disease; Providing different Treatment options based upon Race; and/or completely excluding certain Historically Marginalised Groups from Accessing Care.

Also Healthcare has always relied on datasets derived from Historical access disparities, so if AI algorithms train on these datasets, the algorithms can inadvertently mimic and therefore validate Historical Disparities. The classic example is providing Healthcare Spending as a Proxy for Medical Need which results in a racial bias allocated to Treatment.[28] Because of the Indirect Discrimination that is created under Article 14 of the Constitution, this could clearly establish a Constitution violation. As such, the Opaque Nature of Black Box Algorithms

[28] Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm. *Science, 366*(6464), 447–453.

increases the level of Risk associated with inadvertently creating Biased Outcomes since detecting, Challenging and/or Remedying Bias in Black Box Algorithms is Extremely Difficult.

7.4 Righting the Wrong of Without Due Process: The Right to Information

When a hospital or healthcare organization creates an algorithm (commonly called a "black box"), the patients using these automated processes will not know how those decisions were made and what implications those decisions will have on their own health. Patients have rights, and the process of making decisions using algorithms in the healthcare sector should include transparency and reasoned decision-making.

When patients do not have access to the reasoning used to make a decision via an algorithm, they are unable to hold providers accountable by questioning decisions that affect their health. Likewise, without transparency, courts will have little ability to review or adjudicate claims against negligent healthcare providers. Therefore, it is important for policymakers and scholars to take steps to implement a right to information (Explanatory Rights) for those affected by the decisions made by algorithms in their lives. While complete transparency may not always be feasible, states will be able to implement a standard of explainability sufficient to hold accountable the algorithm developers and to ensure adequate judicial review.

7.5 Comparative Perspectives: EU GDPR and AI Act

Comparative legal frameworks serve as a helpful tool for bridging the transparency gap. The European Union (EU) General Data Protection Regulation (GDPR) provides users with protections against unaccounted-for decision-making through automated means, as well as the right to receive "meaningful information about the logic involved" in the automated decision-making process, though still subject to potential challenges, for example - defining the concept of “meaningful.”

Most importantly, the EU AI Act utilizes a risk-based approach in creating a legal framework for AI and considers Medical AI to be “High Risk”. As such, Medical AI is required to have certain legal obligations to ensure that there are mechanisms in place to allow for human oversight, regulation of data use, and accountability of algorithms. By establishing a clear and transparent mechanism for operating in high-risk domains, the EU ensures compliance with

fundamental rights protection and addresses the black box problem.

By employing a combination of Constitutional Principles along with Regulatory Oversight and Technical Safeguards, transparency in healthcare must be a requirement and not an option as required by Articles 14 and 21 of the Constitution. In order to protect the concepts of equal treatment, dignity, and trust in AI generated medical decisions, we must infuse auditability, explainability and oversight into the Governance Structures of AI Technology.

8. Comparative Legal Framework Governing AI in Healthcare: An Indian Constitutional Perspective

When considering the Governance of Artificial Intelligence (AI) in the field of Healthcare, we must not only think about the technical or regulatory issues but also how they impact the very foundation of our society. We must ask ourselves what are some of the most important constitutional principles that govern us, such as Equality, Dignity, and the Right to Life. The way in which we deliver healthcare in India is deeply influenced by these principles under Articles 14 and 21 of the Constitution. Therefore, we cannot simply evaluate the deployment of AI systems from only the standpoint of innovation and efficiencies.

Rather, we must evaluate AI systems in terms of Constitutional Standards, which include Fairness, Reasonableness and Accountability. Furthermore, by undertaking a comparative study of both India's, the United States', and the United Kingdom's respective approaches towards AI and healthcare, we can identify not only Regulatory Gaps but also Normative Opportunities for Reformation.

8.1 Indian Legal Position: Constitutional Silences and Regulatory Fragmentation

There are currently no dedicated statutory provisions governing the use of artificial intelligence in healthcare in India. The absence of a comprehensive statutory framework for regulating AI-mediated medical harm means that the relevant legal frameworks governing this area are spread across tort law, consumer protection law, medical professional ethics, and data protection law. While existing legal frameworks do provide limited safeguards, these frameworks were never meant to regulate the area of medical decision making that is algorithmic, autonomous, opaque, and evolving.

This situation represents a significant gap in regulatory oversight from a constitutional

standpoint and raises serious issues regarding Article 21 of the Constitution of India. Article 21 guarantees Indian citizens the right to life and personal liberty and has been consistently interpreted by the Supreme Court of India to include the right to health, the right to access medical services, and the protection of human dignity. When AI systems impact the correct diagnosis of a patient, the prioritising of medical treatments for a patient, or the outcome of a surgical procedure, it is clear that failures to provide appropriate accountability directly impact the authority and interpretation of these constitutional issues.

Medical negligence law in India, as established by the case of Indian Medical Association - vs. V.P. Shantha -, assumes that there is a source of identifiable human agency and explainable reasoning for the actions taken in medical treatment. However, in the context of AI-assisted diagnosis and robotic surgery, the reliance on identifiable human agency and explainable reasoning is disrupted, making it difficult to identify sources of harm from algorithmic recommendations that even the best-trained clinicians cannot explain completely. As a result, traditional standards for determining negligence are not effective in pointing to a source of fault, presenting a constitutional violation due to a lack of adequate remedy, thereby violating the principle of substantive due process under Article 21.

Article 14 creates yet another complication for whether AI-based systems will lead to unfair results. Algorithm-trained on biased or unrepresentative datasets run greater risk of being inherently biased against vulnerable populations with a long history of limited access to healthcare. Additionally, AI will be used to discriminate against specific populations (i.e., women) and provide unfair treatment; therefore, the use of AI-based systems could produce arbitrary results that cannot be justified. Therefore, as the Supreme Court has ruled numerous times, Article 14 requires that discrimination based on the use of biased AI systems be construed as arbitrary discrimination.

The Digital Personal Data Protection Act, 2023 has advanced the understanding of personal informational autonomy for the first time in India. However, the Act fails to address automated decision-making, algorithmic explainability, and continuous learning. This lack of attention to these important issues may weaken the constitutional right of privacy granted to individual citizens by the Supreme Court in Justice K.S. Puttaswamy v. Union of India.

8.2 United States: Regulatory Oversight Without Constitutional Anchoring

The United States uses a sector-specific regulatory model, with the Food and Drug Administration (FDA) in charge of AI-based medical devices. The FDA's recognition of adaptive AI and its post-market monitoring show a practical understanding of technological risk. However, U.S. healthcare AI governance lacks a rights-based constitutional framework like India's Articles 14 and 21.

The Blueprint for an AI Bill of Rights[29] outlines principles such as non-discrimination, transparency, and human oversight. While these principles are important, they do not create enforceable legal rights. Consequently, accountability relies heavily on agency discretion and state tort law, which results in uneven protection.

For India, the U.S. experience highlights the importance of technical oversight and lifecycle regulation, but it also shows the limits of regulatory models that do not rest on constitutional guarantees of equality and dignity.

8.3 United Kingdom: Principles-Based Regulation and Institutional Accountability

The AI regulation principles established in the AI Regulation White Paper (2023)[30] by the United Kingdom focus on establishing basic ideals of safety, transparency, and accountability. Medical AI systems are governed by the MHRA and other sectoral regulators who focus primarily on a risk-based system of governance and adaptability.

While there is no legally binding provision in the UK framework, the emphasis placed on contestability and explainability fits well with the Indian Constitution's concept of enabling procedural fairness. The Indian Supreme Court has consistently stated that Procedural Fairness is a core tenet of Article 21. In this regard, the UK Model provides an excellent framework for strengthening procedural safeguards without hindering innovation.

8.4 Comparative Insights and Constitutional Lessons for India

Through comparative analysis, we can see that while AI governance in health care has not been perfected in any jurisdiction, India's Constitution has a framework with certain normative

[29] White House Office of Science and Technology Policy. (2022). *Blueprint for an AI Bill of Rights.*
[30] UK Department for Science, Innovation and Technology. (2023). *AI regulation: A pro-innovation approach.*

advantages. In particular, Articles 14 and 21 provide a foundation for a principled demand for transparency, non-discrimination and accountability for AI decision-making in the health care sector.

In the absence of statutory translation, constitutional principles may remain abstract. The challenge for India will be to develop enforceable regulatory frameworks based on these constitutional commitments within the context of medical AI.

9. Toward a Constitutionally Informed Hybrid Liability and Regulatory Model

9.1 Rationale for a Hybrid Approach

The distributed nature of AI decision-making requires moving away from purely fault-based liability models. However, an overly strict liability system may discourage innovation in a resource-limited healthcare system like India's.

A hybrid model based on constitutional principles provides a balanced solution. This model reflects the Supreme Court's views on reasonableness, proportionality, and fairness, which state that regulatory burdens should match the nature and severity of risk.

9.2 Key Elements of the Proposed Model

First, responsibility must be shared among doctors, hospitals, manufacturers, and developers based on how much control they have and their ability to prevent harm. This reflects Article 21's focus on real justice instead of just assigning blame.

Second, explainability should be a fundamental requirement rather than a technical option. Functional explainability allows for judicial review, helps patients make informed choices, and prevents arbitrary decisions, reinforcing Article 14.

Third, we should introduce mandatory algorithmic impact assessments, especially for high-risk medical AI systems. These assessments should look into bias, safety, and proportionality, similar to the constitutional test used in rights-based decisions.

Fourth, we need to create patient-focused compensation systems. No-fault compensation programs or mandatory AI insurance can provide timely remedies and meet the constitutional duty to offer effective relief for violations of rights.

9.3 Institutional Oversight and Policy Implementation

Establishing a specialized Medical AI Regulatory Authority would improve institutional accountability. This body could certify AI systems, require audits, and investigate negative outcomes. It would importantly connect constitutional norms with technology governance.

From a policy viewpoint, India needs to go beyond ethical guidelines and create enforceable standards. It is essential to include AI governance in public health policy, especially in government hospitals and insurance plans, to prevent weakening constitutional rights through technological outsourcing.

10. Conclusion

Though artificial intelligence has great potential to improve the Indian healthcare system, without sufficient oversight and regulation of its use, the ability of people to access equal treatment within the health care system (as mandated by Article 14) and to receive dignified, compassionate care (as protected by Article 21) may be impaired. In this paper, we contend that while technological challenges exist for AI in healthcare, there are fundamental constitutional issues related to its implementation. By linking AI governance to Articles 14 and 21 of the Constitution, India can develop an effective regulatory framework that encourages technological innovation while protecting individual rights. We propose a combination of liabilities (a mixed liability and regulatory model) and regulation pursuant to constitutional law, which has been shown to be effective. This model will create accountability and inspire progress as more algorithms are used to inform clinical decision making. Additionally, the law must adapt to address the use of AI in the delivery of healthcare whilst maintaining respect for the constitutional value of access to quality, compassionate healthcare. Only then can AI serve the public interest rather than inadvertently lead to greater injustice.

References

1. Abbott, R. (2020). Who's liable when AI causes harm? *Harvard Journal of Law & Technology, 33*(1), 1–55. https://jolt.law.harvard.edu/assets/articlePDFs/v33/Who-is-Liable-When-AI-Causes-Harm.pdf

2. Abbott, R. (2020). The reasonable computer: Disrupting the paradigm of tort liability. *George Washington Law Review, 86*(1), 1–44.

3. Bennett Moses, L., & Chan, J. (2018). Algorithmic prediction in policing: Assumptions, evaluation, and accountability. *Policing and Society, 28*(7), 806–822. https://doi.org/10.1080/10439463.2016.1253695

4. Bolam v. Friern Hospital Management Committee, [1957] 1 WLR 582 (UK).

5. Burrell, J. (2016). How the machine "thinks": Understanding opacity in machine learning algorithms. *Big Data & Society, 3*(1), 1–12. https://doi.org/10.1177/2053951715622512

6. Calo, R. (2015). Robotics and the lessons of cyberlaw. *California Law Review, 103*(3), 513–563.

7. Citron, D. K., & Pasquale, F. (2014). The scored society: Due process for automated predictions. *Washington Law Review, 89*, 1–33. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2376209

8. European Parliament and Council of the European Union. (2016). *General Data Protection Regulation (Regulation (EU) 2016/679).* Official Journal of the European Union.

9. European Parliament and Council of the European Union. (2024). *Artificial Intelligence Act.* Official Journal of the European Union.

10. Food and Drug Administration. (2019). *Proposed regulatory framework for modifications to artificial intelligence/machine learning-based software as a medical device.* U.S. Department of Health and Human Services.

11. Goodman, B., & Flaxman, S. (2017). European Union regulations on algorithmic decision-making and a "right to explanation." *AI Magazine, 38*(3), 50–57. https://doi.org/10.1609/aimag.v38i3.2741

12. Indian Medical Association v. V. P. Shantha, (1995) 6 SCC 651 (India).

13. Jacob Mathew v. State of Punjab, (2005) 6 SCC 1 (India).

14. Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., & Wang, Y. (2017). Artificial intelligence in healthcare: Past, present and future. *Stroke and Vascular Neurology, 2*(4), 230–243. https://svn.bmj.com/content/2/4/230

15. Kesan, J. P., & Hayes, C. M. (2019). Mitigative counterstrike: Self-driving cars, regulatory models, and tort law. *Iowa Law Review, 104*(5), 1799–1871.

16. Kusum Sharma v. Batra Hospital & Medical Research Centre, (2010) 3 SCC 480 (India).

17. Maneka Gandhi v. Union of India, AIR 1978 SC 597 (India).

18. Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society, 3*(2), 1–21. https://doi.org/10.1177/2053951716679679

19. NITI Aayog. (2021). *Responsible AI for all.* Government of India.

20. Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science, 366*(6464), 447–453. https://doi.org/10.1126/science.aax2342

21. Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information.* Harvard University Press.

22. Pasquale, F. (2019). *New laws of robotics: Defending human expertise in the age of AI.* Harvard University Press.

23. Price, W. N., II. (2017). Black-box medicine. *Harvard Journal of Law & Technology, 28*(2), 419–471.

24. Puttaswamy (Justice K. S.) v. Union of India, (2017) 10 SCC 1 (India).

25. Scherer, M. U. (2016). Regulating artificial intelligence systems: Risks, challenges, competencies, and strategies. *Harvard Journal of Law & Technology, 29*(2), 353–400.

26. Topol, E. (2019). *Deep medicine: How artificial intelligence can make healthcare human again.* Basic Books.

27. UK Department for Science, Innovation and Technology. (2023). *AI regulation: A pro-innovation approach.* HM Government.

28. Veale, M., Van Kleek, M., & Binns, R. (2018). Fairness and accountability design needs for algorithmic support in high-stakes public sector decision-making.

Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, 1–14. https://doi.org/10.1145/3173574.3174014

29. Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why a right to explanation of automated decision-making does not exist in the GDPR. *International Data Privacy Law, 7*(2), 76–99. https://doi.org/10.1093/idpl/ipx005

30. White House Office of Science and Technology Policy. (2022). *Blueprint for an AI bill of rights.* https://www.whitehouse.gov/ostp/ai-bill-of-rights/

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