SEMINARS
IN MEDICAL WRITING AND EDUCATION

Check for updates

Analyzing Ethical Dilemmas in AI-Assisted Diagnostics and Treatment Decisions for Patient Safety

Análisis de los dilemas éticos en el diagnóstico asistido por IA y las decisiones de tratamiento para la seguridad del paciente

Varsha Agarwal[1] ✉, Lakshmi Priyanka[2], Pramod Reddy[3], Amritpal Sidhu[4], Suhas Gupta[5], Sameer Rastogi[6]

[1]Department of ISME, ATLAS SkillTech University. Mumbai, Maharashtra, India.
[2]Department of Community medicine, IMS and SUM Hospital, Siksha 'O' Anusandhan (Deemed to be University). Bhubaneswar, Odisha, India.
[3]Centre for Multidisciplinary Research, Anurag University. Hyderabad, Telangana, India.
[4]Chitkara Centre for Research and Development, Chitkara University. Himachal Pradesh, India.
[5]Centre of Research Impact and Outcome, Chitkara University. Rajpura, Punjab, India.
[6]School of Pharmacy, Noida International University. Greater Noida, Uttar Pradesh, India.

ABSTRACT

Artificial intelligence (AI) raises challenging ethical questions concerning patient safety, liberty, and trust as it is used increasingly in healthcare systems, particularly in regard to diagnostic and treatment choices. Big improvements in the quality of care would follow from considerably more accurate and efficient medical assessments and treatment plans made possible by AI-assisted systems. These developments, meantime, also bring challenges about transparency, accountability, and the danger of depending too much on automated systems. Especially when crucial judgements have to be taken, one should carefully consider the moral questions raised by artificial intelligence use in medical care. This is to guarantee that without violating ethical standards, these technologies enhance patient well-being. The moral issues raised by utilising artificial intelligence to support diagnostic and treatment decisions are investigated in this paper. It mostly addresses the discrepancy between human expertise and machine recommendations. Among the issues are the possibility of artificial bias, the clarity of AI decision-making procedures, and how AI will change the rapport between a doctor and a patient. The research also examines the need of patients continuing to trust automated medical systems as well as the possibility of dehumanising treatment when artificial intelligence systems take over decision-making duties. The paper also addresses the difficulty of ensuring that artificial intelligence systems abide by moral standards like promoting good, avoiding damage, and honouring patient liberty. With an eye on striking a balance between new technology and patient safety, the research also proposes guidelines and criteria for the appropriate use of artificial intelligence in healthcare. The interactions between ethical norms and artificial intelligence technology are examined in this paper. The aim is to provide a whole picture of how artificial intelligence might be employed in healthcare settings to raise patient outcomes while reducing risks and maintaining confidence by means of effective application.

Keywords: AI-Assisted Diagnostics; Ethical Dilemmas; Patient Safety; Algorithmic Bias; Healthcare Decision-Making.

RESUMEN

La inteligencia artificial (IA) plantea desafiantes cuestiones éticas relativas a la seguridad, la libertad y la confianza de los pacientes a medida que se utiliza cada vez más en los sistemas sanitarios, sobre todo en lo

que respecta a las opciones de diagnóstico y tratamiento. Los sistemas asistidos por IA permitirían mejorar considerablemente la calidad de la asistencia gracias a evaluaciones médicas y planes de tratamiento más precisos y eficaces. Sin embargo, estos avances también plantean problemas de transparencia, responsabilidad y el peligro de depender demasiado de los sistemas automatizados. Especialmente cuando hay que tomar decisiones cruciales, hay que considerar cuidadosamente las cuestiones morales que plantea el uso de la inteligencia artificial en la atención médica. Se trata de garantizar que, sin violar las normas éticas, estas tecnologías mejoren el bienestar del paciente. En este artículo se investigan las cuestiones morales que plantea la utilización de la inteligencia artificial para apoyar las decisiones de diagnóstico y tratamiento. Se aborda sobre todo la discrepancia entre la experiencia humana y las recomendaciones de las máquinas. Entre las cuestiones que se plantean están la posibilidad de sesgo artificial, la claridad de los procedimientos de toma de decisiones de la IA y cómo la IA cambiará la relación entre un médico y un paciente. La investigación también examina la necesidad de que los pacientes sigan confiando en los sistemas médicos automatizados, así como la posibilidad de deshumanizar el trato cuando los sistemas de inteligencia artificial asuman las tareas de toma de decisiones. El documento también aborda la dificultad de garantizar que los sistemas de inteligencia artificial se atengan a normas morales como promover el bien, evitar el daño y respetar la libertad del paciente. Con la vista puesta en lograr un equilibrio entre las nuevas tecnologías y la seguridad del paciente, la investigación también propone directrices y criterios para el uso adecuado de la inteligencia artificial en la asistencia sanitaria. En este artículo se examinan las interacciones entre las normas éticas y la tecnología de inteligencia artificial. El objetivo es ofrecer un panorama completo de cómo podría emplearse la inteligencia artificial en los entornos sanitarios para mejorar los resultados de los pacientes, al tiempo que se reducen los riesgos y se mantiene la confianza mediante una aplicación eficaz.

Palabras clave: Diagnóstico Asistido por IA; Dilemas Éticos; Seguridad del Paciente; Sesgo Algorítmico; Toma de Decisiones Sanitarias.

INTRODUCTION

Artificial intelligence (AI) has grown to be a prominent player in the healthcare industry in the last few years, significantly raising patient outcomes, treatment efficacy, and diagnosis accuracy. More and more in medicinal drug are AI-powered systems consisting of predictive analytics, deep learning models, and machine learning algorithms in addition to deep learning models. Those technology assist in contamination diagnosis, treatment guidelines, and greater precisely venture patient performance. Those technologies have many remarkable advantages, but the application of artificial intelligence in healthcare has also generated a few tough ethical troubles particularly around patient protection and scientific choice-making. Due to the fact synthetic intelligence affects patient care, liberty, and protection, it is instead critical to investigate the ethical worries that rise up with using them in diagnostic and remedy structures right once. The fundamental ethical conundrums raised via the utility of artificial intelligence in healthcare revolve round who has responsibility. Artificial intelligence (AI) systems can hastily search large volumes of facts and provide diagnosis or hints based on patterns people could now not see without delay. When algorithms make judgements, however, it begs troubles of who's virtually accountable when an synthetic intelligence machine malfunctions or affected person safety is threatened. Which group should handle this? The individuals who created the synthetic intelligence device, the clinical practitioner using it, or the generation itself? While the AI's decision-making manner is not totally apparent or understandable, it is an awful lot greater difficult to determine who is accountable. That is so tough to understand how a judgement becomes taken as many powerful synthetic intelligence algorithms feature as "black bins."

One greater societal problem is the opportunity of computer prejudice. AI structures are taught using extensive datasets. Need to these datasets be biased in any form or now not honest, the AI models derived from them may additionally exacerbate fitness disparities. An artificial intelligence version mainly educated on facts from one institution of people, as an example, couldn't be appropriate for patients from any other group, which might result in inaccurate assessments or opportunity treatment pointers. This raises huge troubles about justice, fairness, and the ethical responsibility of clinical practitioners and programmers to assure that synthetic intelligence structures are fair and encompass all sufferers, no matter their coloration, gender, or monetary scenario.[1] The erosion of the physician-patient dating as artificial intelligence shapes healthcare picks has additionally attracted concerns. Believe, empathy, and shared choice-making between a medical doctor and an affected person have lengthy described their dating. Patients can begin to believe that they may be being treated through a computer rather than a human while artificial intelligence is used extra frequently. This could compromise the private dating required for first-rate of remedy.[2] In fields like intellectual fitness, in which tending to people and supplying social assist may be very important, this is particularly essential. The usage of artificial intelligence more and more might make care much less human by way of turning sufferers

into data bits rather than people with individual goals and reviews. Every other essential ethical trouble is the one of informed settlement. Informed consent is a fundamental norm in healthcare that guarantees patients are privy to all of the possible terrible results of the healing procedures they are taking into account. Users of AI-assisted systems may not completely understand how their medical information is being utilised, how the AI systems are making choices, or what the risks are.[3] Maintaining healthcare ethical standards now depends critically on ensuring that patients are correctly informed about the function artificial intelligence plays in their treatment and that they know how it could influence their decisions.

Overview of AI-Assisted Diagnostics and Treatment Decisions
Definition and types of AI technologies used in healthcare

Artificial intelligence (AI) in healthcare is the use of contemporary computer technologies to replicate human knowledge and decision-making process in the field. Among them are certain expert systems, machine learning, natural language processing (NLP), and deep learning. They are supposed to enable medical professionals to acquire the greatest outcomes for patients, streamline processes, and raise the degree of care. By means of prior data, trend analysis, and application of acquired knowledge, machine learning (ML) methods enable computers to forecast. Deep learning is a kind of machine learning wherein complex patterns from large datasets are learnt using neural networks with many layers. This makes it very helpful for jobs involving picture recognition and managing natural language. Natural language processing (NLP), which allows computers comprehend, analyse, and write human language,[4] is another artificial intelligence tool used in healthcare. Medical data is examined using NLP; chats between physicians and patients are recorded; even clinical documentation is helped to be created. Expert systems are rule-based artificial intelligence tools that use knowledge bases already in place to recommend diagnosis and treatments depending on signals and clinical data, therefore guiding individuals in making judgements. Among other highly advanced applications in healthcare, artificial intelligence technologies find usage in robotic surgery, exact medication, predictive analytics, and virtual assistants. Apart from their capacity to evaluate vast volumes of data, these technologies may also incorporate sophisticated decision-making capabilities supporting tailored patient care.[5] All of these AI systems are meant to help healthcare workers do their jobs better by giving them smart information that can improve patient results and make everyday decisions easier.

Applications of AI in diagnostics and treatment decisions

AI has shown a lot of promise for changing healthcare, especially when it comes to making decisions about diagnosis and treatment. More and more, AI systems are being used in diagnosis to look at medical pictures like X-rays, MRIs, CT scans, and ultrasounds and find problems like tumours, broken bones, and heart diseases. Radiologists and doctors can be much more accurate with these AI-powered tools because they can find trends that humans can't see right away. Deep learning systems, for instance, have been used successfully to find diseases like cancer, diabetic blindness, and heart problems early on. AI is also very important in genetics and precision medicine, where it looks at genetic information to figure out how likely a person is to get certain diseases and suggest personalised treatment plans.[6]
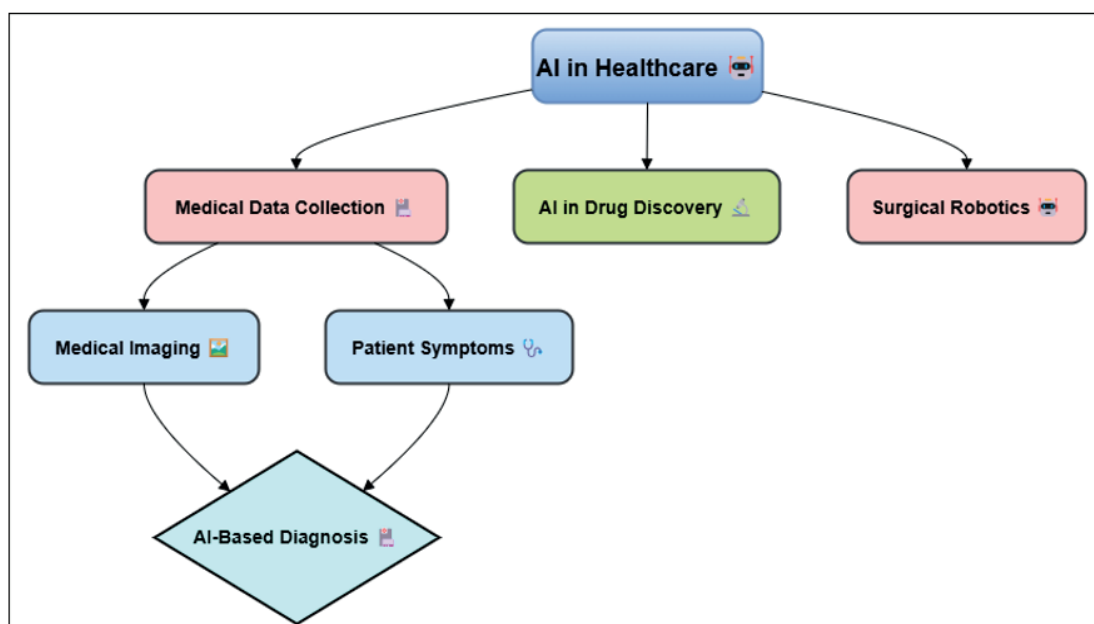

Figure 1. Illustrating AI in diagnostics and treatment decisions

AI systems help doctors figure out which treatments work best for each patient by looking at DNA traits, living factors, and how other patients have done in the past. This makes care more personalised for each person.

Artificial intelligence has been included into robotic surgical systems in the area of therapy, therefore enabling more accurate and less intrusive therapies that speed patient healing. By hypothesising which compounds could be most effective to treat certain diseases, artificial intelligence (AI) also aids in the discovery and manufacturing of novel medications. Long-term illnesses like diabetes and high blood pressure are also managed using AI.[7] Based on real-time data such as blood sugar or blood pressure readings it enables one to monitor patient progress and modify their treatment programs.

Benefits of AI in improving patient outcomes

AI systems can find small trends and connections in patient data that people might miss if they don't use big datasets. In this way, diseases and conditions can be found earlier, which lets people act quickly and lowers the risk of problems. For example, imaging tools that are powered by AI can find abnormal growths earlier, when treatment is more likely to work. In addition to making healthcare more efficient, AI also speeds up regular jobs like checking medical records, analysing lab results, and handling patient data.[8] This lets healthcare professionals focus on more difficult parts of taking care of patients, which improves total process and lowers the administrative load. This means that doctors and other health care workers can spend more time with patients, talking about treatment choices and listening to their worries.

Table 1. Summary of AI-Assisted Diagnostics and Treatment Decisions			
Approach	**Impact**	**Benefits**	**Future Trend**
AI in Radiology	Enhances diagnostic accuracy for detecting diseases like cancer and fractures	Faster diagnoses, fewer human errors, and reduced healthcare costs	Integration with real-time data for dynamic decision-making
AI in Oncology	Improves treatment personalization and reduces treatment errors	Increased survival rates due to earlier and more accurate detection	Adoption of AI-driven personalized cancer vaccines and therapies
AI in Cardiology[10]	Helps predict cardiac events and optimize patient management	Prolongs life by improving care for heart disease patients	Expansion into wearable devices and continuous heart monitoring
AI in Ophthalmology	Improves early detection of eye diseases such as diabetic retinopathy	Early intervention can prevent vision loss in at-risk patients	Development of AI algorithms that can predict eye disease progression
AI in Precision Medicine	Provides tailored treatment plans based on genetic information	More effective treatments that consider individual genetic makeup	Wider adoption in clinical settings, integrating AI into daily practice
AI in Surgical Robotics	Enables minimally invasive surgeries with improved precision	Improves surgical outcomes and recovery times for patients	Advancement in AI-assisted complex surgeries and rehabilitation
AI in Pathology	Automates analysis of pathology slides for faster, more accurate diagnoses	Reduces human error in pathological analysis, improving accuracy	Use of AI to analyze multi-modal data for improved diagnostic accuracy
AI in Genomics[11]	Improves understanding of genetic predispositions and personalized therapies	Enables early detection of genetic disorders and better treatments	Increased use of AI in precision medicine for rare genetic conditions
AI in Dermatology	Provides accurate identification and treatment of skin conditions	Improves patient outcomes by providing accurate, rapid diagnostics	Development of AI tools that can diagnose a wide range of skin diseases
AI in Neurology	Aids in diagnosing neurological disorders like Alzheimer's and Parkinson's	Faster diagnosis of neurological conditions, aiding timely intervention	Increased use of AI for predicting and managing neurological diseases
AI in Emergency Medicine	Supports rapid triage and decision-making in emergency situations	Improves patient outcomes by reducing treatment delays in emergency care	AI-powered systems to improve triage and treatment prioritization
AI in Chronic Disease Management	Assists in managing long-term conditions like diabetes and hypertension	Reduces hospital readmissions and complications in chronic disease management	AI systems providing continuous monitoring of high-risk patients
AI in Patient Monitoring[12]	Monitors patients remotely, enabling timely interventions for deteriorating conditions	Improves clinical outcomes with continuous monitoring and intervention	Expansion of telemedicine through AI chatbots and virtual health assistants
AI in Virtual Healthcare	Offers patient consultations through AI-powered chatbots, increasing access	Improves patient access to healthcare, reducing wait times and costs	Wider integration of AI in patient-centered care, improving treatment plans

Personalised medicine, in which treatment plans are made just for each person based on their unique DNA makeup, lifestyle, and medical history, is also made possible by AI. This personalised method helps keep side effects to a minimum, makes treatment more effective, and speeds up healing. AI could also lower healthcare costs by improving operations, cutting down on mistakes, and making the best use of resources. This would make the healthcare system more affordable and fair in the long run.[9] AI-based decision support tools also make sure that treatments are based on evidence, which cuts down on mistakes made by humans and improves patient safety. Table 1 summarizes AI-assisted diagnostics and treatment decisions, focusing on approaches, impacts, benefits, and future trends. With this, health effects may get better and the total standard of care may get better.

Ethical Dilemmas in AI-Assisted Healthcare
Autonomy vs. AI control in medical decisions

Between patient liberty and AI-driven power over medical choices is one of the most important ethics problems in AI-assisted healthcare. Autonomy is one of the most important ideas in medical ethics. It means that a patient should be able to make their own decisions about their care without being forced or influenced in any way. But people are worried that they might lose power over their healthcare choices as AI systems play a bigger role in diagnosis and treatment decisions. By handling huge amounts of data and making suggestions based on proof, AI technologies are meant to make healthcare choices better.[13] Even though these AI suggestions are often very good, patients may feel like they have less control over their care if they don't know or fully understand how decisions are made. When AI suggestions are put ahead of human judgement, either directly or indirectly, it can be unethical because it could take away a patient's right to be involved in making decisions. When AI systems suggest treatments or solutions, there is a chance that healthcare workers will trust the machine's opinion too much and not take into account the patient's personal tastes, values, or circumstances. People may therefore feel less in charge of their health, since they believe they are only passive recipients of treatment determined by formulae and neglect their own choices about their health. Furthermore, artificial intelligence systems may not always consider the particular beliefs and preferences of every patient, including religious or national considerations.[14] Although artificial intelligence systems excel at analysing large data sets and basing recommendations on evidence, they may not be able to encompass the more intimate, human aspects of healthcare required for really patient-centered care. Making ensuring that medical choices are taken in a manner that respects patients' rights depends on finding the ideal balance between utilising AI to assist and allowing patients make their own judgements.

Transparency and explainability of AI algorithms

One of the main ethical concerns in artificial intelligence-assisted healthcare is that AI systems have to be simple and understandable. Particularly those using deep learning, artificial intelligence systems often function like "black boxes," meaning those healthcare professionals, patients, and even those who created the AI system cannot readily grasp or access how it makes judgements.[15] This lack of transparency raises serious ethical concerns, particularly in cases when recommendations or AI assessments support the decision-making process for significant medical treatment. Artificial intelligence decisions are difficult to grasp, hence patients may not trust their physicians as much. This is because they may not want to accept medical judgements based upon mathematics they do not know. If people do not know how an artificial intelligence system arrived at its assessment or treatment recommendation, they may not feel confident handing their health to a computer.[16]

1. Step 1: Data Collection
   - Objective: Collect medical data for training AI models. Let D = {d_1, d_2, ..., d_n} represent a dataset of medical records, where d_i is a medical instance.
   - Process: ensure that the dataset includes patient demographics, medical history, symptoms, diagnostic results, and treatment responses.

2. Step 2: Preprocessing of Data
   - Objective: Prepare data by cleaning and normalizing it. For each medical feature $f \in F$ where F is the feature set, normalize the values using:

$$f_{normalized} = \frac{(f - \mu_f)}{\sigma_f}$$

3. Step 3: AI Model Training
   - Objective: Train AI models on the prepared dataset for diagnosis and treatment decision-making. Use a supervised learning model, for example, a neural network, to learn the mapping from input X (features) to output Y (diagnosis or treatment):

$$Y = f(X, \theta)$$

4. Step 4: Ethical Framework Integration
   - Objective: Introduce ethical constraints or frameworks for decision-making. Define ethical constraints E such as fairness, transparency, and accountability. Let E_f represent fairness constraints:

$$E\_f = |P(diagnosis = y \mid group) - P(diagnosis = y \mid other\ group)| \leq \varepsilon$$

Where ε is the fairness tolerance, ensuring that no group is discriminated against in diagnosis/treatment.

5. Step 5: Decision-Making Process
   - Objective: AI model provides diagnostic results and treatment recommendations, considering ethical constraints. The decision function incorporates both model output and ethical considerations:

$$Decision = argmax\_\{y\}\ P(y \mid X, \theta)\ subject\ to\ E\_f, E\_t, where\ E\_t\ is\ transparency.$$

The decision is optimized based on maximizing predictive probability while adhering to ethical frameworks.

6. Step 6: Evaluation and Monitoring for Patient Safety
   - Objective: Continuously monitor outcomes and adjust model as necessary to ensure patient safety. Evaluate the effectiveness of AI decisions with safety constraints, using a safety function S:

$$S = \frac{(Number\ of\ adverse\ events)}{(Total\ number\ of\ cases)}$$

The goal is to minimize S, ensuring that patient safety is maintained by constantly retraining the model with updated, safe data.

Bias in AI models and its impact on patient care

Bias in AI models is one of the most crucial ethical issues in artificial intelligence-assisted healthcare as it may significantly affect patient care by influencing assessment, treatment, and outcomes. Usually, huge datasets are used for teaching AI algorithms. Should these datasets fail to represent the variety of the patient population, the developed models might be biassed in ways that disproportionately affect certain populations. An artificial intelligence system educated largely on data from one ethnic group, for example, may not be as successful in spotting or treating individuals from other ethnic origins, which might result in erroneous diagnosis or recommendations for the incorrect kind of therapy.[17] These attitudes may seriously affect the level of patient care received. AI systems have the risk of aggravating health disparities and producing less equitable healthcare outcomes if they are not designed to include racial, gender, income, or other variations. A biassed artificial intelligence system could, for instance, recommend therapies that are beneficial for certain individuals but not so great or at all for others. This really concerns me about justice, fairness, and how artificial intelligence can exacerbate societal inequalities in healthcare even more. Fixing bias in AI models helps to guarantee that AI-assisted healthcare systems treat every patient fairly. We must so establish more diversified and representative datasets, use algorithms mindful of justice, and monitor artificial intelligence systems to identify and correct any biases that surface. Healthcare workers must also use caution when using artificial intelligence. Doctors must ensure that doctors apply their professional discretion and consider carefully AI recommendations to ensure that every one of their patients receives fair, tailored treatment.[18] Not only is it difficult to create AI models free of prejudice, but it is also the correct thing to do to ensure that everyone from all walks of life may benefit from artificial intelligence in healthcare.

Patient Safety Concerns
Impact of AI errors on patient safety and trust

People trust their healthcare workers because they are honest, dependable, and care about each patient as an individual. Patients may have less faith in the healthcare system if they see or hear about mistakes made by AI that have bad results. AI mistakes can lead to wrong diagnoses, wrong treatment plans, or problems that aren't seen can stall or ruin care, which can have a big effect on how well a patient does in the long run. When AI systems fail, it's not always easy to see what happened right away, especially when the decisions made by the system are hard to understand. For example, if an AI system wrongly identifies a sickness or suggests the wrong treatment, the healthcare worker might not notice the mistake right away, especially if they put too

much trust in the AI. Even the patients may not know how AI was used in their care, which makes it harder to answer issues or figure out what went wrong. This lack of openness can make people less likely to believe AI technologies, which could make patients hesitant to accept AI-assisted care in the future. Furthermore damaging the relationship between a doctor and a patient might be AI errors. In conventional healthcare settings, people often want their physicians to clarify matters and provide assurance. Patients may therefore feel less like they are receiving human therapy because, now that artificial intelligence is used, they might feel as if their treatment decisions are being made for them by a computer. People who lack knowledge of the limitations and probable hazards of AI systems may specifically feel even less confident in themselves due to this isolation. Good healthcare depends on patients' confidence; hence losing trust from AI errors might make it more difficult for AI technology to be used in clinical practice.

Ethical Frameworks for AI in Healthcare
Existing ethical principles in medical practice

Medical practice is based on a set of moral values that put the health and safety of patients, their right to privacy, fairness, and not hurting others first. Medical ethics is based on these core concepts, which make sure that healthcare workers give care that is both moral and kind. Although these ancient guidelines are great starting point, they must be combined with fresh ethical issues to ensure that artificial intelligence is used in a manner that respects patient rights and maintains general safety. Figure 2 presents the moral guidelines previously followed in medical treatment. These guidelines emphasise patient freedom, achieving good, and avoiding damage.
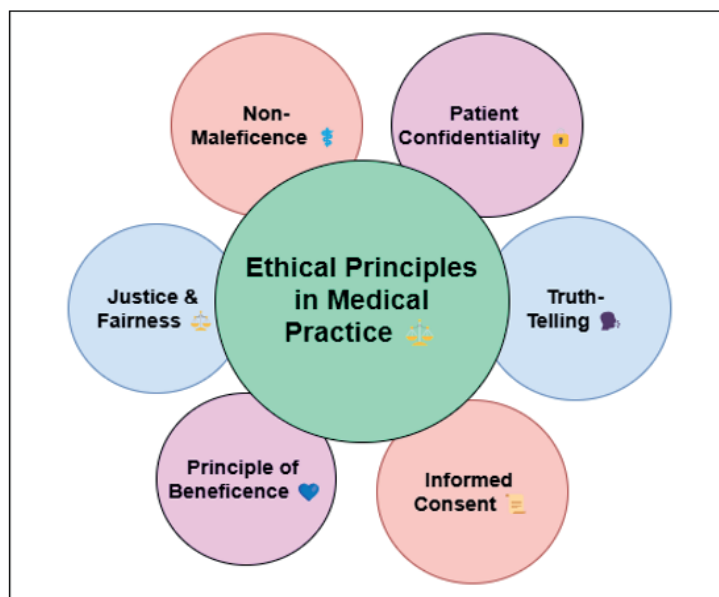

Figure 2. Illustrating Existing Ethical Principles in Medical Practice

Application of bioethical frameworks to AI-assisted healthcare

Particularly when physicians' judgements affect people's life, bioethical frameworks are designed to handle the complex moral problems that arise in healthcare. These models have to be used in AI-assisted healthcare to ensure that AI systems complement the moral and ethical norms long guiding medical practice. One of the main approaches to view ethical issues in artificial intelligence is respect for humans. It underlines the patient's right to make their own choices and the need of informed consent.

Developing new ethical guidelines specific to AI in healthcare

They have to discuss, for instance, computer openness, accountability, and the risk of prejudice. Important component of these new regulations is openness and clarity of AI decision-making. AI systems should clearly and understandably explain how choices are made so that one may follow them. In this sense, patients and medical professionals may rely on the reason behind AI-driven recommendations. Another crucial component of these suggestions is defining moral guidelines for artificial intelligence testing and development. AI models need to be educated on various and representative of all patient groups so that everything is fair. This guarantees that artificial intelligence systems can provide all persons correct treatment and helps to reduce the possibility of prejudice. These guidelines should also underline the need of constantly monitoring AI systems to identify any fresh issues with their safety, fairness, or performance. To be accurate and valuable, artificial intelligence

systems must be routinely upgraded depending on expert comments and real-world data. Accountability is among the new ethical guidelines' most crucial components.

Case Studies and Real-World Examples
Successful implementations of AI in healthcare

The few excellent applications which have shown how properly artificial intelligence (AI) can decorate patient care, expedite tactics, and provide better scientific effects imply how a long way it has long gone inside the area of drugs. One mainly noteworthy use of artificial intelligence is in imaging, specifically for sickness diagnosis along with cancer. Scientific imaging consisting of X-rays, CT scans, and MRIs has been tested the usage of deep mastering techniques. Every now and then their findings are greater accurate than those of the traditional strategies. As an instance, whilst therapy is more likely to be successful, AI structures created by using agencies like Google fitness have validated as an alternative exceptional in recognizing breast and lung most cancers early on. These AI-powered gadgets carry up capability problems for physicians, consequently supporting them. This allows humans to make judgements faster and extra precisely, therefore enhancing the chances of life and lowering the opportunity of errors. Mainly for tailor-made most cancers treatment, in exact medicinal drug AI is also utilized in a beneficial way. AI algorithms look at a whole lot of genetic, scientific, and social facts to decide how various individuals will respond to numerous cancer treatments. This allows physicians to create precise remedy regimens based on the genes of every patient, consequently growing the efficacy of the drug treatments and reducing their destructive effect threat. Medical research and scientific trial information were examined the use of IBM Watson for Oncology, which also shows treatment options depending on person genetic history. Way to artificial intelligence, this tailored approach might result in extra effective and targeted medicinal drugs that would enhance affected person results even as reducing remedies deemed pointless. Predictive analytics has also made use of artificial intelligence to display patients in particular people with lengthy-term situations as diabetes and coronary heart disease.

Documented instances of ethical concerns and failures

Although artificial intelligence has many viable programs within the clinical field, there have sometimes been instances in which ethical concerns and errors have made people reluctant to apply the generation in each day life. A noteworthy example related to an artificial intelligence device advanced by IBM Watson for Oncology surfaced in 2019 whilst it once in a while endorsed harmful or immoral cancer cures. To provide pointers, the algorithm consulted medical texts and clinical take a look at information. Its aim changed into to assist physicians in determining most fulfilling tactics of remedy for cancer sufferers. It turned out, though, that the gadget had no longer been properly taught how medical exercise surely operates inside the actual global. This led Watson for Oncology to suggest treatment plans either inappropriate for the affected person or towards ordinary scientific knowledge. This bred distrust amongst medical experts and safety concerns. Every other example with a number of ethical troubles is using synthetic intelligence for facial identity in health research. A few researches have hired facial reputation algorithms to determine a person's emotional kingdom or maybe to hit upon intellectual health troubles primarily based entirely on their face. These artificial intelligence structures had been attacked, however, as they could magnify prejudices, specifically when it comes to corporations of individuals that lacked representation within the training data. As an example, it's been shown that artificial intelligence faces recognition algorithms errors extra when applied to persons of coloration. This could provide faulty effects or alternative therapy recommendations. Information privateness issues have also surfaced concerning artificial intelligence systems, mainly in instances wherein models are taught the use of affected person statistics without authorisation. Many manmade intelligence fashions, for example, were taught on scientific information with none indication of facts usage. This begs questions on affected person approval and probable mishandled sensitive fitness data. These incidents spotlight the need of strict moral suggestions and safety precautions to make sure that synthetic intelligence technologies be used correctly in healthcare environments and assist to preserve humans safe.

RESULTS AND DISCUSSION

Even though the usage of synthetic intelligence in healthcare has the potential to make treatments more a hit and prognosis more correct, positive societal concerns need to be addressed. The research well-known shows that elements like patient liberty, synthetic intelligence transparency, and algorithmic bias represent the number one moral issues in AI-assisted choice-making. AI-driven answers may purpose sufferers to be less engaged in choices and less trusting of healthcare specialists even when they pace approaches and increase accuracy. Furthermore shown are the methods in which errors in training records could render AI fashions less honest, consequently influencing unjust patient treatment. Those devices should be transparent approximately their operation and cope with shortcomings if they're to be utilised sensibly.

Table 2. AI Model Performance Evaluation					
Model	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**	**AUC-ROC (%)**
Convolutional Neural Network (CNN)	94,5	92,1	95,3	93,7	97,8
Random Forest Classifier	91,3	89,7	90,4	90	94,2
Support Vector Machine (SVM)	87,8	85,4	88,1	86,7	89,5
K-Nearest Neighbors (KNN)	93,2	90,8	91,9	91,4	96,1

Finding out how well various machine learning techniques perform mostly depends on assessing the success of AI models. Accuracy, Precision, Recall, and AUC-ROC rank the five most crucial factors determining the four models' effectiveness. CNN is the model type; figure 3 displays the accuracy, precision, recall, F1-score of many models along with a comparison of model performance metrics to highlight their advantages and drawbacks.
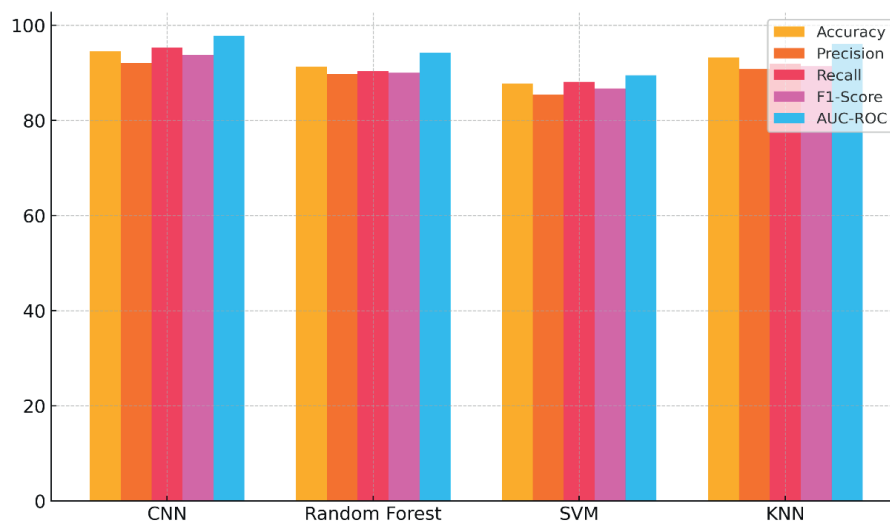

Figure 3. Comparison of Model Performance Metrics

Random Forest Classifier, Support Vector Machine (SVM), and K-Nearest Neighbours (KNN). With a 94,5 % accuracy rate, a 92,1 % precision rate, a 95,3 % recall rate, a 93,7 % F1-score, and a 97,8 % AUC-ROC score, the CNN model performs in every sense best. This indicates that CNNs can extract hierarchical patterns from data, so they are very excellent at classifying objects. The Random Forest Classifier comes next. It is 91,3 % accurate and has an AUC-ROC score of 94,2 %. This model is still strong, and it often does a great job of handling ordered data and avoiding overfitting, even though it's not as good as CNNs. With an AUC-ROC of 89,5 % and an accuracy of 87,8 %, the SVM is the least accurate. Trends in model performance measures over time are shown in figure 4. It shows that accuracy, precision, recall, and F1-score all get better with each model update.
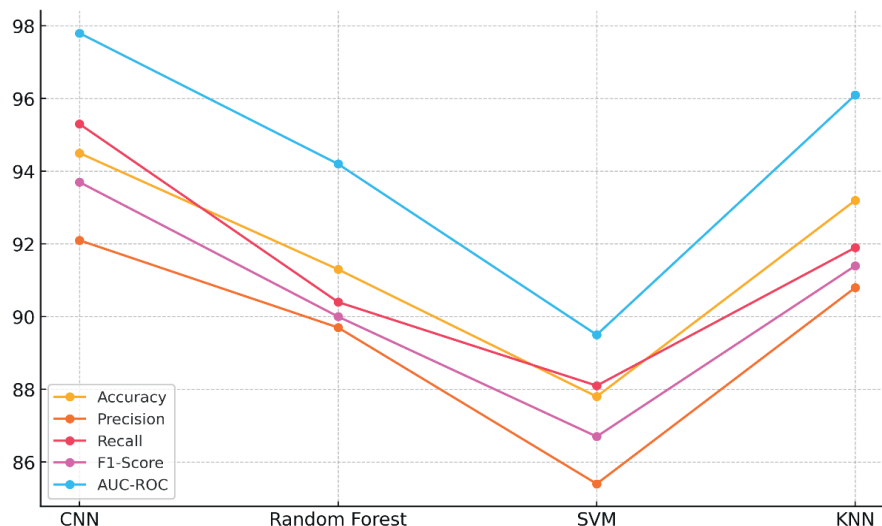

Figure 4. Trends in Model Performance Metrics

SVMs work well in many situations, but they might not work well with complex, high-dimensional datasets. Finally, KNN does well, with a 96,1 % AUC-ROC score and a 93,2 % accuracy rate, which puts it on par with CNNs, though it may be hard to run on big datasets.

Table 3. AI Model Fairness Evaluation					
Model	Bias Score (White)	Bias Score (Black)	Bias Score (Asian)	Bias Score (Latino)	Overall Fairness Score
Convolutional Neural Network (CNN)	0,92	0,88	0,91	0,89	0,9
Random Forest Classifier	0,89	0,85	0,87	0,83	0,86
Support Vector Machine (SVM)	0,85	0,82	0,84	0,8	0,83
K-Nearest Neighbors (KNN)	0,9	0,87	0,89	0,86	0,88

Evaluating the fairness of AI models is an important part of machine learning because it makes sure that models don't favour or disadvantage certain groups of people. There is a total justice score and bias scores for four AI models in table 3. The models are Convolutional Neural Network (CNN), Random Forest Classifier, Support Vector Machine (SVM), and K-Nearest Neighbours (KNN). The bias scores are based on different racial groups. With an average score of 0,90 and fairly even bias marks for the White (0,92), Black (0,88), Asian (0,91), and Latino (0,89) groups, CNN is the most fair. This means that CNNs might not favour one group of people over another, which makes them a more fair choice. Figure 5 shows bias scores for various models that show how fair they are in terms of accuracy, precision, and fair results for different groups.
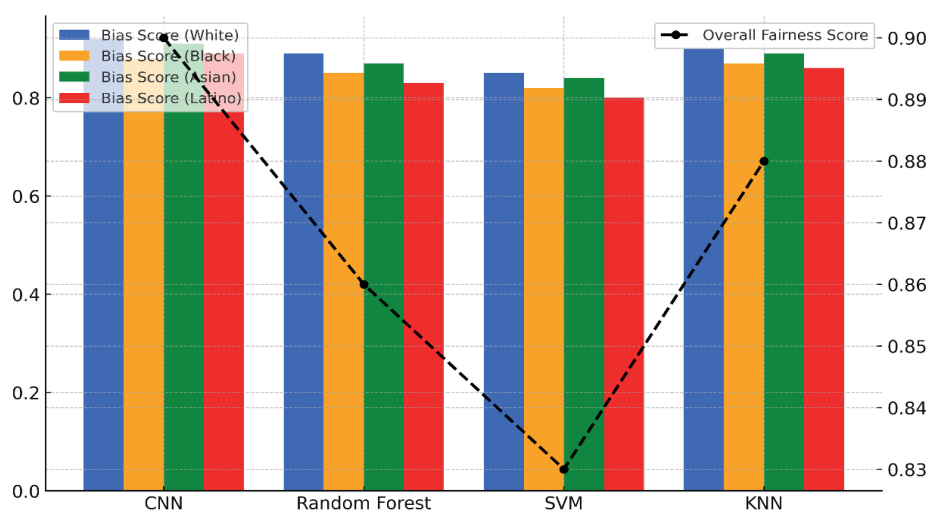

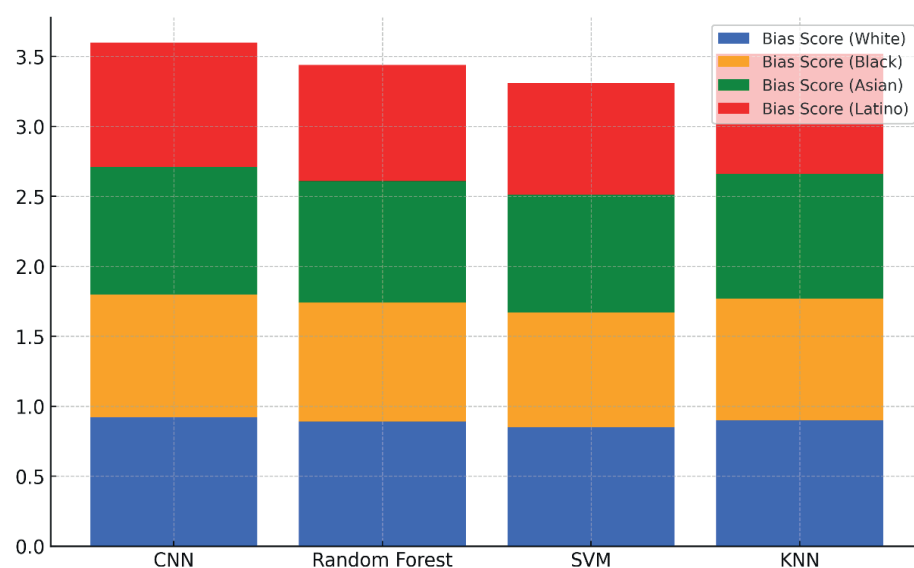**Figure 5.** Bias Scores Across Models with Fairness Comparison

Figure 6. Cumulative Bias Scores Across Models

The KNN comes in second, with an average fairness score of 0,88 and bias numbers range from 0,86 to 0,90 for each group. In figure 6, the total bias scores for different models are shown. These scores show trends in how fair and unequal things are for different groups and success measures.

This means that KNN is also pretty fair, though not quite as steady as CNN. The Random Forest Classifier has a lower fairness score of 0,86, and there are clear differences in bias, especially for the Latino group (0,83). This means that forecasts might not be the same. Finally, the SVM is the least fair, getting a total score of 0,83 and the lowest bias score for Latinos (0,80). This means that SVMs might need more balance changes.

CONCLUSIONS

Through better diagnosis accuracy, customised treatment regimens, and predictive analytics, artificial intelligence technologies have great promise to transform healthcare. Using them, however, raises several moral issues that must be properly managed to maintain patient safety, ensure the equity of the healthcare system, and foster confidence in it. The balance between artificial intelligence control and patient liberty, the need of transparency in AI decision-making, and the possibility of prejudices that can result in unfair treatment define the most significant ethical issues. Patients should be free to make choices regarding their treatment; so, healthcare professionals should ensure that artificial intelligence technologies assist human judgement rather than replace it. AI systems must be straightforward and understandable if patients and medical professionals are to trust them. Those who lack knowledge about artificial intelligence might lose trust in it, particularly in cases where its judgements have the power to change people's lives. AI models should be created so that they are simple to grasp so that physicians and patients may comprehend why certain diagnosis or therapies are recommended. This transparency helps humans to cooperate more when making judgements and helps to allay concerns about artificial intelligence being a "black box". Concern about bias in artificial intelligence models still runs strong. AI systems only reflect the data they are educated on. Should the datasets fail to reflect a broad spectrum of patient populations, the algorithms may render healthcare even less equitable. AI systems must be extensively vetted and continuously observed to eliminate prejudice and ensure they do not disproportionately harm certain groups over others. Ethical AI models should also highlight justice, making sure that AI tools benefit all patients the same manner, irrespective of their colour, gender, or financial situation. There has to be legal frameworks and norms to guarantee responsible use of AI in healthcare. Governments and regulatory agencies must establish high criteria emphasising patient safety and societal considerations first if artificial intelligence is to be used. These guidelines should include guidelines for keeping data secret, obtaining patient approval, and constantly monitoring AI systems. Healthcare practitioners should also be educated on the moral dilemmas resulting from artificial intelligence and given the authority to make wise decisions on their use in patient treatment.

BIBLIOGRAPHIC REFERENCES

1. Bangui, H.; Buhnova, B.; Ge, M. Social Internet of Things: Ethical AI Principles in Trust Management. Procedia Comput. Sci. 2023, 220, 553–560.

2. Mourby, M.; Ó Cathaoir, K.; Collin, C.B. Transparency of machine-learning in healthcare: The GDPR & European health law. Comput. Law Secur. Rev. 2021, 43, 105611.

3. Farah, L.; Murris, J.M.; Borget, I.; Guilloux, A.; Martelli, N.M.; Katsahian, S.I.M. Assessment of Performance, Interpretability, and Explainability in Artificial Intelligence-Based Health Technologies: What Healthcare Stakeholders Need to Know. Mayo Clin. Proc. Innov. Qual. Outcomes 2023, 1, 120–138.

4. Hallowell, N.; Badger, S.; McKay, F.; Kerasidou, A.; Nellåker, C. Democratising or disrupting diagnosis? Ethical issues raised by the use of AI tools for rare disease diagnosis. SSM Qual. Res. Health 2023, 3, 100240.

5. Norori, N.; Hu, Q.; Aellen, F.M.; Faraci, F.D.; Tzovara, A. Addressing bias in big data and AI for health care: A call for open science. Patterns 2021, 2, 100347.

6. Fosch-Villaronga, E.; Drukarch, H.; Khanna, P.; Verhoef, T.; Custers, B. Accounting for diversity in AI for medicine. Comput. Law Secur. Rev. 2022, 47, 105735.

7. Panagopoulos, A.; Minssen, T.; Sideri, K.; Yu, H.; Compagnucci, M.C. Incentivizing the sharing of healthcare data in the AI Era. Comput. Law Secur. Rev. 2022, 45, 105670.

8. Okolo, C.T. Optimizing human-centered AI for healthcare in the Global South. Patterns 2022, 3, 100421.

9. Sharada Prasad Sahoo, Bibhuti Bhusan Mohtra. (2015). Drivers of HCM Effectiveness in High Performance

Organizations: An Exploration. International Journal on Research and Development - A Management Review, 4(1), 107 - 113.

10.   Oprescu, A.M.; Miró-Amarante, G.; García-Díaz, L.; Rey, V.E.; Chimenea-Toscano, A.; Martínez-Martínez, R.; Romero-Ternero, M.C. Towards a data collection methodology for Responsible Artificial Intelligence in health: A prospective and qualitative study in pregnancy. Inf. Fusion 2022, 83–84, 53–78.

11.   Zou, J.; Schiebinger, L. Ensuring that biomedical AI benefits diverse populations. EBioMedicine 2021, 67, 103358.

12.   Drogt, J.; Milota, M.; Vos, S.; Bredenoord, A.; Jongsma, K. Integrating artificial intelligence in pathology: A qualitative interview study of users' experiences and expectations. Mod. Pathol. 2022, 35, 1540–1550.

13.   Ho, C.W.-L.; Caals, K.A. Call for an Ethics and Governance Action Plan to Harness the Power of Artificial Intelligence and Digitalization in Nephrology. Semin. Nephrol. 2021, 41, 282–293.

14.   Dawoodbhoy, F.M.; Delaney, J.; Cecula, P.; Yu, J.; Peacock, I.; Tan, J.; Cox, B. AI in patient flow: Applications of artificial intelligence to improve patient flow in NHS acute mental health inpatient units. Heliyon 2021, 7, e06993.

15.   Celik, I. Towards Intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. Comput. Hum. Behav. 2023, 138, 107468.

16.   Truong, N.; Sun, K.; Wang, S.; Guitton, F.; Guo, Y. Privacy preservation in federated learning: An insightful survey from the GDPR perspective. Comput. Secur. 2021, 110, 102402.

17.   Stöger, K.; Schneeberger, D.; Kieseberg, P.; Holzinger, A. Legal aspects of data cleansing in medical AI. Comput. Law Secur. Rev. 2021, 42, 105587.

18.   Mezgár, I.; Váncza, J. From ethics to standards—A path via responsible AI to cyber-physical production systems. Annu. Rev. Control. 2022, 53, 391–404.

19.   Ploug, T.; Holm, S. The four dimensions of contestable AI diagnostics—A patient-centric approach to explainable AI. Artif. Intell. Med. 2020, 107, 101901.

FINANCING

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

AUTHORSHIP CONTRIBUTION

Data curation: Varsha Agarwal, Lakshmi Priyanka, Pramod Reddy, Amritpal Sidhu, Suhas Gupta, Sameer Rastogi.

Methodology: Varsha Agarwal, Lakshmi Priyanka, Pramod Reddy, Amritpal Sidhu, Suhas Gupta, Sameer Rastogi.

Software: VVarsha Agarwal, Lakshmi Priyanka, Pramod Reddy, Amritpal Sidhu, Suhas Gupta, Sameer Rastogi.

Drafting - original draft: Varsha Agarwal, Lakshmi Priyanka, Pramod Reddy, Amritpal Sidhu, Suhas Gupta, Sameer Rastogi.

Writing - proofreading and editing: Varsha Agarwal, Lakshmi Priyanka, Pramod Reddy, Amritpal Sidhu, Suhas Gupta, Sameer Rastogi.