Journal of Computing & Biomedical Informatics
ISSN: 2710 - 1606
Volume 10 Issue 01
2025

Review Article
https://doi.org/10.56979/1001/2025/1139

Ethical Considerations of Large Language models in medical practice: A Review in Healthcare

Ramsha Saeed[1], Rabia Tehseen[1,2*], Anam Mustaqeem[3], Muhammad Waqas Nadeem[4], Usman Aamer[1], and Jawad Hassan[5]

[1]Department of Computer Science, University of Central Punjab, Lahore, 54782, Pakistan.
[2]Department of Computer Science, University of Management & Technology, Lahore, 54782, Pakistan.
[3]Department of Software Engineering, University of Central Punjab, Lahore, 54782, Pakistan.
[4]Faculty of Information and Communication Technology, Universiti Tunku Abdul Rahman, Kampar, Perak, Malaysia.
[5]Department of Applied Computing Technologies, University of Central Punjab, Lahore, 54782, Pakistan.
[*]Corresponding Author: Rabia Tehseen. Email: rabia.tehseen@ucp.edu.pk

Received: August 31, 2025 **Accepted:** November 25, 2025

Abstract: The introduction of large language models (LLM) into healthcare has attracted acute ethical issues, as well as the possibilities of enhancing clinical decision making and patient care. The article in question comprises a systematic literature review (SLR) of ethics concerns regarding the use of LLM in healthcare and the problems of bias, transparency, accountability and confidentiality. We consider those published since 2016 and 2024, interpret the ethical aspects of the LLM in different medical uses, such as clinical decision support, interaction with patients and medical research. The work of this review has a rich taxonomy of ethical concerns, exploration of a gap in the existing literature, and a recommendation on how responsible use of LLM may be applied in health care. The purpose of this article is to give a reflection to the healthcare practitioners and decision-makers regarding the ethical issues of incorporating the use of LLM in clinical practice.

Keywords: LLM; Medicine; Healthcare; NLP; ML; DL; Generative AI; Medical AI

1. Introduction

The role played by Artificial Intelligence (AI) in the field of healthcare has brought considerable progress, especially the use of large language models (LLMs). The models have proven their ability to perform impressive clinical decision-making, medical research, and patient care capabilities including natural language processing (NLP), named entity recognition, and clinical decision support. The capacity of the LLM to work through extensive medical information, write like humans, and make decisions is a big potential to enhance the delivery of healthcare [1]. Nevertheless, in addition to the fact that they can transform healthcare, LLMs also present dangerous ethical issues that are to be taken into consideration prior to their mass application in clinical practice [2].

Concerns like bias, transparency, privacy and accountability are major challenges in the incorporation of the LLC into the healthcare system. To illustrate, inappropriate treatment recommendations or misdiagnosis can be caused by bias in LLM particularly when the training data lacks representativeness of various patient populations [3]. It is also significant to ensure transparency because physicians and patients should be aware of how AI models make decisions that change health outcomes [4]. In like fashion, the matter of patient privacy is also a problem, since, based on the sensitive medical information, LLCs will produce some data security problems and conflicts with law, including HIPAA and GDPR [5]. Moreover, inadequate accountability strategies complicate the process of locating responsibility when medical diagnosis or treatment suggestions by LLMs are erroneous [6]. These are the ethical issues that are to be resolved so that the implementation of LLM can be implemented into clinical practice safely and responsibly.

The article is a systematic literature review (SLR) of ethical evaluations associated with the utilization of LLM in the health system. The review synthesises the research published since 2016 and presents a current issue with the bias, transparency, accountability and privacy as one of the main ethical issues to address, as well as provides a thorough analysis of the challenges and solutions suggested in the literature [7]. This is to give a systematic guide to these ethical issues, suggest ways of mitigating the risk as well as the future research directions.

Some of the input of this review is the creation of a taxonomy of ethical issues in the field of LLM healthcare, exploration of gaps in the existing literature and a series of practical recommendations to practitioners, researchers and policy makers [8]. By so doing, the article attempts to inform the ethically sound application of LLMs in healthcare in a manner that the quite potent tools are applied ethically to enhance patient care and safeguard the patient rights and justice in healthcare [9].

Similarly, the trends have been addressed in figure 1 on parallel basis with reference to LLMs.

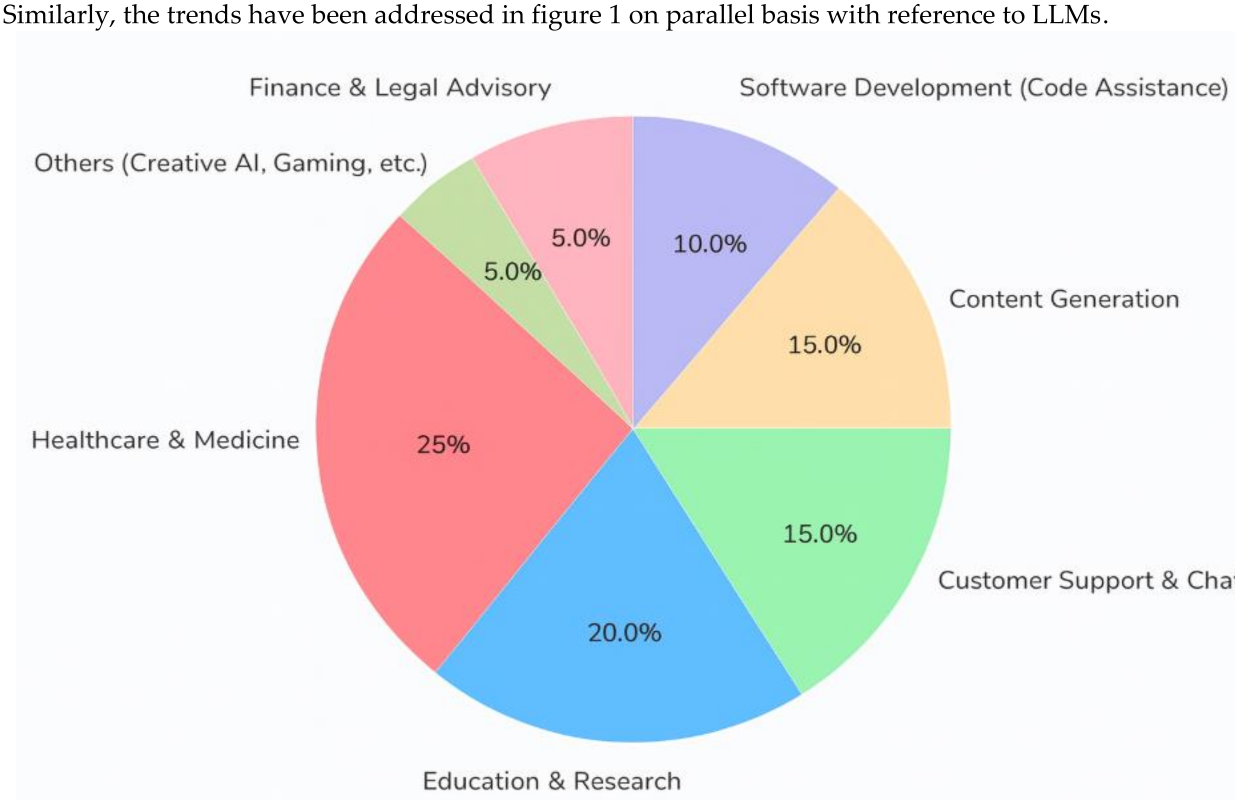

Figure 1. Trends in LLM's

As noted previously, the big advantage of LLMs is that they are able to deeply analyze and synthesize large amounts of medical literature, patient records, and their clinical research, which is perennially being updated. Because of the underlying intricacy, heterogeneity, and omnipresent volume of data in healthcare [8, 9], managing the information in a timely manner is still a big problem. Automating the analysis of medical texts, extracting pertinent information, and applying that knowledge in research and patient care greatly advances medical practice. Integration of LLMs into healthcare systems is capable of bringing tremendous improvements. The overview of LLM applications is provided in figure 2.

The recent state-of-the-art models, namely, the GTP-3.5 [10], GTP-4 [11], and Bard by Google, which have demonstrated a variety of successes in performing various tasks in the field of natural language processing (NLP), make a strong contribution to the fact that the overwhelming attention is paid to LLMs in AI and promulgate its implication in the healthcare sector. Their potential to analyze and produce text that resembles those of human beings is likely to revolutionize healthcare operations especially in areas where effective communication and processing of information are key.

The NLP history [12] has been marked with significant events, which continue to contribute to the discipline as we see it today. Initial to the AI, the contextual dependencies in text were captured as a building step with recurrent neural networks or RNNs. However, RNNs were associated with the issues of long-range dependencies that were limiting. The invention of transformer architectures was a game changer because they overcame such problems and enabled more advanced models to be produced. It was owing to this structure that Llama2 and GPT-4 could be developed [12, 13]. After being trained on several datasets, these models transcended the boundaries imposed to NLP and endeavored to reach human like conception and text generation.

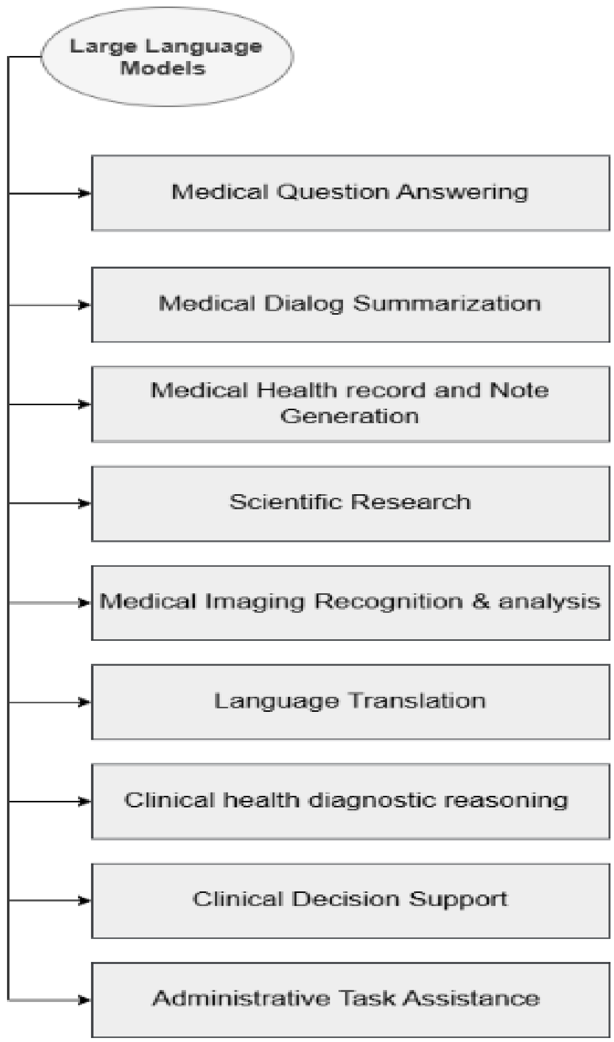

Figure 2. Overview of the Application of LLM's.

In BioMedicine, BSCs of transformer based models, such as BioBERT and ClinicalBERT [14], have been created to address the issues of medical language. These models are specifically structured to facilitate domain specific problems like heavy use of medical terminology, linguistic ambiguity and diction variation. However, the use of LLM in the highly sensitive, regulated field of medicine contains highly specific threats to ethics, privacy, and security. The basic components that need to be addressed are the protection of the data of the patient and the removal of bias and harm that the LLAMs might cause.

Despite these challenges, the current research and development work in the field continues to concentrate on the abilities of the LLMs with regard to enhancing the healthcare services, patient outcomes, and the development of medicine.

This is intended to fill the gaps in a domain like every other field does. This review [16] will serve medical researchers and healthcare experts who are interested in streamlining their operational research endeavors and clinical workflows by applying the use of LLMs. Our purpose is to assist in the determination of best LLM, depending on specific clinical needs. We deconstruct the technology of LLMs, usages, and potential applications in the healthcare sector in general, as well as crucial issues like fairness, prejudice, privacy, transparency, and ethical principles. Keeping these focus areas, we will address how the use of LLMs can be applied to the healthcare field and make the approach is ethical, fair, and transparent in the way it will impact the patients and the healthcare services providers.

This paper is structured in the following manner: Section II will provide an overview of other studies conducted in this field, Section III will cover the methodology that should be followed to find and shortlist the relevant research articles, Section IV will present the findings and discussions, and Section V will give the conclusion of this research that highlights the most important findings.

2. Related Work

Initial applications of the LLM were in the field of natural language processing (NLP) and text generation in general. Recurrent neural networks (RNN) models and the long-term memory (LSTM) models have played a central role in modeling dependence of texts in terms of long-range contextual dependence deficits, but they were not able to do so. architecture of transformer introduced [13]. Another important innovation was made in 2017 where self-awareness mechanisms were employed to allow models to perceive and narrative more segmented text as more efficient. These models form the basis of the creation of the pre-trained Transformer model, which formed the backbone of GPT-based models, BERT and state-of-the-art LLMs used in healthcare today.

The use of LLM in the health care system has been encouraging particularly in the clinical decision support and patient care. The formation of special models that include ClinicalBERT or BioBERT has facilitated the process of management of medical literature and clinical records by the models more effectively, giving information on disease prediction, medical text mining and management of patient data [14] [15]. Medical datasets are also fine-tuned on these models to enhance their capabilities to perform domain-specific tasks, such as answering medical questions and clinical documentation assistance.

Nevertheless, irrespective of these developments, the use of LLCM in the healthcare sector is replete with ethical dilemmas that go unnoticed. The literature on most existing surveys is related to technical development of LLCMs, model archetypes or performance indicators. Nonetheless, it is increasing as an understanding that the ethical considerations of these models like prejudice, privacy, accountability and transparency must be scrutinized with care before they can be safely implemented in the health care delivery environment [15] [16].

An example of a severe ethical issue can be bias in the LLM. A number of studies have cautioned on the danger of discriminating outcomes with regard to prejudicial training data. In the context of healthcare where models are applied to assist in clinical decision-making, any bias in the model may result in improper treatment advice and particularly in and underrepresented patient groups [17]. The research also emphasized the ways in which biases within medical AI models may strengthen health care disparities, particularly when models are trained using non-representative datasets. On the same note, research studies have established that models that are trained using historical health data that in many cases depict gender and racial inequalities may widen them unless appropriate measures are put in place [18] [19].

Another severe aspect of the implementation of an LLM in healthcare is privacy. Since these models deal with a lot of confidential patient information, it is of extreme importance that the confidentiality of data is upheld. A number of studies have pointed out the possible risks of patient privacy particularly in clinical decision support systems that deal with sensitive health information [20]. Federated learning is suggested as the means of enhancing data privacy because the model may be trained with the data that is not required to leave the local environment. Nonetheless, encryption of data and anonymization have remained a significant issue in the deployment of LLMs particularly in the wake of regulations around GDPR and HIPAA [21] [22].

Moreover, responsibility and transparency of LLM on the healthcare system have been criticized in the new literature. In case of the involvement of LLMs in making decisions, it is of importance that the healthcare professional comprehends how these models come up with the recommendations. Transparency (model transparency) can have considerable impacts on trust in such systems. Indicatively, attachments and attention mechanisms can be applied to Shapley to make the LLM more readable, which can make practitioners aware of the key aspects of the information that the model has concentrated on in its decision-making process [23]. Regardless of these developments, not all models are available as black boxes, and there are concerns regarding how they can be interpreted and the dangers of implementing AI in high-risk settings, including the healthcare sector.

Along with bias, one can refer to the problem of privacy and accountability, as well as regulatory gaps in AI healthcare. The regulation of AI in healthcare is an ever-changing subject, but, at the same time, there remains no unified framework covering ethical and legal issues regarding LLM in clinical practice. According to the research conducted, it is important to have uniform regulations to warp AI technologies in the healthcare sector to avoid abuse and safe patient care [24]. The authors state that the absence of clear ethical and legal guidelines reduces the chances of the adequate control of the use of LLCs, and as a result, it may provoke some ethical breaches or even negative consequences.

The proposed review intends to fill these literature gaps by synthesizing ethical evaluations of LLM in healthcare in a systematic manner. This review gives a comprehensive picture of how to solve ethical concerns about LLMs by comparing bias, confidentiality, transparency and accountability in studies in detail. Such a framework will assist healthcare professionals, policy makers, and AI developers to see the ethical path forward in implementing AI in healthcare as such potent tools can be utilized in a responsible manner to enhance patient care without breaching the ethical norms [25][26].

3. Methodology

The paper is an exploratory research that discusses the ethical aspects of Large Language Models (LLMs) in medicine as a field of study; it will utilize a mixture of the Systematic Literature Review (SLR) and qualitative interviews to consider the fundamental ethical issues. The purpose of the research is to understand the issues related to the ethical use of LLMs in the healthcare sector, which are the themes of bias, privacy, accountability, and transparency. The study technique was developed to help triangulate data of more than one source to come up with a holistic picture of the ethical issues. Methodology of this SLR has been presented in figure 3.

The initial step in the study was to perform an SLR as a means of gaining a theoretical basis of the ethical concerns surrounding the topic of LLMs in healthcare. The SLR summarized the findings of published works that were published and published within the year 2016-2024 on the subject matter in terms of ethical considerations, including bias, privacy, accountability, informed consent, and the possibility of misusing LLM in the healthcare environment. This step of the research provided the foundation of the qualitative interview, as the gaps in the literature were pointed out, and the crucial ethical principles were mentioned.

The SLR incorporated findings of studies in academic databases that are considered reputable, including PubMed, IEEE Xplore, ScienceDirect, and Google Scholar so that the scope of the topic could be covered in the results. The search strategy involved a search of healthcare related LLM terms, and high inclusion and exclusion criteria were followed. Peer-reviewed articles published between 2016 and 2024 were eligible and the articles had to be concerned with ethical issues occurring directly related to the practice of LLMs. A total of 50 studies were then selected to be reviewed in detail after the application of the criteria.

3.1. Inclusion Criteria (IC):

1. IC-1: Studies that specifically discuss ethical issues related to the use of Large Language Models (LLMs) in healthcare.

2.IC-2: Studies published between 2016 and 2024, ensuring the inclusion of the most current research.
3.IC-3: Peer-reviewed journal articles, conference papers, or technical reports that provide ethical analysis and insights into the use of LLMs in healthcare.

3.2. Exclusion Criteria (EC):

1.EC-1: Studies that focus solely on the technical aspects of LLMs, such as model performance or architecture, without discussing ethical issues.
2.EC-2: Studies published before 2016, as they do not meet the scope of this review.
3.EC-3: Studies that do not specifically address ethical issues or healthcare applications (e.g., studies on general NLP models or unrelated medical fields).

The selected studies were evaluated for quality using a quality scoring framework (discussed below), and only studies that met a minimum quality threshold were included in the final analysis.

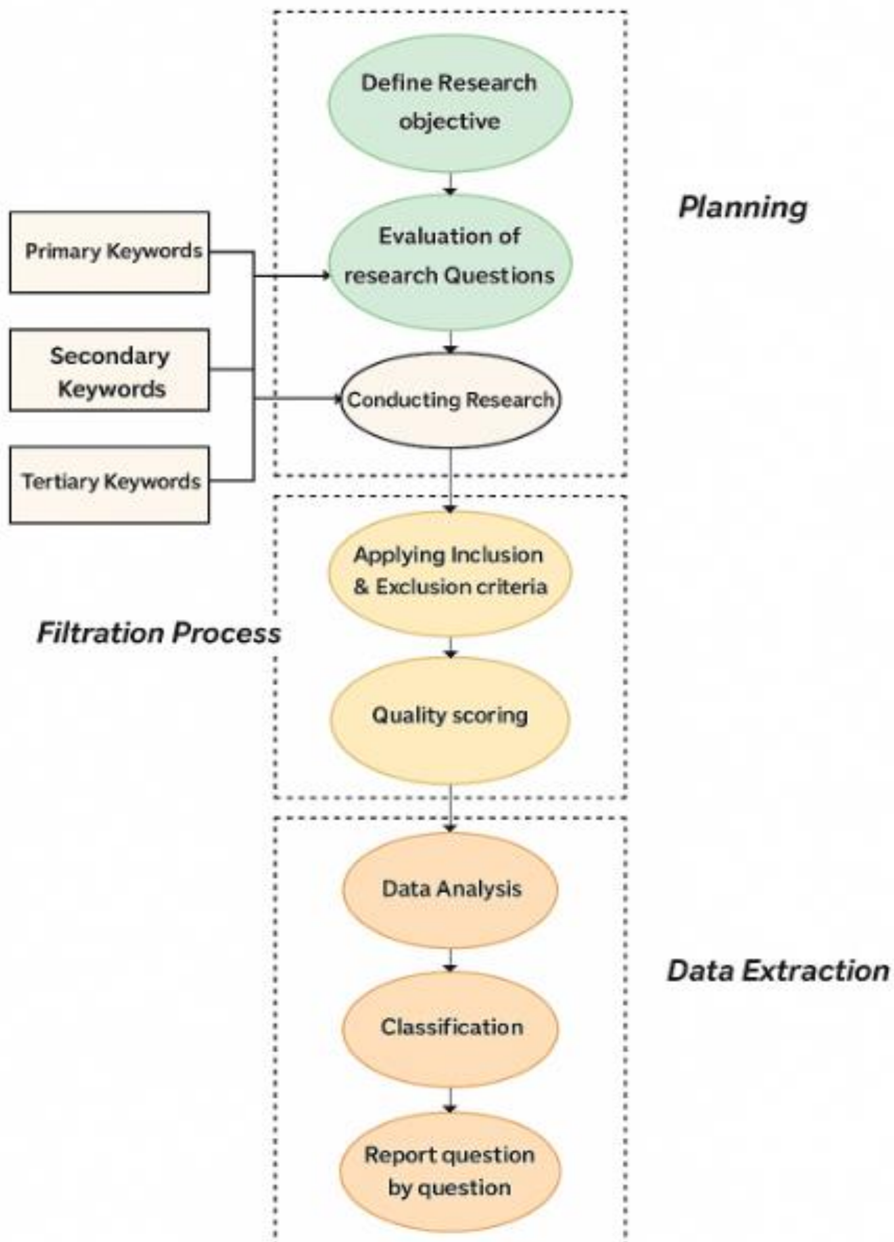

Figure 3. Research Methodology

3.3. Qualitative Interviews

Along with SLR, 40 healthcare professionals who have experience using LLM technologies in clinical practices were interviewed qualitatively. The healthcare professionals including physicians, and clinical researchers had gone through a training course on LLMs in healthcare in 2022. The interviews were to be conducted to gain first-hand experience and perception on ethical issues in the utilization of the LLMs in clinical decision-making.

The interviews were transmitted in 2023 based by semi-structured interviews. The interviews were face to face or video conferencing starting with 30 minutes up to 60 minutes. The participants were questioned in an open-ended question on their experience with LLCs and ethical concerns related to the policy of data privacy, transparency, accountability, and determining the role of AI in doctor-patient relationships. During all interviews, the interviews had been recorded confidentially and with the permission of the participants.

Thematic analysis was then conducted on the transcripts, and thematic patterns emerged on recurring ethical issues of using LLMs in healthcare and novel knowledge was gained. It was an inductive method that enabled the ethical concerns to rise directly out of the responses of the participants and the themes were combined with the results of the SLR to bring a more delicate view of the ethical concerns surrounding the implementation of LLM in clinical practice.

3.3.1. Selection Flow

The study selection process for the SLR was as follows:

1. Initial hits: Over 1,000 studies were retrieved across the selected databases.
2. Title and Abstract Screening: 400 studies were screened for relevance based on titles and abstracts.
3. Full-text Review: After applying the inclusion and exclusion criteria, 100 full-text articles were reviewed.
4. Final Selection: A total of 50 studies were selected for detailed review and analysis.

3.3.2. Quality Scoring

The quality scoring framework was applied to assess the reliability and rigor of the studies included in the SLR. The scoring criteria included:

- Relevance to the ethical issues addressed in the study (e.g., bias, privacy, accountability).
- Methodological rigor of the study, including clarity of ethical discussions and the appropriateness of research design.
- Impact and applicability of the findings in real-world healthcare settings.

Each study was assigned a score based on these criteria, and only those that met the minimum threshold were included in the final analysis

A. Research Objectives

The study proposed will aim to conduct an examination of the ethics of Large Language Models usage in medicine in the sense of its implications towards practitioners and patients and implications on medical decision-making. It would involve a systematic literature review (SLR) with the perspective of gaining insight into the ethical concerns, advancements, and overarching concerns regarding the application of the LLMs to the consideration of healthcare facilities. Patient privacy, bias, accountability, informed consent, and reliability of AI-related medical recommendations are issues that are likely to be handled by the study. The findings of the collected review will provide insights that will be significant in guiding the research in the future and the appropriate progress and practices of LLMs in healthcare.

RO1: To determine how the use of LLMs by medical practices relates to the ethical concerns of autonomy, privacy, and patient decision-making.

RO2: To examine circumstances under which transparency, explain ability and accountability tend towards the application of the LLCM in medical decision-making.

RO3: To determine possible regulatory and legal frameworks that will enable the process of solving ethical issues that arise because of the utilization of LLCs and offer suggestions concerning the prevention of the emergence of ethical risks and reasonable utilization of LLCs in the medical sphere.

B. Research Questions

This is in response to some of the most important questions on the ethics that will be used by LLMs in medical practice particularly how they will impact care providers, patients and the medical decision-making process. The research will use a Systematic Literature Review (SLR) to find meaningful ethical issues, developments, and concerns related to the use of LLMs in the medical field. It will investigate such burning questions as patient privacy, prejudice, responsibility, informed consent, and the credibility of AI-based medical recommendations. The study seeks to answer these questions in order to provide useful information, which will shape future research and the sustainable development and regulations of using LLM in medical practice.

RQ1: What are the ethical issues caused by the implementation of LLMs during patient care and medical decisions?

RQ2: How transparency and explain ability could be guaranteed in the medical decision support systems which are driven by LLM to keep the trust between healthcare providers and patients?

RQ3: Who is liable when LLMs commit medical diagnosis, treatment advice, or even interaction errors with patients, as well as how the punishment is to be handled?

C. Search String

The next important activity in a Systematic Literature Review (SLR) is the development of a systematic search plan that can be used to identify relevant studies in a particular field. In the current review, the sources were found in online repositories, such as Springer link, IEEE Xplore, Wiley, and Academics. The search strings were developed using the keywords as shown in Table 1, whereas Table 2 defines the search strings used in various repositories.

Table 1. Keywords used for searching

Primary Keywords	Secondary Keywords	Tertiary Keywords
Large language model	Natural Language Processing	Generative AI
Medicines	Machine Learning	Medical AI
Healthcare	Deep Learning	

D. Inclusion and Exclusion Criteria

In order to have a focused and comprehensive review, this paper will have some inclusion and exclusion criteria to choose relevant literature regarding the marketability of the ethical considerations of Large Language Models (LLMs) in medical practice. Peer-reviewed articles, conference papers, and reputable reports describing ethical issues and regulatory frameworks, patient privacy, bias, and accountability associated with the use of LLMs in healthcare will be given priority during the inclusion process. Included in the scope of this study are studies which have been released in English in the past ten years; as this category is believed to present new developments and new areas of concern. On the other hand, documents that are not directly related to medical practice, offer no ethical consideration, or only emphasize technical-related issues without touching on the ethical issues are not inclusive. Such requirements contribute to the relevance of the study and high quality synthesis of the existing knowledge.

Table 2. Search strings with respect to digital repositories from 2016 to 2024

Repository	Search Keywords	Search Strings	No of Papers
ACM Digital Library	"Large language model" OR "medicines" OR "healthcare" AND "Natural language processing" OR "Machine learning"	([All: large language model] OR [All: medicines] OR [All: healthcare]) AND ([All: natural language processing] OR [All: machine learning]) AND [All: Publication Date: (01/01/2016 TO	112,534

		10/31/2024)]	
Elsevier	"Large language model" OR "medicines" OR "healthcare" AND "Natural language processing" OR "Machine learning"	"Large language model" OR "medicines" OR "healthcare" AND "Natural language processing" OR "Machine learning"	10,110
Springer	"Large language model" OR "medicines" OR "healthcare" AND "Natural language processing" OR "Machine learning"	"Large language model" OR "medicines" OR "healthcare" AND "Natural language processing" OR "Machine learning"	4,730
IEEE Xplore	"Large language model" OR "medicines" OR "healthcare" AND "Natural language processing" OR "Machine learning"	(All Metadata: "large language model" OR "All Metadata": "medicines" OR "All Metadata": "healthcare") AND (All Metadata: "Natural language processing" OR "Machine learning")	29,083
Science Direct	"Large language model" OR "medicines" OR "healthcare" AND "Natural language processing" OR "Machine learning"	"Large language model" OR "medicines" OR "healthcare" AND "Natural language processing" OR "Machine learning"	1,000,000

E. Quality Score

Evaluating the quality of included studies is a crucial phase in SLR. The selected studies experienced a quality assessment, and their quality was evaluated using the specified criteria as presented in Table 3.

Table 3. Quality Scoring

Criteria	Description	Rank	Score
	Internal Scoring		
a)	Did the abstract clearly define the method of proposed solution?	Yes / Partially / No	1.5 / 1 / 0
b)	Did the study show comparison of the particular method with previously defined methods?	Yes / Partially / No	1.5 / 1 / 0
c)	Was methodology clearly defined?	Yes / Partially / No	1.5 / 1 / 0
d)	Was the conclusion based on	Yes / Partially / No	1.5 / 1 / 0

Criteria	Description	Rank	Score
	results?		
	External Scoring		
e)	What is the ranking of the publication source?	Q1 / Q2 / Q3 / Q4 / Core / A Core / B Core / C	2 / 1.5 / 1 / 1 / 1.5 / 1 / 0.5 / —

Classification of the literature studied is presented in figure 4 that categorizes studies based on investigation aspects and quality ratings, noting "None" where information is lacking.

In the classification Table 4, a systematic overview of different research papers in the area of machine learning and natural language processing (NLP) is given. It contains such important information as the reference number, the channel of publication (journal or conference), year of publication, dataset or study title, area of application, classification model applied, particular methodology, and architecture of the model. The areas covered include electronic health records (EHR) and predicting a disease, text mining in biomedical, and NLP assignments. BERT, GPT, and RoBERTa as well as BioBERT are several examples of these classification models that have a variety of applications. The table also draws examples of various methodologies including compact representations, deep learning, few-shot learning, and knowledge integration along with the various types of model architectures i.e. LSTM RNN and unsupervised feature learning. Arranging all these information, the table will assist in grasping the history of classifications used in various fields of research as well as give a comparative profile of the methods utilized and thus it is more straightforward to notice certain trends and progressions in the subject.

Table 4. Classification of studies

Ref	Channel	Year	Domain	Tool/ Technique
[1]	J	2016	EHR	Compact Representations
[2]	J	2017	Disease Prediction Relation Classification	Machine Learning
[3]	J	2017	Biomedical NER	Convolutional neural network LSTM RNN
[4]	J	2018	Language understanding	BERT
[7]	C	2018	Language understanding	Generative Pre-trained Transformer (GPT)
[9]	J	2019	NLP tasks	RoBERTa
[11]	J	2020	Multitask learning	Language models
[12]	J	2020	Facial movement analysis	Language models
[13]	J	2020	Few-shot learning	Language models
[14]	J	2021	Biomedical Text Mining	BioBERT
[15]	J	2021	Biomedical language models	Pretrained language models
[16]	J	2024		Machine Learning
[17]	J	2019		Machine learning models
[20]	J	2022	NLP, Deep Learning, Large Language Models (LLMs)	Natural language processing (NLP)
[21]	J	2022	Scalable NLP, Mixture-of-Experts, LLMs	Finetuned language models
[22]	C	2023	NLP, Transfer Learning, Language Models	Dual-view model

Ref	Channel	Year	Domain	Tool/ Technique
[23]	J	2021	NLP, Few-shot/Zero-shot Learning, LLM Generalization	3D-Shift Graph Convolution Network
[24]	C	2022	Computer Vision, Generative Models, Diffusion Models	Physics-informed neural networks (PINNs)
[25]	J	2024	Biomedical NLP, Health Informatics, Medical AI	Switch transformer
[26]	C	2017	Deep Learning, NLP, Transformer Architecture	Mixture-of-experts
[27]	J	2024	Clinical NLP, Healthcare AI, Electronic Health Records	BioGPT
[28]	J	2023	AI Ethics, NLP, Foundation Models, Model Evaluation	Clinical language model
[29]	J	2023	Multimodal AI, Vision-Language Models, Generative AI	BioBART
[30]	J	2023	Multimodal Learning, Vision-Language Pre-training, Foundation Models	Modality unifying network
[31]	C	2024	Multimodal AI, Computer Vision, Mixture-of-Experts, Vision-Language Models	Conditional diffusion model

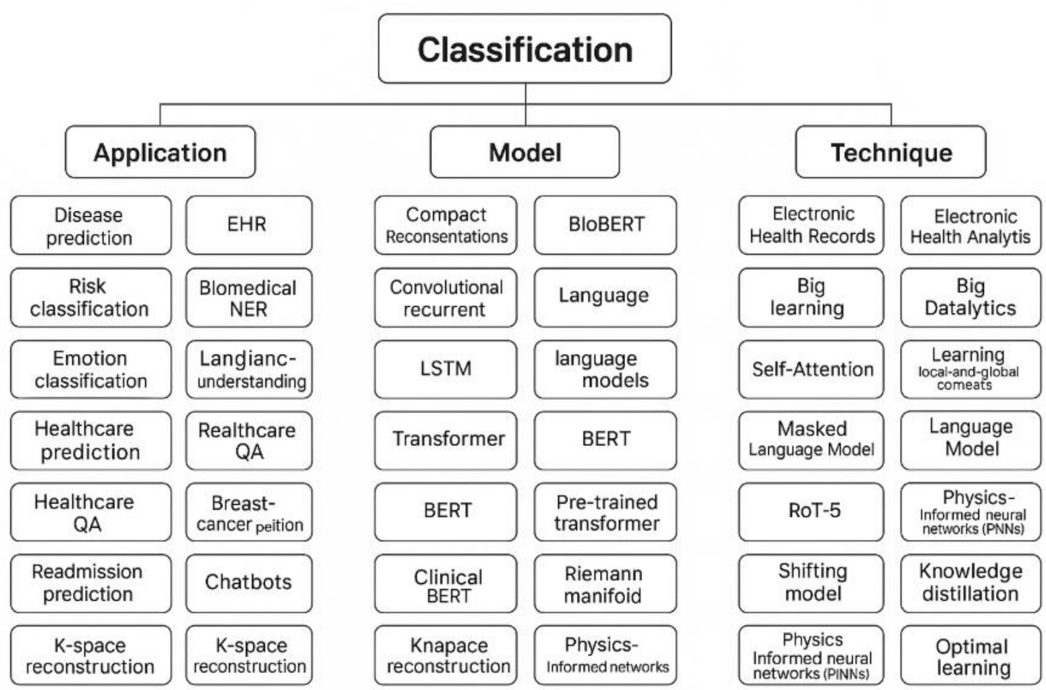

Figure 4. Classification Breakdown

4. Results

The selected articles undergo data extraction and synthesis in accordance with the ethical considerations of large language models (LLMs) in medical practice as delineated in this investigation. Figure 5 illustrates the distribution of LLM applications in healthcare over the given time period. The summarization of the years in which the chosen studies were published indicates a noteworthy upward trend in research on this subject, especially beginning in 2019. The years 2021 and 2022 exhibit the highest concentration of publications addressing ethical implications in medical AI.

Out of the 50 papers that make up the review, 45 (or 90%) are journal-published, while the remaining 05 (or 10%) are conference presentations. Notably, journal publications are more common in the following years: 2019, 2020, and 2021, reflecting the growing academic interest in the ethical deployment of LLMs in healthcare settings.

Criteria A is based on whether the paper is based on some technical algorithm, model, or is a review paper. No score is awarded for SLR, review paper, feasibility studies, or collaborative studies. Criteria b is based on whether or not a framework or architecture model of the system is presented in the document. Criteria c is based on limitations. No score has been awarded if the paper has not discussed any limitations of their study. Criteria d is based on the accuracy of the model presented. For accuracies above 80, 1 mark is awarded, for accuracies below 80, 0.5 is awarded, and 0 is awarded if the accuracy score is not discussed in the article.

The data extracted through research question 1 has been summarized in Figure 6. It shows the relationship between years of publications and publication venues from where the paper was extracted. Figure 6 shows the relationship between the research type and empirical type of the research. Research type is divided into solution and evaluation. The empirical type is further divided into experiment, comparative study and survey. The solution type is further divided into experiment.

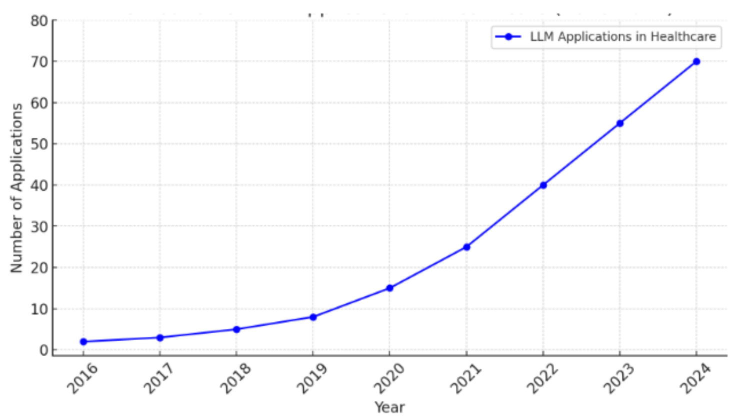

Figure 5. Distribution of LLM applications in healthcare over years.

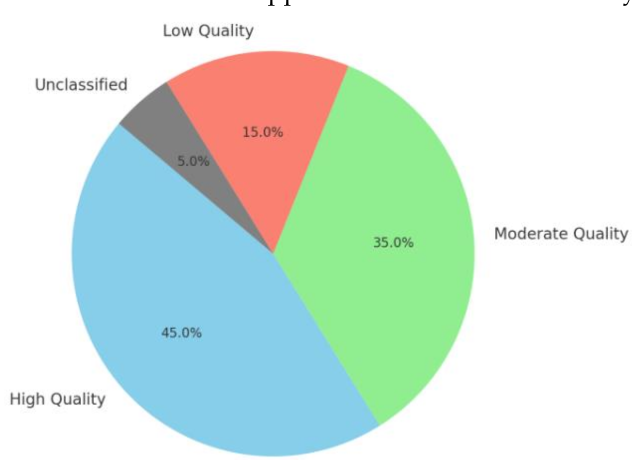

Figure 6. Quality scoring classification analysis.

Research Question-1: Which are the ethical issues in relation to the application of the LLMs to the care of patients and to medical decisions making?

There is also monumental ethical dilemma of ensuring the safety of patients, accuracy and accountability of the medical decision-making process with the use of the LLMs in the patient care and in medical decision-making process. A major problem is that it may result in the creation of misinformation and/or biased advice in which LLMs deliver answers using huge amounts of data that may either be erroneous or biased. A wrong or misleading medical advice might lead to misdiagnoses and/or wrong treatment or ill clinical decision making, not to mention the issue of not having some accountability. With AI, such consequences get unclear when it comes to medical ethics:

To address these risks, there is a need to ensure that healthcare professionals interact in the independent evaluation of the AI-enhanced recommendations alongside the transparency of the model decision-making.

The other side of the ethical issues mentioned above is associated with patient privacy and data security. To be effective, LLDMs require massive datasets, which at times contain sensitive medical information that without a proper protection might violate the privacy of patients. One has to abide with regulations which may include HIPAA and GDPR. It would be hard to guarantee patient information safety at all times in relation to the unseen mechanisms of AI-powered models. The other ethical implication of concern is that there exists a risk that LLMs may intensify the existing inequalities in healthcare due to training data biases, as there is an issue of unfairness and equity of healthcare. In the absence of stringent supervision and ethical principles, there is a threat that the subgroups will receive poor care because of the bias of algorithms. To curb these issues, it is important to have a multidisciplinary approach where ethical frameworks, regulatory policies and ongoing model evaluation are considered to ensure that there is responsible and fair utilization of LLMs in healthcare.

Research Question-2: What can be done to ensure transparency and explain ability in decision support systems that are driven by LLM and retain trust between healthcare providers and patients?

The stakeholders of healthcare facilities and patients rely on the quality of these medical decision support systems, and transparency and explainability are essential to ensure the system allied to the LLM is trusted. The development of interpretable AI can be considered one of the avenues; attention-based methods, the feature attribution approach, and model-agnostic explanation tools, such as SHAP and LIME, are included. By using these approaches, clinicians are able to see the path that the model undertakes to arrive at its recommendations and this results in the assurance that making of recommendations based on the medical knowledge and ethical standards. Additionally, fully documented LLMs, with the help of such information as the origins of the model training data, biases, and shortcomings would promote accountability and help healthcare professionals evaluate the credibility of these technologies before adopting them into a clinical workflow.

The human involvement when it comes to medical decisions that are supported by AI is another secret to retaining such trust. The medical staff will be required to confirm and compare the recommendations of the model with the medical practice and guidelines. This should be accompanied by an interactive interface in which a clinician can formulate questions that will prompt the system to provide more explanations and other suggestions to make the system more transparent and make clinicians responsible in exercising medical reasoning. Other ethical concerns, including the attitude to patient privacy and the reduction of bias, should be actively discussed by regular scrutiny of the model predictions and its further development based on evidence provided in practice. Ethical considerations ensured through the focus on explain ability and the reliability of AI systems as a helpful aid and not a decision-maker will enable the medical application of an LLM to maintain moral values and increase the level of trust in healthcare institutions.

Research Question-3: What is the most appropriate way to hold people liable when LLMs commit medical diagnosis, treatment prescription, or patient communication errors, and how is the term of legal accountability to be introduced?

Institutional accountability ought to occur. It would need a complex trade-off between morality and law in terms of accountability relating to medical diagnosis, treatment recommendations or patient interactions that are performed by large language models (LLMs). The responsibility spreads to involve the stakeholders of AI development like healthcare providers, creators of AI applications, and even regulatory authorities. When a physician and other health practitioners make clinical decisions, based on the AI-generated findings, they should perform clinical judgment by reviewing the recommendations against medical guidelines and other patient-specific variables. The issue of responsibility does not belong to AI makers and organizations alone. They need to make sure that their LLMs are well trained on quality, balanced and current information in medical disciplines. Development of clear guidelines by the regulatory bodies will also significantly contribute to the safety of the further implementation of the LLMs in the clinical practice i.e. in this respect the transparency, the explainability requirements and the constant monitor systems are also relevant to mitigate possible damages. In a very much-needed structure, the responsibility is to be shared in placement of legal protection and ethical checkpoint in order to deal with the liability. A policy concerning AI governance should be established so that the healthcare institutions utilize LLMs as decision-support systems and not as decision-makers. Further, legal mechanisms need to establish the division of liability, which includes what failures can be related to systems, biased training data or even hallmarks according to which the human oversight is lost. A highly potent combination of professional responsibility and institutional responsibility (may contribute to the harm of a patient), corporate responsibility (may lead to the flawed model design), and the enforcement of the regulations (may lead to the violation of the safety standards) will be needed. Also, the idea of patient consent and AI transparency must be highlighted, and patients should be aware of the extent to which LLMs are involved into their health and they can make informed decisions regarding the practice of medicine. The comparison of articles included in this research on the aspect of their dimensions is presented in Table 5.

Table 5. Comparison with existing solutions

Aspect	Current Techniques	Existing Solutions	Key Differences	Ref.
Bias and Fairness	Fine-tuning LLMs with diverse datasets, bias detection tools	Rule-based systems, decision trees	Current models focus on continuous learning from diverse data, existing systems are more rigid.	[11],[17], [23],[35], [47], [51]
Data Privacy and Security	Differential privacy techniques, encrypted data storage, federated learning	Traditional data encryption, secure servers	LLMs use federated learning for decentralized training, unlike traditional centralized systems.	[7-8], [13],[19], [22],[25], [38],[53]
Transparency and Explain ability	Explainable AI techniques (e.g., LIME, SHAP) to interpret LLM predictions	Rule-based AI, decision trees	Traditional solutions provide more transparency, while LLMs often operate as "black boxes."	[3],[15], [24],[26], [31],[39] [45]

Aspect	Current Techniques	Existing Solutions	Key Differences	Ref.
Accountability and Responsibility	Clear attribution of decisions to developers and AI systems, audit trails	Accountability through human oversight and regulatory bodies	LLMs may lack clear accountability pathways, requiring new regulatory frameworks.	[1-2],[4-5],[14],[16],[44]
Informed Consent	AI-assisted informed consent using chatbots for clarity	Manual consent process, paper-based documentation	LLMs enable dynamic, personalized consent conversations, unlike static documents.	[6],[9],[10],[12],[18],[33],[42]
Clinical Decision Support	Integration of LLMs in clinical decision support systems (CDSS)	Expert systems, clinical guidelines	LLMs offer more flexible, adaptive support compared to static expert systems.	[20-21],[27],[29],[34],[49]
Human-AI Collaboration	AI as a tool for supporting healthcare professionals (e.g., augmented intelligence)	AI as a decision-making partner, limited collaboration	LLMs focus on enhancing decision-making, while existing systems are more directive.	[30],[32],[36],[40],[43]
Regulatory Compliance	Ongoing development of AI-specific regulations and standards for healthcare	Compliance with HIPAA, GDPR, and other medical data regulations	LLMs require new, specific regulatory frameworks for ethical AI deployment in medicine.	[37],[41],[46]
Trust and Acceptance	Transparency efforts, patient education, controlled deployments	Limited patient interaction with AI, manual error checking	LLMs aim to foster trust through transparency and active patient engagement.	[48],[50],[52]

5. Discussion

Large Language Models (LLM) can be used in the healthcare sector to significantly optimize clinical decision-making processes and improve the quality of patient care, as well as simplify administrative workflows. These technologies, however, present important moral issues that should be dealt with care. The essential issues that can be gained in the Systematic Literature Review (SLR) and the qualitative interviews are bias, privacy, accountability, and transparency. Such problems should be addressed to make the use of LLMs ethical and responsible in the healthcare sector. Taxonomy of LLM has been presented in Figure 7.

Prejudice in the case of LLMs is a crucial ethical issue. The results of many studies [52-58] reviewed point out that the use of LLMs that have been trained on non-representative data sets may produce discriminatory healthcare outcomes. In the case of using the LLMs in clinical decision support systems, the bias of the model can lead to unequal treatment suggestions made to underrepresented groups of patients. Indicatively, AI models created in biased datasets might amplify healthcare inequities, especially to racial minorities and women[59-64].

Another major challenge is the privacy issue. As the LLAMs operate based on large volumes of sensitive patient information, it is essential to make sure that the data is secured in terms of privacy and confidentiality. Numerous research works, underline the role of federated learning and data encryption techniques to ensure that the privacy of patients remains intact when training a model. Nevertheless, confidentiality breaches and data breach are also some issues that may occur and any strong security measures can be used.

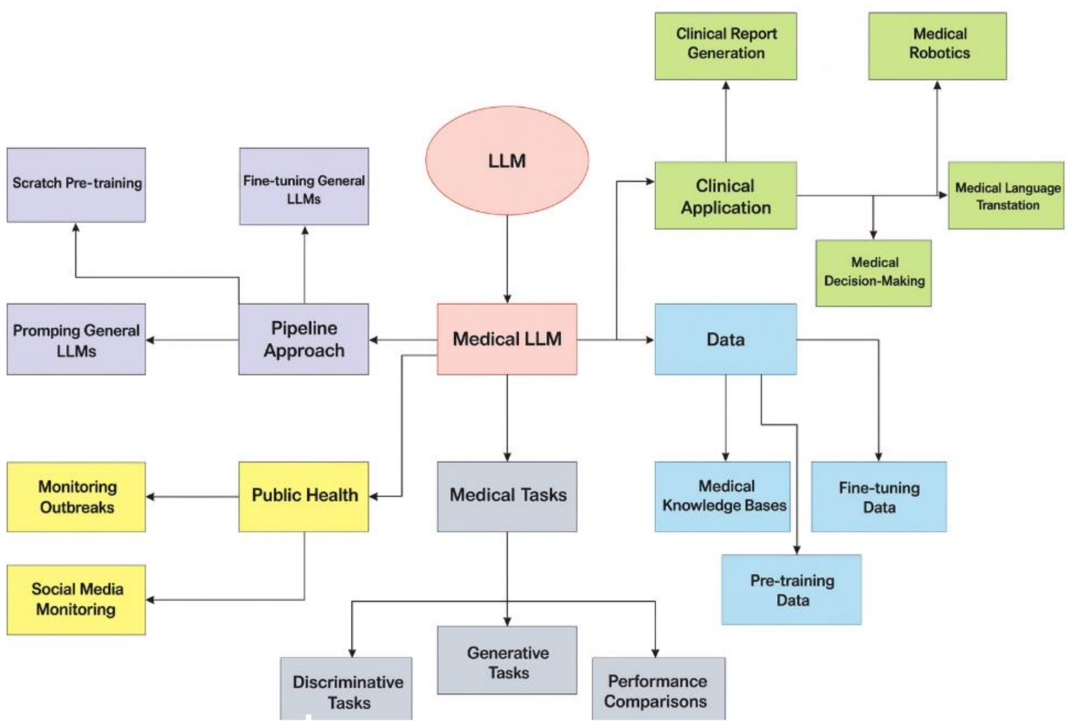

Figure 7. Taxonomy

There is also a need to hold oneself accountable in the implementation of LLM. Although the medical practitioners will be ultimately answerable to the patient care, the organizers of the LLM technologies should too be liable every time the AI system fails to provide the expected results calculations. Other researchers emphasize that regulatory frameworks should be in place that should establish clear accountability between healthcare professionals and AI developers whereby both the former and the latter assume responsibility in the decision-making process.

Lastly, transparency can be a key element to the trust in LLMs. Many AI models are opaque in nature, which would compromise the capacity of clinicians in comprehending the manner in which a model makes its decisions. The researchers suppose that make such LLMs easier to interpret, features of explainability, e.g., Shapley values and attention mechanisms may be considered.

5.1. Limitations

Although Large Language Models (LLMs) have created many opportunities and can have a significant positive impact on healthcare, multiple limitations to be resolved should be highlighted to consider wider use of the concept. The issue of clinical validation of these models is one of the main concerns. The majority of the LLMs have not been rigorously tested in practical medical settings. Consequently, the outputs by the LLMs tend to reflect statistical estimates as opposed to clinical determinations, which may cause misdiagnoses, incorrect treatment prescriptions and dissatisfactory contact with patients. Such absence of practical testing presents enormous concerns regarding the reliability, safety and regulatory compliance of LLMs especially in risky utilization in the medical domain.

The other important question is prejudice and equality of several AI-driven medical systems. Limited datasets are used to train LLPs, which risk having biases affecting the decision-making of the models. To illustrate this, some demographic categories, rare conditions or specialized medical literature might be poorly represented in the training data, resulting in a disparity in the treatment recommendations favoring some populations over others.

Elucidation and visibility are also the main issues. LLMs tend to be black-box models and, in such a case, clinicians and researchers find the way the models generate their recommendations hard to comprehend. This interpretability problem is a weakness and hinders the utilization of LLCMs in the clinical environment, where clarifiability is essential to make an informed decision. Lack of the capacity to justify predictions despite being right is a force that poses a hindrance to the integration of LLMs in the healthcare processes.

There are also data privacy and security issues which are a major challenge. The training of LLMs demands massive medical data in large quantities, creating ethical and legal concerns and considerations in terms of patient privacy, regulatory compliance, like HIPAA and GDPR. Possible confidentiality breaches, attacks of model inversion, or even unauthorized access can jeopardize the environmentally friendly use of LLMs in the area of medicine.

LLMs are associated with training and deployment costs that are heavy. Small outpatient clinics and low-resource medical settings do not easily have access to and implement such models due to the resources that are required which include, but are not limited to, energy consumption and data storage. Moreover, a constant revision of the model to maintain its correspondence with the changes in medical knowledge and clinical practice may be a complicated and multi-resourced process.

Lastly, the regulatory and ethical ambiguity of LLMs is another major obstacle to the integration of the technology in the clinical practice. Although LLM has immense potential in the clinical decision support field, in autopilot documentation and medical research, their applicability in autonomous decisions is extremely debated. Insufficient standardization of guidelines, absence of defined accountability models and strict regulatory regulations makes the assimilation of LLMs into the clinical processes a challenging task and questions the ethical use of AI it.

6. Conclusion & Future Directions

Huge advances made in the ability of language LLMs have generated radical developments in technology within the medical sphere, and multiple prospects exist to improve patient care and clinical decision-making, as well as simplify healthcare management in multiple spheres. However, the ethical and regulatory challenges that are brought by these developments are innumerable. These issues would be highly model-specific (LLM-based models) as compared to the issues raised against traditional machine learning or deep learning models. As powerful data processing systems that are able to generate human-like text and a variety of medical uses, LLMs have raised new concerns of bias, fairness, data privacy, and responsibility. Due to the inherently more fluid model dynamics of the issue of LLM in comparison to choice algorithms and the adaptive behaviour that may emerge during their implementation, unlike when traditionally applied to advanced models, the algorithms implemented by LLM are harder to predict and be in control of. These cases create challenges in the development of models and their regulation and bring up as ethical issues the ways to employ these models to medical practice. Regulators and developers are urgently required to establish a strong collaboration in order to deal with the intricacy of these issues. Mechanisms that guarantee the safety of patient information, equity in AI outputs and responsibility must be put in place. Simultaneously, this control mechanism should not deter the creative use of LLCs that could transform medicine in the world one way or another in the future in a manner that our wildest dreams could not imagine. There should be a meticulous balance in favouring innovation and enforcing moral principles that promote patient safety and protection of information and equity. To continue with the process of integrating LLMs into the practice of medicine, one will need to anticipate such challenges and collaborate in their resolution. Regulation avenues should have strong principles, carry out supervision, and ensure the open operation of the AI within which future synthesis of the LLMs will be answerable and advantageous to patients and healthcare professionals. The only way to effectively deploy LLC in a responsible manner within health care will be a compromise that will enable the opportunities and potential of LLC to be taken into account without forgetting about ethical concerns.

The applications of Large Language Models (LLMs) in the healthcare sector are still developing and thus offer an opportunity to make new investigations in most areas. The development of medical logic and causality knowledge in the future should focus on enabling the LLM to incorporate causal inference and diagnosing capabilities with implications of clinical use of decision support beyond the detection of simple patterns. The second critical pathway would be in multimodal LLMs that could integrate textual data with images, electronic health records (EHRs), and genomic data and wearable devices output to conduct the analysis comprehensively. An additional consideration to do is to address the problem of algorithmic biases to make sure that all parties become fairly represented as an LLC can disproportionately affect underrepresented groups. Avoiding generalization by encouraging the use of LLMs in certain medical specialties, including oncology, cardiology, and neurology, will increase their accuracy and reliability and, therefore, increase physician trust. Moreover, the development of few-shot and zero-shot learning will decrease the requirement of huge amounts of annotated datasets and allow LLMs to be more generalized on new medical conditions.

Acknowledgments:

The authors gratefully acknowledge the continuous support of senior research fellows for their expertise and assistance throughout all aspects of our study and for their help in writing the manuscript.

Contribution:

- **Ramsha Saeed (RS)** led the research design, including the **Systematic Literature Review (SLR)** and analysis of the **ethical issues** related to **Large Language Models (LLMs)** in healthcare. She was responsible for drafting the manuscript and synthesizing key findings on the ethical considerations of LLMs, with a focus on **bias**, **privacy**, and **accountability**.
- **Rabia Tehseen (RT)** contributed to the development of the **qualitative interview** framework, overseeing the **interview process** with healthcare providers. She analyzed the **interview data** using **thematic analysis** and provided insights into the practical **ethical challenges** experienced by clinicians using LLM technologies in healthcare.
- **Anam Mustaqeem (AM)** assisted in the **SLR** process, ensuring the relevant studies were selected based on the ethical issues identified. She also contributed to the interpretation of ethical challenges identified in the **literature** and **interviews**.
- **Usman Aamer (UA)** provided support in proofreading, revising, and refining the manuscript for consistency and clarity, ensuring that all sections cohesively addressed the **ethical issues** surrounding **LLMs in healthcare**.
- **Jawad Hassan (JH)** helped in writing and proofreading of the manuscript.

- **Patient consent**: Patient informed consent was not required due to the study design (survey of systematic reviews/literature).

Declaration of conflicting interests: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethical approval: Ethical approval was not required as this research was a survey of the literature.

Funding: The author(s) received no financial support for the research, authorship, and/or publication of this article

Guarantor: RT

References

1. Multiscale 3D-shift graph convolution network for emotion recognition from human actions. IEEE Intelligent Systems, 2022. 37(4): p. 103-110.
2. Modality unifying network for visible-infrared person re-identification in Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023.
3. Micro-expression action unit detection with dual-view attentive similarity-preserving knowledge distillation in 2021 16th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2021). 2021. IEEE.
4. Characterizing subtle facial movements via Riemannian manifold. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), 2019. 15(3s): p. 1-24.
5. A survey of large language models for healthcare: from data, technology, and applications to accountability and ethics. arXiv preprint arXiv:2310.05694, 2023.
6. Are large language models ready for healthcare? A comparative study on clinical language understanding in Machine Learning for Healthcare Conference. 2023. PMLR.
7. Leveraging generative AI and large Language models: a Comprehensive Roadmap for Healthcare Integration in Healthcare. 2023. MDPI.
8. M. Hussain, W. Sharif, M. R. Faheem, Y. Alsarhan, and H. A. Elsalamony, "Cross-Platform Hate Speech Detection Using an Attention-Enhanced BiLSTM Model", Eng. Technol. Appl. Sci. Res., vol. 15, no. 6, pp. 29779–29786, Dec. 2025.
9. Learning optimal k-space acquisition and reconstruction using physics-informed neural networks in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022.
10. Generating realistic brain mris via a conditional diffusion probabilistic model in International Conference on Medical Image Computing and Computer-Assisted Intervention. 2023. Springer.
11. Language models are few-shot learners. arXiv preprint arXiv:2005.14165, 2020.
12. Gpt-4 technical report. arXiv preprint arXiv:2303.08774, 2023.
13. One small step for generative ai, one giant leap for agi: A complete survey on chatgpt in aigc era. arXiv preprint arXiv:2304.06488, 2023.
14. Llama: Open and efficient foundation language models. arXiv preprint arXiv:2302.13971, 2023.
15. BioBERT: a pre-trained biomedical language representation model for biomedical text mining. Bioinformatics, 2020. 36(4): p. 1234-1240.
16. Z. Awais et al., "ISCC: Intelligent Semantic Caching and Control for NDN-Enabled Industrial IoT Networks," in IEEE Access, vol. 13, pp. 169881-169898, 2025, doi: 10.1109/ACCESS.2025.3614984.
17. Clinicalbert: Modeling clinical notes and predicting hospital readmission. arXiv preprint arXiv:1904.05342, 2019.
18. Transformers and large language models in healthcare: A review. Artificial Intelligence in Medicine, 2024: p. 102900.
19. Language models as knowledge bases? arXiv preprint arXiv:1909.01066, 2019.
20. Improving language understanding by generative pre-training. 2018.
21. Palm: Scaling language modeling with pathways. Journal of Machine Learning Research, 2023. 24(240): p. 1-113.
22. Switch transformers: Scaling to trillion parameter models with simple and efficient sparsity. Journal of Machine Learning Research, 2022. 23(120): p. 1-39.
23. Zubair, M.; Hussain, M.; Albashrawi, M.A.; Bendechache, M.; Owais, M. A comprehensive review of techniques, algorithms, advancements, challenges, and clinical applications of multi-modal medical image fusion for improved diagnosis. Computer Methods and Programs in Biomedicine. 2025, 272, 109014. https://doi.org/10.1016/j.cmpb.2025.109014.
24. Hussain, M., Chen, C., Hussain, M. et al. Optimised knowledge distillation for efficient social media emotion recognition using DistilBERT and ALBERT. Sci Rep 15, 30104 (2025). https://doi.org/10.1038/s41598-025-16001-9
25. Glam: Efficient scaling of language models with mixture-of-experts in International Conference on Machine Learning. 2022. PMLR.
26. Pre-trained language models and their applications. Engineering, 2023. 25: p. 51-65.
27. Finetuned language models are zero-shot learners. arXiv preprint arXiv:2109.01652, 2021.

28. High-resolution image synthesis with latent diffusion models in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022.
29. Utilizing natural language processing and large language models in the diagnosis and prediction of infectious diseases: A Systematic Review. American Journal of Infection Control, 2024.
30. Attention is all you need. Advances in Neural Information Processing Systems, 2017.
31. Critical care studies using large language models based on electronic healthcare records: A technical note. Journal of Intensive Medicine, 2024.
32. A survey of hallucination in large foundation models. arXiv preprint arXiv:2309.05922, 2023.
33. Visual chatgpt: Talking, drawing and editing with visual foundation models. arXiv preprint arXiv:2303.04671, 2023.
34. Blip-2: Bootstrapping language-image pre-training with frozen image encoders and large language models. in International conference on machine learning. 2023. PMLR.
35. Mova: Adapting mixture of vision experts to multimodal context. arXiv preprint arXiv:2404.13046, 2024.
36. Moe-llava: Mixture of experts for large vision-language models. arXiv preprint arXiv:2401.15947, 2024.
37. Cumo: Scaling multimodal llm with co-upcycled mixture-of-experts. arXiv preprint arXiv:2405.05949, 2024.
38. Large language models in medicine. Nature medicine, 2023. 29(8): p. 1930-1940.
39. Toward controlled generation of text in International conference on machine learning. 2017. PMLR.
40. Humangan: A generative model of human images in 2021 International Conference on 3D Vision (3DV). 2021. IEEE.
41. FloWaveNet: A generative flow for raw audio arXiv preprint arXiv:1811.02155, 2018.
42. Ai-generated content (aigc): A survey arXiv preprint arXiv:2304.06632, 2023.
43. Pre-trained models: Past, present and future. AI Open, 2021. 2: p. 225-250.
44. Distributed training of large language models in 2023 IEEE 29th International Conference on Parallel and Distributed Systems (ICPADS). 2023. IEEE.
45. Ruder, Universal language model fine-tuning for text classification arXiv preprint arXiv:1801.06146, 2018.
46. Multimodal large language models: A survey in 2023 IEEE International Conference on Big Data (BigData). 2023. IEEE.
47. Zubair, M., Owais, M., Hassan, T. et al. An interpretable framework for gastric cancer classification using multi-channel attention mechanisms and transfer learning approach on histopathology images. Sci Rep 15, 13087 (2025). https://doi.org/10.1038/s41598-025-97256-0
48. Jumping NLP curves: A review of natural language processing research. IEEE Computational intelligence magazine, 2014. 9(2): p. 48-57.
49. A comprehensive survey and analysis of generative models in machine learning. Computer Science Review, 2020. 38: p. 100285.
50. Class-based n-gram models of natural language. Computational linguistics, 1992. 18(4): p. 467-480.
51. Hidden markov models. Lecture notes, August, 2004. 15(18-19): p. 48.
52. A review of recurrent neural networks: LSTM cells and network architectures. Neural computation, 2019. 31(7): p. 1235-1270.
53. Markov chains: Structure and applications. Handbook of statistics, 2001. 19: p. 817-851.
54. Diversified hidden Markov models for sequential labeling. IEEE Transactions on Knowledge and Data Engineering, 2015. 27(11): p. 2947-2960.
55. An HMM posterior decoder for sequence feature prediction that includes homology in formation. Bioinformatics, 21 (1): i251-i257. Bioinformatics, 2005. 21(1): p. i251-i257.
56. The exploding gradient problem demystified-definition, prevalence, impact, origin, tradeoffs, and solutions arXiv preprint arXiv: 1712.05577, 2017.
57. Natural language statistical features of LSTM-generated texts. IEEE Transactions on Neural Networks and Learning Systems, 2019. 30(11): p. 3326-3337.
58. M. R. Faheem, T. Anees, M. Hussain, A. Ditta, H. Alquhayz and M. A. Khan, "Indexing in WoT to Locate Indoor Things," in *IEEE Access*, vol. 11, pp. 53497-53517, 2023, doi: 10.1109/ACCESS.2023.3272691.
59. Alhijawi, B., Kilani, Y., & Alsarhan, A. (2020). Improving recommendation quality and performance of genetic-based recommender system. International Journal of Advanced Intelligence Paradigms, 15(1), 77-88.

60. Almomani, O., Arabiat, A., Al Tayeb, M., Almaiah, M.A., Obeidat, M., Aldhyani, T.H., Shehab, R. and Rowad, M., 2025. A Robust Model for Android Malware Detection via ML and DL classifiers. Mesopotamian Journal of Big Data, 2025, pp.261-277. DOI:https://doi.org/10.58496/MJBD/2025/017.
61. Alqura'n, R., AlJamal, M., Al-Aiash, I., Alsarhan, A., Khassawneh, B., Aljaidi, M., ... (2024). Advancing XSS detection in IoT over 5G: A cutting-edge artificial neural network approach. IoT, 5(3), 478-508.
62. Masadeh, R., Almomani, O., Zaqebah, A., Masadeh, S., Alshqurat, K. et al. (2025). Narwhal Optimizer: A Nature-Inspired Optimization Algorithm for Solving Complex Optimization Problems. Computers, Materials & Continua, 85(2), 3709–3737. https://doi.org/10.32604/cmc.2025.066797.
63. Aljaidi, M., Alsarhan, A., Samara, G., AL-Khassawneh, Y. A., Al-Gumaei, Y. A., Aljawawdeh, H., & Alqammaz, A. (2022, November). A critical evaluation of a recent cybersecurity attack on itunes software updater. In 2022 International Engineering Conference on Electrical, Energy, and Artificial Intelligence (EICEEAI) (pp. 1-6). IEEE.
64. Shambour, Qusai, Mahran Al-Zyoud, and Omar Almomani. "Quantum-Inspired Hybrid Metaheuristic Feature Selection with SHAP for Optimized and Explainable Spam Detection." Symmetry 17, no. 10 (2025): 1716.https://doi.org/10.3390/sym17101716.