

Review began 10/23/2024
Review ended 11/16/2024
Published 11/19/2024

© Copyright 2024
Lee et al. This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

DOI: 10.7759/cureus.73994

Evaluating the Impact of Artificial Intelligence (AI) on Clinical Documentation Efficiency and Accuracy Across Clinical Settings: A Scoping Review

Craig Lee¹, Shawn Britto¹, Khaled Diwan¹

1. General Internal Medicine, University Hospitals Plymouth NHS Trust, Plymouth, GBR

Corresponding author: Shawn Britto, shawn.britto@outlook.com

Abstract

Artificial intelligence (AI) technologies (natural language processing (NLP), speech recognition (SR), and machine learning (ML)) can transform clinical documentation in healthcare. This scoping review evaluates the impact of AI on the accuracy and efficiency of clinical documentation across various clinical settings (hospital wards, emergency departments, and outpatient clinics).

We found 176 articles by applying a specific search string on Ovid. To ensure a more comprehensive search process, we also performed manual searches on PubMed and BMJ, examining any relevant references we encountered. In this way, we were able to add 46 more articles, resulting in 222 articles in total. After removing duplicates, 208 articles were screened. This led to the inclusion of 36 studies. We were mostly interested in articles discussing the impact of AI technologies, such as NLP, ML, and SR, and their accuracy and efficiency in clinical documentation. To ensure that our research reflected recent work, we focused our efforts on studies published in 2019 and beyond. This criterion was pilot-tested beforehand and necessary adjustments were made. After comparing screened articles independently, we ensured inter-rater reliability (Cohen's kappa=1.0), and data extraction was completed on these 36 articles. We conducted this study according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.

This scoping review shows improvements in clinical documentation using AI technologies, with an emphasis on accuracy and efficiency. There was a reduction in clinician workload, with the streamlining of the documentation processes. Subsequently, doctors also had more time for patient care. However, these articles also raised various challenges surrounding the use of AI in clinical settings. These challenges included the management of errors, legal liability, and integration of AI with electronic health records (EHRs). There were also some ethical concerns regarding the use of AI with patient data.

AI shows massive potential for improving the day-to-day work life of doctors across various clinical settings. However, more research is needed to address the many challenges associated with its use. Studies demonstrate improved accuracy and efficiency in clinical documentation with the use of AI. With better regulatory frameworks, implementation, and research, AI can significantly reduce the burden placed on doctors by documentation.

Categories: Healthcare Technology

Keywords: ai and machine learning, artificial intelligence in medicine, electronic health record (ehr), machine learning (ml), natural language programming (nlp), scoping review, speech recognition, ward round

Introduction And Background

Introduction

In our daily ward rounds, we often think about the heavy task of documenting patient information throughout the day, whether this may be progress notes, clerking, or discharge summaries. There were days when records were all handwritten; however, we have now moved towards a digital era where electronic health records (EHRs) are the norm [1-3]. The introduction of computers completely revolutionized healthcare, making it easy for staff to share patient information where needed and providing easy access to patient histories [4-6].

As we considered these advancements, we thought about the next revolutionary change. Could artificial intelligence (AI) enhance our clinical documentation processes similarly to how computers improved healthcare? The potential for AI to reduce workload and documentation burden seems promising [7-11]. This scoping review aims to explore the impact of natural language processing (NLP), machine learning (ML), and speech recognition (SR) on the accuracy and efficiency of clinical documentation across various clinical settings, including hospital wards, emergency departments, and outpatient clinics [12-20]. By having a look

How to cite this article

Lee C, Britto S, Diwan K (November 19, 2024) Evaluating the Impact of Artificial Intelligence (AI) on Clinical Documentation Efficiency and Accuracy Across Clinical Settings: A Scoping Review. Cureus 16(11): e73994. DOI 10.7759/cureus.73994

at the current literature, we seek to discover how AI can support healthcare staff and improve patient care [21-25].

Definitions of Key Terms

AI: AI refers to the simulation of human intelligence processes by machines, especially computer systems. These processes include learning (the acquisition of information and rules for using the information), reasoning (using rules to reach approximate or definite conclusions), and self-correction [26,27].

NLP: NLP is a branch of AI that focuses on the interaction between computers and humans through natural language. It involves the development of algorithms and models that enable computers to process and analyze large amounts of natural language data [12-15].

ML: ML is a subset of AI that involves the use of statistical techniques to give computer systems the ability to learn from data and improve from experience without being explicitly programmed. It includes various methods such as supervised learning, unsupervised learning, and reinforcement learning [10,16,17].

SR: SR is a technology that enables the recognition and translation of spoken language into text by computers. It involves the use of algorithms to process and analyze vocal input and is commonly used in applications such as voice assistants and transcription services [16,20].

Large language models (LLMs): LLMs are a type of AI model designed to understand and generate human-like text based on vast amounts of data. These models, such as GPT-3, are trained on diverse datasets and can perform various language-related tasks, including translation, summarization, and question-answering [23,25,28-30].

Rationale

Shifting from handwritten to digital records via EHRs has created some challenges in clinical documentation, by increasing the number of administrative tasks that healthcare professionals need to perform. This detracts from patient care. There is evidence that suggests that AI can help ease the burden of these administrative tasks in clinical documentation, with the use of ML, NLP, and SR. These technologies are being used with greater efficiency and accuracy as time goes by; however, the literature does highlight the challenges of using AI in this context.

This scoping review will explore the impact of AI on clinical documentation in terms of efficiency and accuracy and look at challenges that arise while using it across various healthcare settings. A scoping review approach is suitable given the exploratory nature of the questions and the diversity of study designs, allowing for the comprehensive mapping of existing literature, identification of research gaps, and examination of complex AI interventions in clinical workflows. This review can offer further insight for stakeholders and may help guide research and integration of AI into healthcare in the future.

Methods

Study Design

This scoping review aimed to investigate the impacts of AI technologies on the accuracy and efficiency of clinical documentation across different clinical settings. Our primary research question was "What are the impacts of AI technologies on the accuracy and efficiency of clinical documentation in various clinical settings?".

Search Strategy

A literature search was done on Ovid, and the search string used to identify pertinent articles is shown below in Table 1.

Step	Search string
1	("artificial intelligence" or "AI" or "natural language processing" or "machine learning").mp.
2	exp Artificial Intelligence/
3	1 or 2
4	((clinical adj document*) or (medical adj document*) or (health adj record*) or (patient adj record*)).mp.
5	exp medical records/ or exp electronic health records/
6	4 or 5
7	3 and 6
8	7 and ("language model" or "LM" or "SR" or "speech recognition").mp.
9	limit 8 to yr="2019 - Current"

TABLE 1: Search strategy for identifying literature on AI in clinical documentation

AI: artificial intelligence

This search strategy yielded 176 articles. We then combined the Ovid search with manual searches in PubMed and BMJ, as well as reference examination, identifying an additional 46 articles. This process resulted in a total of 222 articles. After removing 14 duplicates, we screened 208 articles. Following this, we excluded 154 articles and assessed 54 full texts for eligibility, ultimately narrowing our selection down to 36 studies for inclusion in our scoping review.

Inclusion and Exclusion Criteria

Specific inclusion and exclusion criteria were used to select our 36 articles. The inclusion criteria included empirical research articles (quantitative, qualitative, or mixed methods), case studies, evaluations, experience reports, observational studies, systematic reviews, scoping reviews, meta-analyses, and relevant conference papers that reported on the application and impact of AI technologies in clinical documentation. We focused on studies involving NLP, ML, SR, and other AI technologies used in various clinical settings, including inpatient units, emergency departments, and outpatient clinics. We considered studies published within the last five years to ensure we captured recent advancements in the field and included only those published in English or with an English translation available.

Exclusion criteria filtered out studies not centered on clinical documentation, applications of AI unrelated to documentation (e.g., diagnostic tools, medical imaging), studies conducted in non-clinical settings, and studies with outcomes not related to documentation practices. We also excluded studies published before 2019 or those published in languages other than English without available translations.

Pilot Testing

Prior to finalizing our methods, we conducted pilot testing with a small subset of articles to fine-tune our data extraction process. This helped us find any potential challenges in our review process and refine our research question and inclusion/exclusion criteria.

Data Extraction and Review Process

Multiple reviewers took part in data extraction. Each article was independently read to determine its relevance to our research question. A qualitative assessment was carried out which allowed us to capture subtle insights from the literature. Table 2 shows the data which was extracted from each article.

Data extracted	Details
Study characteristics	Authors, publication year, study design
Clinical setting	Inpatient, outpatient, emergency department
Type of AI technology	Natural language processing, speech recognition, machine learning
Key findings	Impact on documentation accuracy and efficiency
Benefits and challenges	Reported benefits and challenges
Recommendations	Future research and practice recommendations

TABLE 2: Data extraction and review summary of AI in clinical documentation
AI: artificial intelligence

We narrowed down the articles to 36 in total after applying our inclusion and exclusion criteria. The reviewers conducted an inter-rater reliability assessment, achieving a Cohen's kappa score of 1.0, indicating perfect agreement in our initial screening and selection process. We resolved minor discrepancies through consensus, to ensure the integrity of the data extraction process and a thorough understanding of the literature.

Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Guidelines

We followed the PRISMA checklist to guide our review process, ensuring a comprehensive and systematic examination of the current landscape of AI technologies in clinical documentation across various clinical settings. A PRISMA flowchart illustrating the selection process is included (Figure 1).

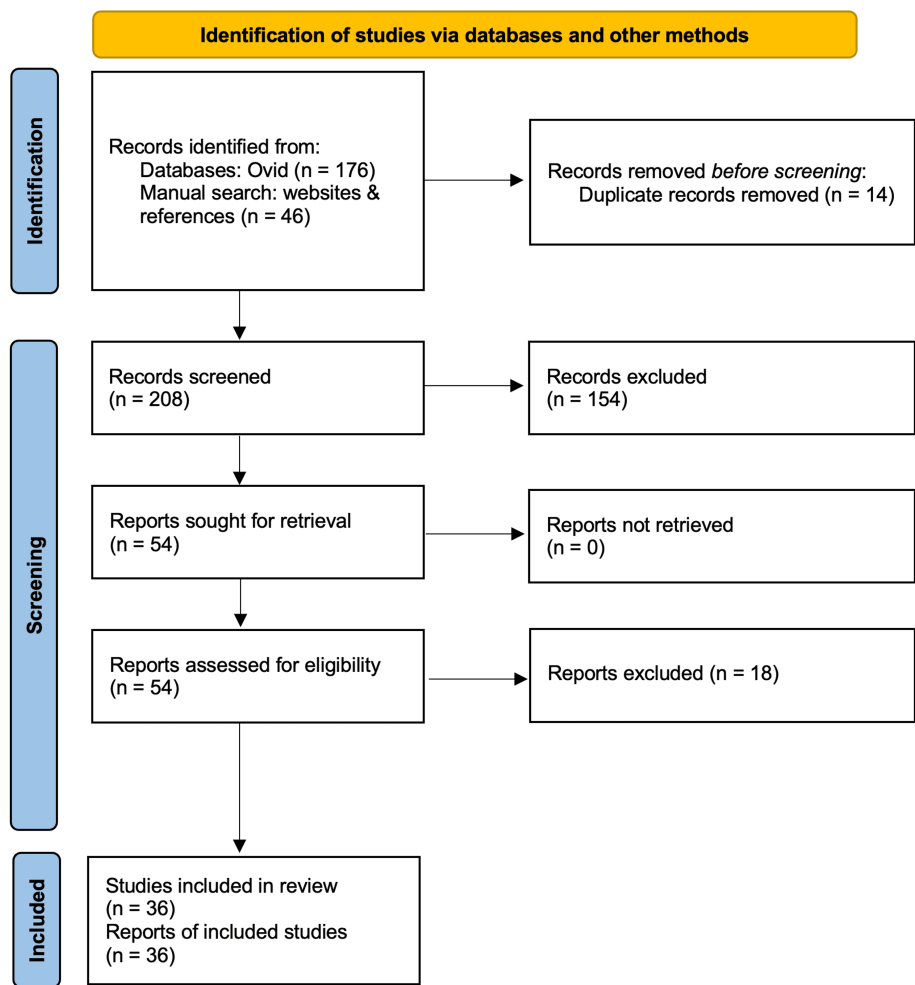


FIGURE 1: PRISMA flow diagram for the scoping review on evaluating the impact of AI on clinical documentation efficiency and accuracy across clinical settings

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses; AI: artificial intelligence

Reference: [31]

Review

Overview of studies on AI in clinical documentation

The studies reviewed span various healthcare settings and explore different aspects of using AI, specifically LLMs, NLP, ML, and SR, in clinical documentation. This broad analysis includes observational studies, scoping reviews, systematic studies, and experience reports. Table 3 shows findings from these articles which are relevant to this review. Together, these studies provide a thorough overview of the current state, benefits, challenges, and future directions of AI-driven clinical documentation.

Aspect	Finding	Citation
Efficiency and user experience	Speech recognition improves workflow efficiency but requires editing by clinicians for accuracy.	[7,20]
	LLMs like GPT-4 make discharge summaries more readable, even though they still require clinicians to edit them afterwards.	[3,5,17,19,23,25,29]
	LLMs facilitate better communication between physician and patient. It also improves medical training.	[3,5,17,25]
	Ambient AI scribes reduce clerical workload, by reducing the time clinicians spend documenting ward rounds. This improves clinician efficiency and patient interaction.	[32]
	Speech recognition reduces documentation time and improves workflow efficiency in various clinical settings.	[7,20]
Accuracy and error management	LLM-generated summaries sometimes contain misinterpretations, certainty illusions, fabricated information, and attribute errors. This is why clinicians may still be needed to review them.	[11,15,23]
	Speech recognition systems have been shown to improve documentation time but still need significant editing to ensure the reliability of its content.	[7,20,24]
	Various studies highlight that LLMs require careful review to ensure clinical accuracy and prevent misinformation.	[11,23,24,30]
Clinical utility and safety	Integrating LLMs into workflows can reduce clinician documentation burden, though this may come with its own risks such as fabricated information.	[11,25,29,32,33]
	LLMs can generate patient-friendly discharge summaries, enhancing patient-centered care.	[3,5]
	LLMs improve medical documentation but still require human oversight to ensure patient safety.	[4,15,32,34,35]
	AI-driven documentation tools like ambient scribes enhance clinical utility by reducing time spent on clerical tasks.	[32,36]
Patient-centered care	AI enhances patient health literacy and treatment adherence by providing information in an easier-to-understand language.	[3,5,8,32,35]
	Patient-friendly discharge summaries created using LLMs show improvements in readability and understandability.	[3,5,35]
	Ambient AI scribes allow clinicians to focus more on patient care by reducing the amount of clerical work. Patient-physician interactions are then improved.	[32,36]
Liability and ethical considerations	Legal uncertainties exist regarding liability for AI-generated medical errors, necessitating clear regulations.	[33,34]
	Ethical considerations such as data privacy and model interpretability are crucial for safe AI implementation.	[17,33,34]
	Balancing technical benefits with ethical considerations is essential, requiring more empirical research on LLMs' clinical utility.	[17,34]

TABLE 3: Summary of key findings in AI-enhanced clinical documentation

LLMs: large language models; AI: artificial intelligence

Key findings

Efficiency of AI Documentation

SR has been shown to be beneficial in reducing documentation time and improving workflow efficiency in various clinical settings. Many clinicians find this technology satisfactory and believe it improves efficiency [7,20].

Accuracy and Error Management

GPT-4 generated more accurate summaries compared to GPT-3.5 Turbo, with fewer inaccuracies and hallucinations, but errors like omissions were still a concern. Studies also found that LLM-generated

summaries sometimes contain misinterpretations, certainty illusions, fabricated information, and attribute errors [6,11,15,23,29]. While SR have been shown to improve documentation speed, they still require significant editing to correct errors [7,20].

Clinical Utility and Safety

Integrating LLMs into clinical workflows can help reduce the documentation burden on clinicians and allow more time for patient care, though fabricated information poses safety risks [4,11,25,29,32,33].

Patient-Centered Care

Patient-friendly discharge summaries created using LLMs have shown improvements in readability and understandability [3,5,8,34,35]. This can then improve patient health literacy and even treatment adherence. Ambient AI scribes have improved patient-physician interactions by reducing clerical work and allowing physicians to focus more on patient care [32,36].

Legal and Ethical Considerations

There will always be concerns surrounding AI as it is implemented into healthcare systems; therefore, regulations are needed to address ongoing legal uncertainties about liability for medical errors caused by AI systems. Ethical considerations such as data privacy are also crucial for the safe implementation of AI in healthcare [17,33,34].

Comparative studies and innovations

Studies comparing AI-generated summaries with those by senior internal medicine residents found similar performance levels, suggesting AI's potential utility in medical documentation [37]. LLM-generated notes from doctor-patient conversations are often preferred nearly as much as human-written notes, indicating their potential for automated clinical documentation [15].

Modular summarization techniques like CLUSTER2SENT and fine-tuning transformer models (e.g., BART) show promise in generating high-quality clinical summaries from conversation transcripts [24,30]. Pre-trained models like BERTSUM and BART have been effectively used for summarizing medical dialogues, with GPT-3.5 generating more comprehensible summaries according to human evaluators [29].

Implementation challenges

Technical improvements and customization are required for the effective integration of AI tools with existing EHR systems, as this remains a challenge [19,36].

To ensure the maximal benefit of these AI technologies are obtained, extensive training of healthcare professionals is required. Only then will there be adequate adoption of AI tools in clinical practice. Studies highlight the importance of providing adequate training to clinicians to maximize the benefits of AI technologies [9,21]. Early adopters have reported improvements in documentation efficiency and accuracy after proper training [27].

Discussion

Several studies show that AI in clinical documentation can reduce the administrative burden on healthcare professionals. Today, this technology is becoming increasingly integrated into clinical settings and is already positively affecting staff and patients in some of the areas it is being used. It does this while keeping the accuracy and integrity of records. The result is reduced clinician workload and better workflow efficiency through increased speed of documentation and less administrative work [7,8,18,20,36]. We will also discuss some of the challenges and future potential and hope for AI in clinical documentation. Figure 2 provides an outline of the strengths, challenges, and future directions of AI in clinical documentation.

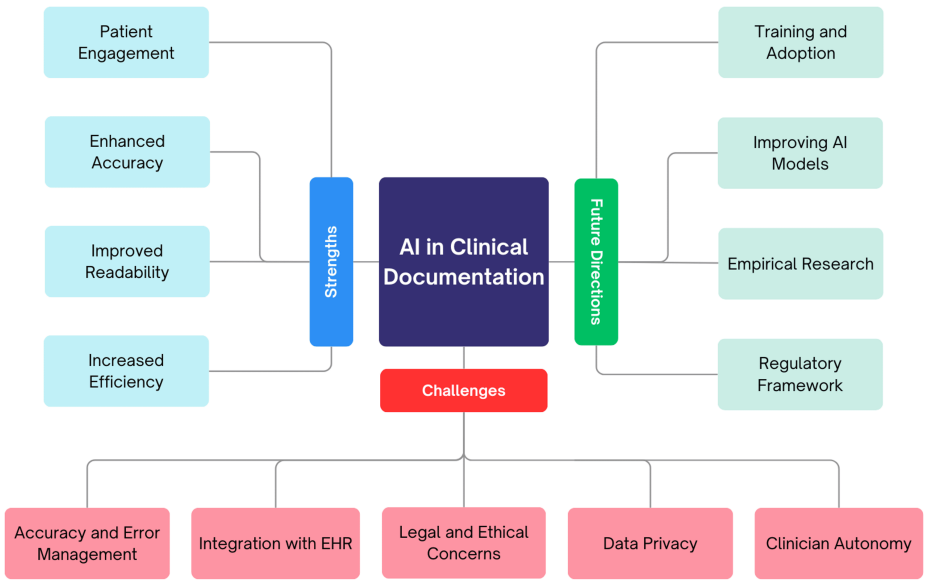


FIGURE 2: Outline of the strengths, challenges, and future directions of AI in clinical documentation

AI: artificial intelligence

Strengths of AI in Clinical Documentation

One of the first strengths we will explore is increased efficiency in documentation. AI can complete tasks and process data faster than humans, and we see this positively impacting workflow efficiency, through a reduction in documentation time [19,20,27]. Reducing the time clinicians spend on documentation allows us to focus more on patient care [7,18,20,36].

One benefit for the patients is having a better understanding of the documents given to them by health institutions. In studies focused on AI-generated clinical summaries, it is clear that AI can enhance the readability and understandability of these documents, improving patient engagement and adherence to treatment [3,5,8,35]. Good readability is crucial for patient care, where clear communication can lead to better health outcomes and patient satisfaction [3,5]. LLMs have shown the potential to simplify complex medical jargon, making it easier for patients to understand their health status and follow medical advice [5].

Moreover, AI can produce comprehensive and accurate clinical summaries when effectively managed. Studies comparing AI-generated content with human experts have shown that AI can match or even exceed human accuracy [6,11,15,19,37]. There can also be errors in manual documentation, and in comparison studies between GPT-4 and GPT-3.5, we see the potential of AI to reduce these types of errors; however, AI does produce errors of its own [6,11,23]. Even though AI systems have been found to improve the speed and accuracy of clinical documentation, they will still require human oversight to ensure those errors do not occur [19,24,27].

Challenges With AI in Clinical Documentation

One significant challenge is error management. Errors such as hallucinations (factually incorrect), omissions (leaving out important information), and fabricated information in AI-generated summaries pose significant risks and require rigorous validation and continuous oversight by clinicians [6,11,23]. For instance, studies have noted that while GPT-4 shows improvement in documentation, it still introduces errors that can affect clinical documentation quality and, consequently, clinical outcomes [6,14,16,26]. The main concern with these errors is their potential to compromise patient safety and treatment efficacy [33]. We therefore need to ensure that we have extensive error management strategies when integrating AI systems.

Effective integration of AI tools with existing EHR systems remains another challenge, requiring ongoing technical improvements and customization [19,36]. This is important as studies have highlighted the need for seamless integration to maximize the utility of AI in clinical documentation [36].

Bias and inequality are additional concerns. It is therefore important to address these biases to ensure fairness and equity in clinical documentation [17]. AI systems can perpetuate biases present in the data they are trained on, leading to unequal documentation outcomes. To prevent disadvantages to certain patient

groups, there must be strategies in place to mitigate these biases [11,17].

Furthermore, determining accountability for decisions made by AI systems in clinical documentation is not clear-cut. There is a need for clear guidelines on the responsibilities of developers, healthcare providers, and institutions in the context of AI usage for clinical documentation [34]. Studies have called for frameworks that delineate the accountability of different stakeholders involved in the deployment and use of AI in clinical documentation [32,34].

There must always be measures in place to prevent unauthorized access and breaches. Without these measures, we may lose patient trust and undermine the integrity of clinical documentation systems. Hence, it is important that we comply with legal standards and ensure that the AI systems in use prioritize data security [15,17].

There are also healthcare providers who are hesitant to accept and trust AI systems. Efforts to improve the explainability and transparency of AI decision-making processes are crucial for their widespread adoption in healthcare [17,35]. Clinicians will begin to trust this technology as long as it maintains transparency. Only then will AI be able to facilitate more effective clinical documentation processes.

Balancing the use of AI with clinician autonomy is essential. AI should augment, rather than replace, clinical judgment to ensure that healthcare providers remain in control of clinical documentation decisions [34]. Maintaining this balance can help integrate AI into clinical documentation practice without undermining the expertise and decision-making capabilities of healthcare professionals [36].

Future Directions and Insights

There are several key areas in which research and development can focus on to improve AI in clinical documentation. Future models should be more reliable. ML can make this possible by allowing AI models to better understand context in medical language over time [2,5].

For a deeper understanding of AI in clinical documentation, further studies will need to be done. More randomized controlled trials, observational studies, and longitudinal studies are needed to assess clinical outcomes. Rigorous empirical research will provide robust evidence to support the integration of AI in clinical practice [17,32].

It is vital to set up clear legal and ethical guidelines to manage the use of AI in healthcare, particularly concerning liability for errors and data privacy concerns [15,34]. These frameworks should also include guidelines for the ethical use of patient data in training AI models and ensuring transparency in AI decision-making processes [17].

Over time, the needs of healthcare providers and patients are always evolving. Therefore, for these AI systems to be adaptable, user feedback would be crucial where clinicians can report errors and improvements can be made in real-time [8,11].

While AI has some reported benefits in documentation, for it to be used correctly, healthcare staff will need adequate training. If this is not done, the implementation of AI tools will not be successful. Education and support for staff through comprehensive training programs will then be of utmost importance to effectively use the technology [9,14].

Building patient trust in AI is crucial. Ensuring transparency in how AI tools are used and involving patients in the review process can enhance trust and acceptance. Efforts to involve patients in the development and refinement of AI tools can also help ensure that these technologies meet their needs and expectations [36].

For health institutions to implement AI systems successfully, the cost-effectiveness of their use will need to be assessed. Reduced documentation time and increased efficiency can save money [32]. However, a thorough analysis of the long-term financial benefits of AI will need to be done.

Limitations

Several limitations were encountered while performing this scoping review. For instance, we acknowledge that most of the articles included are experience reports, cross-sectional studies, or observational studies. There is a lack of longitudinal studies, which limit understanding of AI's long-term impacts. With adequate funding and the integration of AI systems into healthcare, future research can focus on cohort studies that can determine the long-term effects of using AI in healthcare and assess patient outcomes. Also, due to the nature of our scoping review, we have not broadly covered all AI technology currently in use, as well as emerging AI tools such as predictive analytics and computer vision. We focused on NLP, SR, and ML and their use in various clinical settings, mainly in developed countries, and we also focused on studies published in English. This introduces a cultural and geographical bias; therefore, it is crucial to conduct

research in diverse contexts to understand the impact of AI in healthcare globally. Lastly, to review the latest trends and advancements, we only included articles from the past five years. Research into older studies can provide some historical context and insights into the evolution of AI in clinical documentation.

Conclusions

AI has significant potential to reshape clinical documentation through the enhancement of efficiency, accuracy, and patient engagement. This can be made possible by utilizing LLMs and various other AI technologies. We will see a reduction in documentation time, improved readability, and better patient-centered care. However, several key challenges must be addressed first.

Accuracy and efficiency are the main concerns. We've observed that AI-generated summaries and progress notes sometimes contain errors, known as hallucinations and omissions. Therefore, for this process to be validated, it still requires oversight by clinicians. A comprehensive regulatory framework is therefore needed to address these errors, as there are legal and ethical concerns. Additionally, effective integration of EHRs must make full use of the various AI tools available.

Later efforts should aim at improving AI models by addressing biases, creating better prompt engineering, and developing robust evaluation metrics. More research, such as randomized controlled trials and observational studies, are necessary to improve the validity of AI in healthcare. Clear legal and ethical guidelines will help manage the risks associated with AI use, ensuring data privacy and model interpretability. Continuous evaluation and adaptation of AI tools are essential to meet the evolving needs of healthcare providers and patients. Collaboration between clinicians, AI developers, and policymakers is important to take full advantage of in AI in healthcare. By focusing on these areas, the healthcare industry can effectively leverage AI to reduce the documentation burden on clinicians, enhance patient care, and improve overall healthcare delivery. The path forward requires a balanced approach that integrates technical innovation with ethical and regulatory considerations, ensuring that AI technologies are implemented safely and effectively to support the goals of modern healthcare.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Shawn Britto, Craig Lee, Khaled Diwan

Acquisition, analysis, or interpretation of data: Shawn Britto, Craig Lee, Khaled Diwan

Drafting of the manuscript: Shawn Britto, Craig Lee, Khaled Diwan

Critical review of the manuscript for important intellectual content: Shawn Britto, Craig Lee, Khaled Diwan

Disclosures

Conflicts of interest: In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

References

1. Baughman DJ, Botros PA, Waheed A: Technology in medicine: improving clinical documentation. *FP Essent.* 2024, 537:26-38.
2. Gaffney A, Woolhandler S, Cai C, Bor D, Himmelstein J, McCormick D, Himmelstein DU: Medical documentation burden among US office-based physicians in 2019: a national study. *JAMA Intern Med.* 2022, 182:564-6. [10.1001/jamainternmed.2022.0372](https://doi.org/10.1001/jamainternmed.2022.0372)
3. Kim H, Jin HM, Jung YB, You SC: Patient-friendly discharge summaries in Korea based on ChatGPT: software development and validation. *J Korean Med Sci.* 2024, 39:e148. [10.3346/jkms.2024.39.e148](https://doi.org/10.3346/jkms.2024.39.e148)
4. Ahuja AS: The impact of artificial intelligence in medicine on the future role of the physician. *PeerJ.* 2019, 7:e7702. [10.7717/peerj.7702](https://doi.org/10.7717/peerj.7702)
5. Zaretsky J, Kim JM, Baskharoun S, et al.: Generative artificial intelligence to transform inpatient discharge summaries to patient-friendly language and format. *JAMA Netw Open.* 2024, 7:e240357. [10.1001/jamanetworkopen.2024.0357](https://doi.org/10.1001/jamanetworkopen.2024.0357)
6. Williams CY, Bains J, Tang T, et al.: Evaluating large language models for drafting emergency department discharge summaries [PREPRINT]. *medRxiv.* 2024, [10.1101/2024.04.03.24305088](https://doi.org/10.1101/2024.04.03.24305088)
7. Tran BD, Chen Y, Liu S, Zheng K: How does medical scribes' work inform development of speech-based

- clinical documentation technologies? A systematic review. *J Am Med Inform Assoc.* 2020, 27:808-17. [10.1093/jamia/ocaa020](https://doi.org/10.1093/jamia/ocaa020)
8. Lin S: The present and future of team documentation: the role of patients, families, and artificial intelligence. *Mayo Clin Proc.* 2020, 95:852-5. [10.1016/j.mayocp.2020.01.034](https://doi.org/10.1016/j.mayocp.2020.01.034)
9. Chen M, Decary M: Artificial intelligence in healthcare: an essential guide for health leaders . *Healthc Manage Forum.* 2020, 33:10-8. [10.1177/0840470419873123](https://doi.org/10.1177/0840470419873123)
10. Warner E, Lee J, Hsu W, Syeda-Mahmood T, Kahn CE Jr, Gevaert O, Rao A: Multimodal machine learning in image-based and clinical biomedicine: survey and prospects. *Int J Comput Vis.* 2024, 132:3753-69. [10.1007/s11263-024-02032-8](https://doi.org/10.1007/s11263-024-02032-8)
11. Van Veen D, Van Uden C, Blankemeier L, et al.: Adapted large language models can outperform medical experts in clinical text summarization. *Nat Med.* 2024, 30:1134-42. [10.1038/s41591-024-02855-5](https://doi.org/10.1038/s41591-024-02855-5)
12. Lindvall C, Deng CY, Moseley E, et al.: Natural language processing to identify advance care planning documentation in a multisite pragmatic clinical trial. *J Pain Symptom Manage.* 2022, 63:e29-36. [10.1016/j.jpainsymman.2021.06.025](https://doi.org/10.1016/j.jpainsymman.2021.06.025)
13. Marshall TL, Nickels LC, Brady PW, Edgerton EJ, Lee JJ, Hagedorn PA: Developing a machine learning model to detect diagnostic uncertainty in clinical documentation. *J Hosp Med.* 2023, 18:405-12. [10.1002/jhm.13080](https://doi.org/10.1002/jhm.13080)
14. Liu J, Wang C, Liu S: Utility of ChatGPT in clinical practice . *J Med Internet Res.* 2023, 25:e48568. [10.2196/48568](https://doi.org/10.2196/48568)
15. Giorgi J, Toma A, Xie R, Chen SS, An KR, Zheng GX, Wang B: WangLab at MEDIQA-Chat 2023: clinical note generation from doctor-patient conversations using large language models. *Proceedings of the 5th Clinical Natural Language Processing Workshop. Association for Computational Linguistics, Toronto, Canada; 2023.* 323-34.
16. Kernberg A, Gold JA, Mohan V: Using ChatGPT-4 to create structured medical notes from audio recordings of physician-patient encounters: comparative study. *J Med Internet Res.* 2024, 26:e54419. [10.2196/54419](https://doi.org/10.2196/54419)
17. Meng X, Yan X, Zhang K, et al.: The application of large language models in medicine: a scoping review . *iScience.* 2024, 27:109713. [10.1016/j.isci.2024.109713](https://doi.org/10.1016/j.isci.2024.109713)
18. Waisberg E, Ong J, Masalkhi M, et al.: GPT-4: a new era of artificial intelligence in medicine . *Ir J Med Sci.* 2023, 192:3197-200. [10.1007/s11845-023-03377-8](https://doi.org/10.1007/s11845-023-03377-8)
19. Clough RA, Sparkes WA, Clough OT, Sykes JT, Steventon AT, King K: Transforming healthcare documentation: harnessing the potential of AI to generate discharge summaries. *BJGP Open.* 2024, 8:BJGPO.2023.0116. [10.3399/BJGPO.2023.0116](https://doi.org/10.3399/BJGPO.2023.0116)
20. Goss FR, Blackley SV, Ortega CA, et al.: A clinician survey of using speech recognition for clinical documentation in the electronic health record. *Int J Med Inform.* 2019, 130:103938. [10.1016/j.ijmedinf.2019.07.017](https://doi.org/10.1016/j.ijmedinf.2019.07.017)
21. Sushil M, Kennedy VE, Mandair D, Miao BY, Zack T, Butte AJ: CORAL: expert-curated oncology reports to advance language model inference. *NEJM AI.* 2024, 1:[10.1056/AIIdbp2300110](https://doi.org/10.1056/AIIdbp2300110)
22. Can AI fix medical records?. (2024). Accessed: October 17, 2024: <https://www.nature.com/articles/d41586-019-03848-y>.
23. Tang L, Sun Z, Idnay B, et al.: Evaluating large language models on medical evidence summarization. *NPJ Digit Med.* 2023, 6:158. [10.1038/s41746-023-00896-7](https://doi.org/10.1038/s41746-023-00896-7)
24. Zhang L, Negrinho R, Ghosh A, Jagannathan V, Hassanzadeh HR, Schaaf T, Gormley MR: Leveraging pretrained models for automatic summarization of doctor-patient conversations. *Findings of the Association for Computational Linguistics: EMNLP 2021. Association for Computational Linguistics, Punta Cana, Dominican Republic; 2021.* 3693-712.
25. Roberts K: Large language models for reducing clinicians' documentation burden . *Nat Med.* 2024, 30:942-3. [10.1038/s41591-024-02888-w](https://doi.org/10.1038/s41591-024-02888-w)
26. Patel SB, Lam K: ChatGPT: the future of discharge summaries? . *Lancet Digit Health.* 2023, 5:e107-8. [10.1016/S2589-7500\(23\)00021-3](https://doi.org/10.1016/S2589-7500(23)00021-3)
27. Ando K, Okumura T, Komachi M, Horiguchi H, Matsumoto Y: Is artificial intelligence capable of generating hospital discharge summaries from inpatient records?. *PLOS Digit Health.* 2022, 1:e0000158. [10.1371/journal.pdig.0000158](https://doi.org/10.1371/journal.pdig.0000158)
28. Florig ST, Corby S, Rosson NT, Devara T, Weiskopf NG, Gold JA, Mohan V: Chart completion time of attending physicians while using medical scribes. *AMIA Annu Symp Proc.* 2021, 2021:457-65.
29. Liu Y, Ju S, Wang J: Exploring the potential of ChatGPT in medical dialogue summarization: a study on consistency with human preferences. *BMC Med Inform Decis Mak.* 2024, 24:75. [10.1186/s12911-024-02481-8](https://doi.org/10.1186/s12911-024-02481-8)
30. Krishna K, Khosla S, Bigham JP, Lipton ZC: Generating SOAP notes from doctor-patient conversations using modular summarization techniques. *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers). Association for Computational Linguistics, Stroudsburg (PA); 2021.* 4958-72.
31. Page MJ, McKenzie JE, Bossuyt PM, et al.: The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ.* 2021, 372:n71. [10.1136/bmj.n71](https://doi.org/10.1136/bmj.n71)
32. The AI scribe is the future of medical documentation: reducing clinicians' workloads . (2024). Accessed: October 17, 2024: <https://www.heidihealth.com/blog/ai-scribe>.
33. Preiksaitis C, Sinsky CA, Rose C: ChatGPT is not the solution to physicians' documentation burden . *Nat Med.* 2023, 29:1296-7. [10.1038/s41591-023-02341-4](https://doi.org/10.1038/s41591-023-02341-4)
34. Duffourc M, Gerke S: Generative AI in health care and liability risks for physicians and safety concerns for patients. *JAMA.* 2023, 330:313-4. [10.1001/jama.2023.9630](https://doi.org/10.1001/jama.2023.9630)
35. Zernikow J, Grassow L, Gröschel J, Henrion P, Wetzel PJ, Spethmann S: Clinical application of large language models: does ChatGPT replace medical report formulation? An experience report [Article in German]. *Inn Med (Heidelb).* 2023, 64:1058-64. [10.1007/s00108-023-01600-3](https://doi.org/10.1007/s00108-023-01600-3)
36. Tierney AA, Gayre G, Hoberman B, et al.: Ambient artificial intelligence scribes to alleviate the burden of clinical documentation. *NEJM Catal Innov Care Deliv.* 2024, 5: [10.1056/CAT.23.0404](https://doi.org/10.1056/CAT.23.0404)
37. Nayak A, Alkaitis MS, Nayak K, Nikolov M, Weinfurt KP, Schulman K: Comparison of history of present

illness summaries generated by a chatbot and senior internal medicine residents. JAMA Intern Med. 2023, 183:1026-7. [10.1001/jamainternmed.2023.2561](https://doi.org/10.1001/jamainternmed.2023.2561)