Applied Medical Informatics
Original Research Vol. 46, No. 1 /2024, pp: 5-11

Leveraging ChatGPT for Digital Healthcare Speech Writing

Ariana-Anamaria CORDOŞ [1, 2*], Sebastian-Aurelian ŞTEFANIGĂ[3], and Călin MUNTEAN[4]

[1] Romanian Society of Medical Informatics, 300222 Timişoara, Romania
[2] Department of Practical Abilities, "Iuliu Haţieganu" University of Medicine and Pharmacy, 400337 Cluj-Napoca, Romania
[3] Department of Computer Science, Faculty of Mathematics and Computer Science of West University of Timisoara, 300223, Timisoara, Romania.
[4] Department of Functional Sciences, Medical Informatics and Biostatistics, "Victor Babeş" University of Medicine and Pharmacy, 300172 Timişoara, Romania
E-mails: cordos.ariana@umfcluj.ro; sebastian.stefaniga@e-uvt.ro; cmuntean@medinfo.umft.ro

* Author to whom correspondence should be addressed;

Received: 22 March 2024/Accepted: 31 March 2024/ Published online: 31 March 2024

Abstract
Introduction: Conversational agents, powered by artificial intelligence (AI) and natural language processing (NLP), are gaining traction in healthcare for various purposes. This study aimed to assess the potential of ChatGPT, a large language model, for generating health-related communication. *Methods:* ChatGPT was prompted with identical queries related to healthcare challenges, digital hospital benefits, and principles for organizing a digital hospital. The prompts were presented in January 2023 and 2024, and the resulting outputs were compared. *Results:* ChatGPT's outputs demonstrated a significant increase in complexity and detail between 2023 and 2024. The 2023 content catered to a general audience, offering clear and concise explanations suitable for raising awareness. In contrast, the 2024 outputs delved deeper into the topics, presenting multifaceted solutions and a broader global perspective. Notably, the 2024 outputs incorporated more technical language, potentially limiting accessibility for a lay audience. Both the 2023 and 2024 versions could benefit from including data to support the claims presented. This study suggests that ChatGPT exhibits promise for generating healthcare content. The observed increase in complexity over time aligns with advancements in AI language models. However, considerations like data inclusion, audience adaptation, and the use of clear language remain crucial. *Conclusion:* This research contributes to the growing body of knowledge on large language models in healthcare communication. The findings highlight the potential of ChatGPT for content creation while emphasizing the importance of human oversight and tailoring content for the target audience. Future research can explore these aspects in greater detail to optimize the use of AI in healthcare communication strategies.

Keywords: Evaluation Study; Narrative Medicine; Large Language Models (LLMs); ChatGPT

Introduction

Artificial intelligence (AI), encompassing both the capability of machines to be intelligent and the ongoing effort to create such machines, has become a rapidly growing field with the potential to revolutionize various sectors, including healthcare communication. In the sense of intelligence exhibited by machines, AI refers to the ability of a computer or robot to perform tasks that typically require human intelligence, such as learning, reasoning, problem-solving, and decision-making. This includes activities like recognizing objects in images, translating languages, writing different kinds of creative content, and even controlling robots in complex environments. AI is also a branch of computer science dedicated to developing intelligent machines and algorithms. Researchers in this

field explore various techniques and approaches to enable machines to exhibit intelligent behavior. This includes areas like machine learning, natural language processing, computer vision, and robotics.

Conversational agents represent a subset of digital innovations that can be tailored for tackling contemporary healthcare challenges, such as the scarcity of medical personnel, which restricts healthcare service availability and accessibility [1]. Leveraging AI, including machine learning (ML, a statistical method for training data models to make predictions based on various features) and natural language processing (NLP; the capability to comprehend and analyze speech and text), conversational agents interact with humans through speech, text, or other input/output channels on mobile, web, or audio platforms [2]. Many of these agents utilize NLP to allow users to communicate with them naturally, as they would with another person. Subsequently, the agent analyzes the received information and responds in a conversational manner [3].

Various types of conversational agents possess distinct functionalities, benefits, and constraints. AI chatbots, for instance, are software applications that emulate human-like messaging interactions when engaging with users. They can be embedded in social media platforms, messaging services, websites, mobile applications, and more. AI chatbots commonly handle simple tasks like dispensing information or aiding users in executing administrative tasks without requiring them to switch to another platform [4]. Additionally, conversational agents encompass voice bots [5] and interactive voice assistants [6].

Bots have found utility across diverse areas within the healthcare sector, supporting both professionals and the general audience. For patients, they can address frequently asked medical inquiries and enhance treatment adherence remotely [6]. However, from a physician's viewpoint, current conversational agents are not yet suitable for clinical use [7]. Nevertheless, development guidelines have been clearly delineated: automate repetitive administrative workflows, collate comprehensive patient data, and facilitate optimal clinical decision-making while adhering to ethical regulations.

While conventional chatbots are limited to specific tasks such as understanding and responding to user inputs using pre-established responses and rule-based algorithms, conversational AI encompasses more advanced AI-based chatbots and systems. Chatbots require conversational AI to enhance their ability to comprehend human language, engage in personalized two-way interactions with users, and meet customer expectations [8].

Despite extensive research on chatbot applications in scientific and medical realms, most studies have focused narrowly on specific medical specialties, agent types, or functionalities [1,9,10,11,12].

Since its rollout, in November 2020, ChatGPT [13] received general improvements on performance across various topics, better factuality, and reduced likelihood of refusing to answer questions twice in January 2023 and once in December 2023. In the first January release the ability to stop the generation of response mid-way was also added. In the December release the ability to view past conversations, rename saved conversations and delete unwanted ones was also rolled out.

Large language models (LLMs) such as ChatGPT are catalyzing profound transformations across diverse domains, with healthcare being particularly ripe for innovation. A pivotal review published in 2023 [14] underscored ChatGPT's burgeoning potential in revolutionizing healthcare education, research paradigms, and clinical practices. The multifaceted advantages elucidated emanate from ChatGPT's adeptness in distilling intricate subject matter into concise summaries, conducting exhaustive literature reviews, and even crafting scientific manuscripts. Furthermore, empirical evidence has illuminated ChatGPT's prowess in furnishing precise diagnoses and treatment recommendations, thereby augmenting the clinical acumen of healthcare practitioners (albeit undergirded by indispensable human oversight). Notably, ChatGPT's capacity to sift through voluminous datasets and furnish comprehensive reports holds immense promise in optimizing healthcare workflows. Nonetheless, this seminal research accentuates the imperative of prudently navigating ChatGPT's limitations and fostering its judicious integration to safeguard patient welfare and uphold ethical imperatives. Pertinent concerns encompass its imperfections in factual accuracy, reasoning, and source referencing, warranting concerted efforts toward refinement. Additionally, the specter of inherent biases in training data necessitates vigilant scrutiny. Acknowledging ChatGPT's potential struggles with intricate medical concepts and its divergence from the nuanced understanding of human healthcare professionals is paramount. Moreover, apprehensions surrounding plagiarism, potential educational misuse, and the exigency of robust regulatory frameworks underscore the exigency of delineating clear guidelines for its conscientious utilization.

Effective communication of healthcare research findings is essential for advancing the field. However, not all researchers may have the time or expertise to craft engaging presentations. This study investigates ChatGPT's ability to generate content for oral presentations, exploring its potential to democratize access to high-quality communication tools for healthcare professionals along with fostering the efficiency of healthcare communication.

This research examined ChatGPT's potential as a tool for creating health communication materials. It investigated ChatGPT's ability to generate content suitable for oral presentations, such as research talks delivered at conferences or lectures. To assess this, the quality of text produced by ChatGPT when prompted with specific healthcare-related queries was analyzed.

Materials and Methods

The objective of the study was to perform a quantitative and a qualitative analysis of the text output received from the free available version of ChatGPT at two different points in time. The same study methodology was replicated one year apart. Both times the PowerPoint presentation "Hospital Digitalization: Challenges and Benefits" previously presented by the main author as communication support during the RoHealth event "Romanian Society of Medical Informatics: Who & What" [15] was used.

The PowerPoint presentation consisted of a title slide, an author slide, 18 content slides, and a thank you slide. The content slides were categorized into three sections: challenges, digital hospital, and benefits, each with a corresponding subtitle slide. The challenges section comprised individual slides addressing various issues such as the global healthcare workforce crisis, increasing complexity of care, wasted capacity, and others. The digital hospital section spanned three slides discussing workflow adaptation, digital hospital transformation, and training. The benefits section included four slides, three of which had specific titles highlighting the advantages of hospital digitalization. The fourth slide contained only imagery, and the first sentence of the associated speech was used as the ChatGPT query. Each content slide featured a title and representative imagery, with the presenter reading the accompanying text from the PowerPoint notes.

The experiment involved copy-pasting the title extracted from each PowerPoint slide (Table 1) in the ChatGPT's query filed to generate text narrations.

Table 1. Titles of PowerPoint slides

Titles of PowerPoint slides
The global healthcare workforce crisis
Increasing complexity of care
Wasted capacity
Ever-growing workloads
Medical devices not interconnected
Distractions, false alarms, and alarm fatigue
Workflow inefficiencies
Workflow adaptation
Digital hospital transformation
Training
Increase the quality and accuracy of medical records
Decrease the incidence of medical errors
Avoid deficiencies during handover
Shorten hospital length of stay
A digital hospital should be organized in a way that consistently and sustainably avoids mistakes and inefficiencies.

This experiment was conducted twice, with the first iteration on January 6, 2023, and the second iteration on January 6, 2024. Two separate Excel sheets were created for each experiment, with each query entered in a separate row. The text generated by ChatGPT in response to each query was then recorded in the corresponding Excel cell, resulting in a tabulated dataset for analysis.

The quantitative analysis consisted of the word count of each of the responses which was further described by means of averages and standard deviations, median and interquartile ranges, minimum, maximum and the sum of all the words.

The qualitative analysis of the text from the two instances was performed acknowledging the focus on differences.

Results

Quantitative Analysis

Upon initial review, the 2023 text appeared more succinct, presenting specific factors in bullet points (Table 2). The suggested answers provided were broad and globally limited, conveyed in a more accessible language, easier to understand for a general audience. Conversely, the 2024 text was more extensive, offering detailed analyses and multifaceted solutions with a broader global perspective offering a deeper analysis, suitable for seeking a thorough understanding of the problem and potential solutions.

Table 2. Descriptive of the number of words for each output

	Data output 2023	**Data output 2024**
Mean ± SD*	1368.36 ± 218.07	3598.64 ± 385.14
Median [Q1 - Q3]*	1312.5 [1221 - 1455.75]	3666.5 [3258.5 - 3799.75]
Min – Max*	1095 - 1855	3067 - 4415
Sum	19157	50381
t test (p)	18.85 (<0.0001)	

*SD: Standard Deviation; Q1: quartile 1 ; Q3 : quartile 3 ; Min : minimum ; Max : maximum

Qualitative Analysis

Challenges Section

For the challenges slides, the 2023 text was deemed appropriate for an introductory understanding or awareness-raising, while the 2024 text was better suited for in-depth analysis and exploring solution development, focusing on healthcare systems and management. Both texts discussed the global healthcare workforce crisis as the first challenge, but in slightly different ways. The 2023 output seemed suitable for a quick overview, whereas the 2024 output was more conducive to deeper understanding or presentation. On the 2nd challenge, both texts discussed the rising complexity of healthcare. The 2024 text had a clearer and more organized structure, making it easier to follow the arguments. It also stood apart employing technical language, potentially challenging for a general audience. Both could have benefited from data inclusion to support claims. The 3rd challenge discussed wasted capacity in healthcare. The 2023 text could have benefited from specific examples, while the 2024 text could have been adapted for a wider audience by providing simpler explanations of technical terms. For the fourth challenge, the 2023 text focused on "increasing workload," while the 2024 text used "ever-growing workload." Additionally, the 2023 text highlighted negative consequences for both professionals and patients, whereas the 2024 text emphasized professional impact. Regarding the lack of interconnectedness among medical devices, the 2023 text used "not interconnected focusing more on the impact on patient safety and efficiency," while the 2024 text specified "lack of interconnectedness" and included data management challenges. The 6th challenge discussed distractions, false alarms, and alarm fatigue. The 2023 text mentioned "distractions" first, while the 2024 text listed it last. The 2023 text used simpler terms like "false alarms" and "alarm fatigue", while the 2024 text was more technical like "non-actionable alarms" and "alarm desensitization". For the last challenge on workflow inefficiencies, the 2023 text focused on "workflow inefficiencies," while the 2024 text delved into specific problems like "fragmented systems" and "manual processes" with a focus on organizational aspects. In 2023, it emphasized the impact on both healthcare providers and patients, while in 2024 the focus was more on the organizational and

systemic aspects.

Digital Hospital

Regarding the digital hospital sections, the 2023 text was more appropriate for awareness-raising, while the 2024 text provided a deeper understanding. For the 1st part, both texts discussed workflow adaptation. The 2023 text used the simpler term "workflow adaptation", while the 2024 text focused on the "ability to modify" and provided specific strategies. Text of 2023 focused on generic steps in the adaptation process, while the text of 2024 delved deeper into specific characteristics and strategies for successful implementation. The text generated in 2024 seemed more appropriate for understanding the nuances and strategies for successfully adapting workflows in healthcare. In the second part about digital hospital transformation, the 2023 text emphasized challenges, the role of leadership, vision, and stakeholder involvement, while the 2024 text explored potential benefits, specific technologies, and implementation strategies. Furthermore, the text of 2024 provided a better understanding of the specific strategies and potential of digital technologies to revolutionize healthcare delivery. The last part about training saw the 2023 text focusing on employer roles in providing training emphasizing the importance of staying up to date, while the 2024 text delved into specific skills, knowledge, and behaviors, targeted by training, and its impact on various aspects of healthcare delivery and professional development, with a focus on educational institutions and professional organizations with more specific examples of how training helps achieve that goal.

On the benefits slides, the 2023 text was suitable for awareness-raising, while the 2024 text provided a more detailed introduction. The 1st benefit discussed strategies for improving medical records and decreasing medical errors. The 2023 text focused on the role of digital technologies in reducing errors through accuracy and data quality, while the 2024 text included a wider range of approaches from standardized documentation, training, and patient involvement to proactive error prevention strategies like Bar Code Medication Administration (BCMA) and root cause analysis. The 2nd benefit focused on improving handover processes. The 2023 text emphasized avoidance of deficiencies, while the 2024 text went beyond and highlighted strategies like prioritizing critical information, verifying understanding, and using standardized times to ensure effective communication. The text generated in 2023 used a simpler list of strategies, while 2024's text delved into specific tools and frameworks like SBAR (Situation, Background, Assessment, Recommendation) and structured templates. Also, in 2024 there was a focus on the importance of face-to-face interaction and documentation, while in 2023 this was just briefly mentioned.

The last benefit discussed strategies for shortening hospital length of stay (LOS). The 2024 text elaborated on specific strategies, their detailed implementation, and their impact on improving patient flow and resource utilization and emphasized the importance of continuous quality improvement for ongoing optimization, while the 2023 text focused on internal streamlining and care continuity. In 2023, there was a mention about early mobility but doesn't emphasize its role in patient recovery and LOS reduction, while in 2024 early mobility was highlighted as a separate strategy.

The conclusion talked about the principles and methods for organizing a digital hospital to avoid mistakes and inefficiencies. Text of 2023 emphasized the importance of continuous improvement as a principle, while text of 2024 elaborated on specific methods for implementing ongoing improvement efforts. Text of 2023 briefly mentioned user-centered design, while text of 2024 highlighted its importance and provided specific guidance for incorporating it. The former added emphasis on patient engagement and empowerment, automation and AI, and feedback and communication channels as crucial aspects of effective organization.

A comprehensive analysis of the contrasting disparities identified in the content generated by ChatGPT in 2023 and in 2024 outputs can be found in Table 3.

Table 3. Summary of data output per data input

Data output 2023	**Data output 2024**
The global healthcare workforce crisis	
• Specific factors: aging population, brain drain, lack of funding, inadequate training, pandemics. • Broader solutions: increased funding, incentives, improved training.	• Multi-faceted solutions: addresses education, retention, recruitment, collaboration, technology, and social determinants. • The topic has global coverage.
Increasing complexity of care	
• 4 contributing factors: aging population, medical advancements, chronic conditions, and multimorbidity. • Broad challenges faced by healthcare systems and patients. • Solutions are general suggestions like increased expertise and resources.	• 6 contributing factors: medical technology advancements, chronic disease rise, multimorbidity, patient expectations, care fragmentation, and social determinants of health. • Specific challenges associated with each factor, emphasizing the need for a holistic approach. • Multiple solutions tailored to each challenge: interdisciplinary collaboration, care coordination, technological adoption, workforce development, and patient engagement. • There are diverse needs of various patient populations.
Wasted capacity	
• Resources: beds, operating rooms, and equipment. • Less detailed: summary of reasons with fewer specific factors. • Simple solutions: optimizing resource utilization, improving scheduling, increasing access to care.	• Broad scope: includes time, personnel, equipment, and facilities. • More detailed: 6 specific factors contributing to wasted capacity. • Comprehensive solutions: data-driven approaches, process improvement, and continuous quality improvement. • More technical: terms like "data analytics," "lean management," and "Six Sigma."
Ever-growing workloads	
• 4 contributing factors: aging population, healthcare professional shortage, increasing complexity of care, and poor time management. • Broad solutions: increase workforce, improve time management, optimize resources.	• 6 specific factors: increasing patient demand, rising chronic disease burden, medical technology advancements, administrative burdens, staffing shortages, and unsustainable work environment. • Diverse solutions: workforce investment, administrative streamlining, team-based care, work-life balance initiatives, addressing social determinants of health, and advocating for policy changes. • More technical: terms like "chronic disease burden" and "unsustainable work environment."
Medical devices not interconnected	
• 2 reasons: different technologies/protocols and lack of communication capability. • Focused on benefits: improved patient safety and efficiency. • Broad strategy: adoption of standards and protocols.	• 6 specific factors: proprietary systems, legacy systems, security concerns, regulatory hurdles, vendor lock-in, and lack of standardization. • Diverse solutions: tailored strategies for each factor, emphasizing standardization, regulatory support, manufacturer collaboration, healthcare provider engagement, cybersecurity measures, and education/training. • More technical: terms like "proprietary systems," "vendor lock-in," and "cybersecurity."
Distractions, false alarms, and alarm fatigue	
• Definition of each term offering simple explanations of distractions, false alarms, and alarm fatigue.	• Detailed explanations: goes deeper into the specific causes and consequences of each issue.

Data output 2023	**Data output 2024**
• Consequences highlighting the impact on patient safety and suggesting broad strategies.	• Multi-faceted solutions: tailored specific strategies for addressing each factor, including technology, training, and cultural changes. • More technical: terms like "patient desensitization," "alarm fatigue," and "smart alarms."
Workflow inefficiencies	
• 5 general causes: poor communication, inefficient processes, lack of standardization, technology issues, and lack of training. • Consequences: Briefly mentions negative impacts on both providers and patients. • Broad solutions: Improve communication, streamline processes, standardize practices, use technology effectively, and provide training.	• Detailed explanations going deeper into the specific causes and consequences of each inefficiency. • Multifaceted solutions: tailors specific strategies for each factor, emphasizing technology, communication, workflow analysis, and continuous improvement. • More technical: terms like "fragmented systems," "manual processes," and "workflow optimization."
Workflow adaptation	
• Common reasons for adaptation: changes in healthcare environment, inefficiencies, and bottlenecks. • Offers generic strategies: identifying current workflow, analyzing for improvement, developing and implementing plans, evaluating results.	• Going beyond reasons, focusing on key characteristics of successful adaptation, like flexibility and agility. • Specific strategies: tailored approaches for each characteristic, emphasizing data-driven decision making, collaboration, technology, and change management. • More technical: terms like "interdisciplinary collaboration," "technology integration," and "change management." • Deeper analysis: suitable for those seeking a thorough understanding of how healthcare organizations can effectively adapt their workflows to changing demands.
Digital hospital transformation	
• Focused on broad categories: EHRs, telemedicine, integration of devices, and other digital tools. • Highlight key benefits: Improved patient safety and efficiency of care delivery.	• Described concepts' benefits for various stakeholders. • 10 key areas for implementation, including EHRs, telehealth, CDSS, HIS, patient engagement, workflow automation, AI/analytics, IOT/wearables, cybersecurity, and change management. • Highlight transformative potential: the impact on patient care, operational efficiency, and overall patient experience. • More technical: Uses terms like "clinical decision support systems," "health information exchange," and "artificial intelligence."
Training	
• A basic overview of training in healthcare and its importance for individuals and organizations. • Broad categories: formal programs, different delivery formats, and benefits for both professionals and organizations.	• A comprehensive analysis of various aspects of training in healthcare and its impact on different areas. • 10 key aspects of training in healthcare, including various types of skills, communication, cultural competency, leadership, and ethics. • Specific benefits: the role of training in patient care, technology adoption, collaboration, and overall quality improvement. • More technical: terms like "continuing education," "interdisciplinary training," "cultural competency," and "simulation-based learning."
Increase the quality and accuracy of medical records Decrease the incidence of medical errors	
• A basic overview of the problem and 5 potential solutions using digital technologies.	• 10 specific strategies across various areas.

Data output 2023	**Data output 2024**
• Broad categories: EHRs, CDSS, standardized terminology, interoperability, and data quality.	• Detailed explanations going deeper into each strategy, outlining its benefits and implementation aspects. • More technical: terms like "standardized documentation templates," "clinical decision support systems," and "barcode medication administration." • Deeper analysis: suitable for those seeking a thorough understanding of how to improve medical records and prevent errors.
Avoid deficiencies during handover	
• Broad categories: standardized protocols, checklists, technology, training, and patient involvement. • More accessible: Easier to understand for a general audience.	• Detailed explanations goes deeper into each strategy, outlining specific aspects of implementation and benefits. • More technical: terms like "structured handover tools," "essential information," "SBAR communication framework," and "handover huddles." • Deeper analysis: suitable for those seeking a thorough understanding of improving handover processes and their impact on patient care.
Shorten hospital length of stay	
• Broad categories: streamlining processes, discharge planning, telemedicine, early mobility programs, and care coordination.	• 10 specific strategies across various areas. • Detailed explanations: each strategy elaborately described with its benefits and implementation aspects. • More technical: terms like "early discharge planning," "multidisciplinary rounds," "care pathways," "hospital-at-home programs," and "ambulatory and same-day surgery." • Deeper analysis: suitable for those seeking a thorough understanding of how to shorten length of stay and its impact on healthcare efficiency and patient satisfaction.
A digital hospital should be organized in a way that consistently and sustainably avoids mistakes and inefficiencies.	
• Broad categories: standardization, interoperability, data quality, continuous improvement, and training.	• 10 specific considerations for implementation. • Detailed explanations: each consideration elaborated with its benefits and best practices for implementation. • More technical: terms like "interoperable systems integration," "user-centered design," "clinical decision support systems," and "real-time analytics." • Deeper analysis: suitable for those seeking a thorough understanding of how to optimize organization and avoid mistakes in a digital hospital.

*EHR: electronic health records, AI: artificial intelligence; CDSS: clinical decision support system, IOT: internet of things

Discussions

As to the authors' knowledge, comparing LLM's answers over a period in the space of digital healthcare has not been scientifically documented thus far. The comparison of ChatGPT's outputs from 2023 and 2024 reveals significant differences in the depth and complexity of the generated content on the topic addressed: challenges in healthcare, benefits of digital hospitals, and principles for organizing a digital hospital. This aligns with the expected improvements in AI LLM over time, as they get trained on more diverse and extensive datasets and fine-tuned using advanced techniques [13]. These results further underscore the rapid evolution of ChatGPT and its potential to generate increasingly sophisticated and nuanced content. ChatGPT's responses were generated based on patterns learned from a diverse range of text available, in reputable sources such as scientific literature, textbooks, or websites, up to its last training cut-off in 2022. While ChatGPT strives for accuracy through constant system

updates, its responses may not reflect the most recent advancements in specific domains of care. For healthcare-related inquiries, consulting up-to-date sources like medical professionals, peer-reviewed journals, and reputable health organizations remains paramount. Therefore, ChatGPT could be used as a tool for information and education, not as a replacement for professional medical advice. However, for less rapidly evolving fields like the digitalization of care, where both the challenges and technologies involved progress at a slower pace, ChatGPT's ability to provide well-supported content for forming an opinion remains valuable. As previously shown in the existing literature, also in this research ChatGPT was able to produce understandable reasoning and provided relevant arguments which increased confidence in its trustworthiness and comprehensibility [14]. Furthermore, this hereby research unveils ChatGPT's inability to dynamically adjust the tone of voice and complexity of verbiage based on the intended audience and purpose of the generated content. This underscored the need for human oversight and intervention to ensure the accuracy and appropriateness of the generated content.

This research demonstrates the utilization of ChatGPT presents promising prospects in healthcare communication. However, its application gives rise to distinct ethical considerations. ChatGPT possesses the capability to produce text that exudes authority, even if it lacks factual accuracy. In the absence of explicit disclaimers or human supervision, patients may misconstrue their outputs as professional medical counsel. Trained on an extensive corpus of text and code, ChatGPT's data has a temporal boundary, rendering it potentially outdated in the swiftly evolving landscape of healthcare knowledge. Consequently, there exists a risk of generating erroneous or deceptive information, thereby potentially endangering receivers of the communication produced. Moreover, ChatGPT's responses are influenced by the biases inherent in its training data, inadvertently perpetuating disparities that may disproportionately impact the auditor of that communication. Additionally, broader societal ramifications, including the implications of AI on the scientific, educational, and healthcare workforce, warrant consideration. While LLMs like ChatGPT hold the promise of augmenting efficiency and productivity, they simultaneously raise apprehensions regarding job displacement and the devaluation of human expertise. The ethical deployment of AI in scientific, educational, and healthcare spaces necessitates striking a delicate balance between harnessing technological advancements to enhance patient outcomes and preserving the human-centric ethos of healthcare education and care delivery. Enhancing ChatGPT's contributions to scientific healthcare communication involves several key recommendations. Firstly, regular updates to its training data are crucial to ensure alignment with the latest medical evidence and standards of care. This entails integrating datasets from recent research, clinical databases, and guidelines to enhance accuracy and relevance. Additionally, ensuring diversity and representation within the training data helps mitigate biases and fosters inclusivity in ChatGPT's outputs. Continuous learning mechanisms, such as real-time integration with medical databases or periodic retraining, enable ChatGPT to stay updated. Implementing a fact-checking layer within ChatGPT verifies the accuracy of its outputs against trusted medical sources before presentation to users. A confidence score system can aid users in assessing the reliability of ChatGPT's statements. Moreover, encouraging source citation whenever possible allows for independent verification of claims. In cases where the above stated are not feasible, involving healthcare professionals oversight to review and validate responses is enhancing reliability and appropriateness, especially in critical scenarios. Lastly, providing user education and training materials on interpreting and critically evaluating ChatGPT's outputs empowers users to make informed decisions and discern reliable information from unreliable sources.

Limitations

Already existing literature primarily focuses on comparing different large language models like ChatGPT, their capabilities, applications, and performance. They discuss the fine-tuning of these models on specific datasets and the implications of this for their performance. They also highlight the need for accurately evaluating large language models and understanding their capabilities and risks. On the other hand, the research at hand specifically focuses on comparing the outputs of the freely available ChatGPT at two different points in time (2023 and 2024). The depth and complexity of the generated content was analyzed, its suitability for different audiences, and the potential improvements in the model over time. The analysis performed in this research paper is more specific as it involves a detailed comparison of the content generated by ChatGPT in response to the same prompts in 2023 and 2024, providing a thorough examination of the nuances in the language used, the depth of the content, and the focus of the discussion in each year. The existing literature does not delve into such specific comparisons of the same model's outputs at different points in time. Furthermore, there are several other limitations. The research

methodology is not structured to benchmark for accuracy, reliability, and relevance, as ChatGPT's outputs were not compared to responses from human healthcare experts. Next, the scope of the analysis was confined to the specific prompts extracted from the PowerPoint presentation. Moreover, ChatGPT's generative nature, which relies on random number generators for probability distributions, introduces inherent variability. This means regenerating the response to the same prompt, even at the same time, could result in different content length and structure. Finally, modifying the prompt itself could influence the output's length, complexity, structure, and language register (ranging from technical jargon for experts to simpler language for a lay audience).

Conclusions

From an application area perspective, this is one of the pioneering studies in the healthcare space that analyzed the generated content across time in the context of discussing challenges in healthcare, benefits of digital hospitals, and principles for organizing a digital hospital. The comparison of ChatGPT's outputs from 2023 and 2024 provides valuable insights into the evolution of AI LLMs and their potential applications in healthcare. It also highlights the challenges and considerations in using these models, such as the need for data inclusion, audience adaptation, and human oversight. As AI LLMs continue to evolve, future research could explore these aspects in more detail to maximize the benefits of these technologies.

List of Abbreviations: ML: machine learning, NLP: natural language processing, EHR: electronic health records, AI: artificial intelligence; CDSS: clinical decision support system, IOT: internet of things, LLM: large language model.

Author Contributions: AAC defined the research's aim and the experiment's design. SS carried out the experiments. AAC and SS participated in the design of the study and performed the analysis. CM coordinated and helped to draft the manuscript. All authors read and approved the final manuscript.

Funding: This research received no funding.

Ethics Statement: Not applicable.

Data Availability Statement: Raw data produced by ChatGPT during the experiments is available in Annex 1 attached to this paper.

Acknowledgments: In writing this manuscript ChatGPT (Jan 2023 version [16]) and ChatGPT (Jan 2024 version [17]) were used.

Conflict of Interest: The authors declare no conflict of interest.

References

1. Milne-Ives M, de Cock C, Lim E, Shehadeh M, de Pennington N, Mole G, et al. The Effectiveness of Artificial Intelligence Conversational Agents in Health Care: Systematic Review. J Med Internet Res. 2020; 22(10):e20346.
2. Campbell M. Beyond Conversational Artificial Intelligence. Computer. 2020 Dec; 53(12):121 - 5.
3. Montenegro JL dCCdRRR. Survey of conversational agents in health. Expert Syst Appl. 2019 Sep; 129:56 - 67.
4. Adamopoulou E, Moussiades L. Chatbots: History, technology, and applications. Machine Learning with Applications. 2020; 2.
5. Corkrey R, Parkinson L. Interactive voice response: review of studies 1989-2000. Behav Res Methods Instrum Comput. 2002 Aug; 34(3):342-53.
6. Hoy M. Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants. Med Ref Serv Q. 2018 Jan-Mar; 37(1):81 - 8.
7. Bibault J, Chaix B, Nectoux P, Pienkowsky A, Guillemasse A, Brouard B. Healthcare ex Machina: Are conversational agents ready for prime time in oncology? Clin Transl Radiat Oncol. 2019 Apr; 16:55 - 9.

8. Bhirud N, Tataale S, Randive S, Nahar S. A Literature Review On Chatbots In Healthcare Domain. International Journal of Scientific & Technology Research. 2019; 8:225-231.
9. Oh Y, Zhang J, Fang M, Fukuoka Y. A systematic review of artificial intelligence chatbots for promoting physical activity, healthy diet, and weight loss. Int J Behav Nutr Phys Act. 2021 Dec 11; 18(1):160.
10. Bin Sawad A, Narayan B, Alnefaie A, Maqbool A, Mckie I, Smith J, et al. A Systematic Review on Healthcare Artificial Intelligent Conversational Agents for Chronic Conditions. Sensors. 2022 Mar 29; 22(7):2625.
11. Pham K, Nabizadeh A, Selek S. Artificial Intelligence and Chatbots in Psychiatry. Psychiatr Q. 2022 Mar; 93(1):249 - 53.
12. Chew H. The Use of Artificial Intelligence-Based Conversational Agents (Chatbots) for Weight Loss: Scoping Review and Practical Recommendations. JMIR Med Inform. 2022 Apr 13; 10(4):e32578.
13. OpenAI. ChatGPT — Release Notes. [Online].; 2024 [cited 2024 March. Available from: https://help.openai.com/en/articles/6825453-chatgpt-release-notes.
14. Roumeliotis K, Tselikas N. ChatGPT and Open-AI Models: A Preliminary Review. Future Internet. 2023; 15(6):192.
15. RoHealth The Health and Bioeconomy Cluster; Romanian Society of Medical Informatics. Romanian Society of Medical Informatics: Who & What. [Online].; 2022 [cited 2023 Jan 5. Available from: https://rohealth.ro/en/course-detail/23.
16. OpenAI. ChatGPT. [Online]. [cited 2023 January. Available from: https://chat.openai.com/.
17. OpenAI. ChatGPT. [Online]. [cited 2024 January. Available from: https://chat.openai.com/.