Scholars Journal of Applied Medical Sciences
Abbreviated Key Title: Sch J App Med Sci
ISSN 2347-954X (Print) | ISSN 2320-6691 (Online)
Journal homepage: https://saspublishers.com

OPEN ACCESS

Medicine

Evaluation of the Reliability and Readability of Chatgpt-4.0 Responses Regarding Complaint of Dizziness Cases

Yagmur Gokhan Semerci, MD[1*]

[1]Consultant Emergency Medicine, PHCC, Qatar

DOI: https://doi.org/10.36347/sjams.2025.v13i01.014 | **Received:** 28.11.2024 | **Accepted:** 05.01.2025 | **Published:** 10.01.2025

***Corresponding author:** Yagmur Gokhan Semerci
Consultant Emergency Medicine, PHCC, Qatar

Abstract **Original Research Article**

This study investigates the reliability and readability of ChatGPT-4.0's responses to medical inquiries about dizziness, a condition affecting balance and movement perception, particularly in older populations. The research aims to assess the feasibility of AI applications in healthcare, focusing on their potential to alleviate the burden on healthcare systems by providing timely and accurate medical information. Utilizing Reddit as a data source, the study identifies common public concerns about dizziness and evaluates ChatGPT-4.0's responses using the Modified Dicern criteria, with assessments conducted by an experienced emergency medicine specialist. The findings highlight the AI's ability to deliver clear and relevant information, though they also reveal gaps in source transparency and data timeliness. Addressing these limitations is crucial for enhancing AI's role in healthcare, improving patient care, and reducing healthcare professionals' workload. The study underscores the importance of integrating AI with traditional medical practices to foster a collaborative approach to medical information dissemination and improve healthcare outcomes.

Keywords: ChatGPT-4.0, Dizziness, Medical Inquiries, AI in Healthcare, Reliability & Readability.

INTRODUCTION

Dizziness, often called vertigo, is characterized by a disturbance in the individual's perception of balance or movement and can be caused by a multitude of causes. *(Rana & Morren, 2013*)

This condition is especially common among older populations and significantly affects their quality of life. Research shows that the global prevalence rate of dizziness is around 10-30%, leading to increased visits to emergency departments and losses of both time and manpower. *(Lasisi & Gureje, 2010*) The frequent occurrence of dizziness puts a strain on health systems and reduces productivity in the professional lives of individuals.

In recent years, there has been a notable increase in the application of artificial intelligence (AI) in healthcare, which aims to increase accessibility to services and improve decision-making processes. *(Javanmard, n.d.)*

Artificial intelligence technologies are used in a variety of fields, including diagnosis, disease management, and personalized medical recommendations, thereby easing the workload of health professionals and ensuring that patients have immediate access to accurate information. Research has shown the potential benefits of AI-based systems in assessing complex conditions such as dizziness. *(Liu et al., n.d.)*

A comparative analysis of traditional medical approaches and AI applications can facilitate more effective and efficient outcomes in the delivery of health services *(Terry et al., n.d.)*. In this context, the integration of artificial intelligence promises to reduce costs and improve the quality of patient care, which is crucial for the future of health systems. The aim of this study is to investigate the feasibility and reliability of AI applications focusing specifically on CHATGPT-4.0 in addressing medical questions related to dizziness. By doing so, we aim to ease the burden on healthcare institutions while ensuring patients receive timely, accurate and appropriate medical treatment. To achieve this, we used Reddit, a public online forum, to evaluate the public's processes of seeking information about vertigo. The data collection phase involved identifying the 10 most commented issues related to "vertigo" posted on Reddit between January 1, 2023 and December 31, 2023. These selected questions were then directed to CHATGPT-4.0 with instructions to respond as a physician using medical resources for evaluation. The responses were then evaluated by an emergency

medicine specialist who used modified Dicern (MDICern) criteria to assess the quality of the answers given. This methodology is designed to improve our understanding of public information-seeking behaviors related to health issues and to evaluate the ability of AI applications to provide medical information. The findings aim to provide important data on the interaction between artificial intelligence and traditional medical practices and contribute to the ongoing dialogue about the future of health care. The results of this study could potentially inform the development of more reliable AI tools that can assist healthcare professionals and patients alike, fostering a collaborative approach to medical information dissemination.

METHOD

In this study, the Reddit platform, an online forum, was used to evaluate the public's information acquisition processes regarding bass freezing. The data collection process identified 10 questions that received the most comments using the keyword “dizziness” from January 1, 2023, to December 31, 2023. These selected questions were directed to the ChatGPT 4.0 (Chat Generative Pre-trained Transformer) artificial intelligence application and were asked to “respond to each question with the value of a doctor using medical resources.” In the second stage of the review, the answers obtained were evaluated by an emergency medicine specialist with 5 years of specialized experience in the field and board certification, using the Modified Dicern (mDicern) criteria. Each response was given a score of 1 to 5 (GQS) (see Table 1-2). This methodology was designed to understand the processes of knowledge acquisition in public health issues and to assess the adequacy of AI applications in providing medical information. The findings aim to provide important data for studying the interaction between AI and traditional medical practices.

Table 1: Modified Dicern Criteria

	mDICERN criteria	**score**
1	Are the aims clear?	1–5 point
2	Does it achieve its aims?	1–5 point
3	Is it relevant?	1–5 point
4	Is it clear what sources of information were used to compile the publication (other than the author or producer)?	1–5 point
5	Is it clear when the information used or reported in the publication was produced? 1–5 point	1–5 point
6	Is it balanced and unbiased?	1–5 point
7	Does it provide details of additional sources of support and information?	1–5 point
8	Does it refer to areas of uncertainty?	1–5 point

This table summarizes the modified Dicern criteria for evaluating publications, highlighting clarity, relevance and transparency in information sources. Global Quality Score Poor quality, poor fow of the site, most information missing, not at all useful for patients1Generally poor quality and poor fow, some information listed but many important topics missing, of very limited use to Patients2Moderate quality, suboptimal fow, some important information is adequately discussed but others poorly discussed, somewhat useful ForPatients3Good quality and generally good fow, most of the relevant information is listed, but some topics not covered, useful for patients4Excellent quality and excellent fow, very useful for patients5.

Table 2: Global Quality Score

Global Quality Score	
Poor quality, poor fow of the site, most information missing, not at all useful for patients	1
Generally poor quality and poor fow, some information listed but many important topics missing, of very limited use to patients	2
Moderate quality, suboptimal fow, some important information is adequately discussed but others poorly discussed, somewhat useful for patients	*3*
Good quality and generally good fow, most of the relevant information is listed, but some topics not covered, useful for patients	*4*
Excellent quality and excellent fow, very useful for patients	*5*

Evaluation of the Global Quality Score highlights varying degrees of usefulness and comprehensiveness in patient information and ultimately guides developments in healthcare communication strategies.

RESULT

Table 3: Evaluation according to mDicern criteria

	mDISCERN Criteria	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
1	Are the aims clear?	**4**	**4**	**4**	**5**	**5**	**5**	**5**	**5**	**5**	**5**
2	Does it achieve its aims?	4	4	4	5	5	5	5	5	5	5
3	Is it relevant?	5	5	5	5	5	5	5	5	5	5
4	Is it clear what sources of information were used to compile the publication (other than the author or producer)?	3	2	2	3	2	3	3	3	3	3
5	Is it clear when the information used or reported in the publication was produced?	3	2	2	3	2	3	3	3	3	3
6	Is it balanced and unbiased?	4	4	4	5	5	5	5	5	5	5
7	Does it provide details of additional sources of support and information?	2	2	2	2	3	4	4	4	4	4
8	Does it refer to areas of uncertainty?	3	3	3	4	4	4	4	4	4	4
	GQS Score (Moderate quality)	3.5	3.5	3.5	4	4.25	4.25	4.25	4.25	4.25	4.25

Clarity of views (Score: 4-5): Specialist physician, that the objectives set by AI practices are clear and understandable evaluated. This indicates that AI sets goals that are easily understood by physicians and patients when used in the healthcare field.

Achieving Goals (Score: 4-5): Again, high scores on achieving goals indicates that AI has a strong capacity to deliver expected results.

Interest (Score: 5): Publication or AI applications to the needs of health professionals and the expectations of patients caters to a very high level.

Clarity of Sources (Score: 2-3): Low scores, enough of the resources on which the information offered by AI is based reveals that it is not clearly indicated.

Time of Generation of Information (Score: 2-3): A similar ambiguity is observed here. Balanced and Neutrality (Score: 4-5): The expert physician has assessed that the information offered by AI is balanced and impartial.

Provision of Additional Resources of Support (Score: 2-4): Low score was given at the point of access to additional resources.

Determine Reference to Areas of Uncertainty (Score: 3-4): There was a moderate assessment of addressing uncertainties.

DISCUSSION

The expert physician considered that the objectives set by AI applications are clear and understandable. In this context, transparency and reliable sources of information are essential for healthcare professionals and patients to make the most of AI practices. Therefore, it is necessary to take into account user feedback in the development of AI systems and implement continuous improvement processes. *(Han et al., n.d.)*

These processes will both improve the quality of healthcare and reinforce patients' and experts' confidence in AI technologies. This collaborative approach not only enhances the effectiveness of AI tools but also fosters a culture of accountability and responsiveness within healthcare systems. *(Esmaeilzadeh, 2020)*

Again, getting high scores on achieving goals shows that AI has a strong capacity to deliver the expected results. In this regard, the integration of AI practices in health care will be possible not only with technological advances, but also with the consideration of ethical and legal frameworks. (Terry *et al.*, n.d.)

Therefore, it is critical to adopt a multidisciplinary approach for the effective adoption of AI in the healthcare sector and to ensure the involvement of all stakeholders.

Publication or AI applications address at a very high level the needs of health professionals and the expectations of patients. This will increase the efficiency of the overall health system, ensuring that health services are delivered more effectively and patient-oriented, increasing the efficiency of the overall health system. In this process, the creation of continuing education and

awareness programs to improve the quality of health care is also important. *(Karpov et al., 2023)*

Low scores reveal that the sources on which the information offered by AI is based are not clearly stated enough. Therefore, identifying reliable data sources and presenting this data transparently is vital to increasing health professionals' confidence in AI applications. *(A Narrative Review and Impacts on Trust for Data in the Healthcare Industry Using Artificial Intelligence, 2022)*

In addition, comprehensive training programs should be developed to ensure that health professionals have adequate knowledge of artificial intelligence systems.

A similar uncertainty is observed regarding the production times of information. This uncertainty could hinder the effective use of AI in healthcare and adversely affect decision-making processes. In this context, it is important to organize hands-on trainings for health professionals to understand how artificial intelligence systems work and how data is collected. *("Artificial Intelligence in Healthcare," 2023)*

Such training can improve the quality of patient care and improve the overall efficiency of health care by enabling healthcare professionals to use AI technologies more effectively.

The expert physician assessed that the information provided by AI is balanced and unbiased. This assessment will give healthcare professionals confidence in artificial intelligence systems, while also helping them to develop awareness about the potential limitations of these technologies. In our study, the point of access to additional resources is given a low score. This necessitates the development of strategies to address the lack of access to information and resources necessary for health professionals to integrate with artificial intelligence systems. *(Castagno & Khalifa, 2020)*

There has been a moderate assessment of addressing uncertainties. This indicates that health professionals need a better understanding of AI applications and the need for training programs to help them overcome these uncertainties. *(Kalthoff et al., 2022)*

From another perspective, the reliability and effectiveness of AI applications in assessing complex health problems such as dizziness can be questioned. Artificial intelligence systems, in particular, such as ChatGPT-4, may have some limitations in terms of the capacity to provide medical information. The responses of such systems may contain less reliable and incomplete information when compared to the experience and intuition of human physicians. In addition, the sources of information offered by AI are often unclear, making it difficult for health professionals to have confidence in this information. *(Goldberg & Prutkin, 2022)*

In addition, the integration of artificial intelligence systems into health care brings with it ethical and legal issues. Incorporating such systems into decision-making processes can ignore the individual needs of patients and reduce the importance of personalized care. While it is important to consider user feedback in the development of AI applications, it should also be questioned whether these processes are sufficiently transparent and reliable. In terms of data collection methodology, the use of open platforms such as Reddit can create problems with the accuracy and reliability of information. The information shared by users may be far from medical accuracy, which can negatively affect the quality of medical advice provided by artificial intelligence systems. In addition, the lack of adequate knowledge of AI systems by healthcare professionals can hinder the effective use of these systems and reduce the quality of patient care. As a result, the use of AI applications in healthcare needs to be addressed not only by technological advances, but also by raising awareness about the limitations and potential risks of these systems. The reliability, transparency and accuracy of the information offered by artificial intelligence systems are critical to increasing health professionals' confidence in these technologies. Therefore, the adoption of artificial intelligence in healthcare should be supported by a multidisciplinary approach and the involvement of all stakeholders.

CONCLUSION

As a result, the evaluation of CHATGPT-4's responses to dizziness complaints reveals a significant intersection between AI and healthcare. The findings suggest that although AI applications such as ChatGPT-4 can provide clear and relevant information tailored to the needs of healthcare professionals and patients, there are notable gaps in the transparency of information sources and the timing of the data provided. High scores on clarity of objectives and the ability to achieve them suggest that AI can effectively contribute to the dissemination of medical knowledge. However, lower scores on source clarity and additional support indicate the need for improvements in the reliability of AI-generated information. The integration of AI into healthcare systems has the potential to improve patient care and reduce the workload on healthcare professionals, but requires a multidisciplinary approach that includes ongoing training and feedback mechanisms to build trust and ensure ethical use of technology. Addressing the identified limitations will be crucial to maximizing the benefits of AI in medical contexts, ultimately leading to improved health care outcomes and a more efficient healthcare delivery system. In this *(Foufi et al., 2019)* context, the role of AI is becoming increasingly critical with the digitalization of healthcare.

Therefore, it is important that policymakers, technology developers and clinical experts work collaboratively to set standards and develop best practices so that AI can be effectively implemented in healthcare. This collaboration will facilitate the integration of AI in healthcare and deliver better care experiences by improving patient safety.

REFERENCES

1. Rana, A. Q., & Morren, J. A. (2013). Dizziness and Vertigo. https://doi.org/10.1007/978-1-4471-5191-3_3
2. Lasisi, A. O., & Gureje, O. (2010). Disability and quality of life among community elderly with dizziness: report from the Ibadan study of ageing. *Journal of Laryngology and Otology*. https://doi.org/10.1017/S0022215110000538
3. Javanmard, S. (n.d.). Revolutionizing Medical Practice The Impact of Artificial Intelligence AI on Healthcare.
4. Liu, J., Wang, C., & Liu, S. (n.d.). Utility of ChatGPT in Clinical Practice.
5. Terry, A., Kueper, J., Beleno, R., Brown, J., Cejic, S., Dang, J., Leger, D., Mckay, S., Meredith, L., & Pinto, A. (n.d.). Is primary health care ready for artificial intelligence What do primary health care stakeholders say.
6. Han, R., Acosta, J., Shakeri, Z., Ioannidis, J., Topol, E., & Rajpurkar, P. (n.d.). Randomised controlled trials evaluating artificial intelligence in clinical practice a scoping review.
7. Esmaeilzadeh, P. (2020). Use of AI-based tools for healthcare purposes: a survey study from consumers' perspectives. BMC Medical Informatics and Decision Making. https://doi.org/10.1186/S12911-020-01191-1
8. Karpov, O. E., Pitsik, E., Kurkin, S. A., Maksimenko, V. A., Shusharina, N., & Hramov, A. E. (2023). Analysis of Publication Activity and Research Trends in the Field of AI Medical Applications: Network Approach. *International Journal of Environmental Research and Public Health*. https://doi.org/10.3390/ijerph20075335
9. A narrative review and impacts on trust for data in the healthcare industry using artificial intelligence. (2022). https://doi.org/10.1049/pbhe040e_ch10
10. Artificial Intelligence in Healthcare. (2023). Advanced Technologies and Societal Change. https://doi.org/10.1007/978-981-99-3157-6_7
11. Castagno, S., & Khalifa, M. (2020). Perceptions of Artificial Intelligence Among Healthcare Staff: A Qualitative Survey Study. https://doi.org/10.3389/FRAI.2020.578983
12. Kalthoff, D., Prien, M., & Götz, N.-A. (2022, May 25). Development and Conception of a Learning Programme in Higher and Continuing Education on the Fundamentals, Applications and Perspectives of AI in Healthcare. *Medical Informatics Europe*. https://doi.org/10.3233/SHTI220584
13. Goldberg, S., & Prutkin, M. (2022). Overcoming AI applications challenges in health: Decision system DINAR2. https://doi.org/10.1016/b978-0-12-824521-7.00006-5
14. Foufi, V., Foufi, V., Timakum, T., Gaudet-Blavignac, C., Gaudet-Blavignac, C., Lovis, C., Lovis, C., & Song, M. (2019). Mining of Textual Health Information from Reddit: Analysis of Chronic Diseases with Extracted Entities and Their Relations. *Journal of Medical Internet Research*. https://doi.org/10.2196/12876