RESEARCH ARTICLE

Exploring the adoption and utilisation of ChatGPT in everyday academic practices among medical and dental students of Rawalpindi and Islamabad: a study on the impact and perceived effectiveness

Khadija Qamar[1], Saad Waseem Butt[2], Mehwish Abaid[3], Ahmed Rashid[4], Rafia Tahira[5], Aarij Sohail Iqbal[6]

Abstract

Objective: To assess the usage of chatbot generative pretrained transformer among medical and dental students in Pakistan.

Method: The cross-sectional study was conducted at the Army Medical College, National University of Medical Sciences, Islamabad, Pakistan, from June 15 to July 31, 2023, and comprised medical and dental students in public and private medical and dental colleges of Rawalpindi and Islamabad who had used the chatbot generative pretrained transformer programme at least once for their academics or had prior knowledge of it. Responses were obtained through a structured online survey. Data was analysed using SPSS 27.

Result: Of the 315 respondents, 181(57.5%) were females, 134(42.5%) were males, 260(82.5%) were from public institutions, 55(17.5%) were from private, 259(82.20%) were medical students and 56(17.80%) were dental students. Overall, 192(61%) respondents had used ChatGPT for academics prior to the study. For them, its utility was most for case-based learning, problem-based learning and for seeking clarification of topics. Fruitful acquisition of knowledge was reported by 270(85.7%) of the study participants and 229(72.8%) participants regarded ChatGPT as a valuable resource for improving academic performance. The mean perceived usefulness score was 17.93±5.08 and the mean perceived risk score was 8.46±2.44.

Conclusion: Chatbot generative pretrained transformer was found to improve academic performance of medical and dental students.

Key Words: Academic performance, Medical education, Problem-based learning, Research, Risk.

(JPMA 74: 2101; 2024) **DOI: https://doi.org/10.47391/JPMA.11053**

Introduction

Chatbot generative pretrained transformer (ChatGPT), based on GPT large language model (LLM), is a computer algorithm that processes the input from natural language and from variable sources, and predicts the words, continuing until a complete response to a query is obtained.[1] LLMs process large amounts of data and interconnect the words within the text. They predict the word in a sequence of words.[2] LLMs work in such a way that they use transformer models, and are trained to use enormous datasets, hence they are large. Due to this, they can recognise, translate, predict or generate text and other content.[3]

With its rising popularity, a number of researches have been carried out about its utilisation in the field of education, healthcare and research.[4] Pertaining to education, ChatGPT marks a new milestone in consolidating dispersed data from multiple sources into a unified platform, effectively saving time. Moreover, it proves beneficial in summarising extensive information into concise points, saving time spent on sifting through lengthy texts. It can help formulate topic-oriented multiple choice questions (MCQs) that can help students apply key concepts of their knowledge, and thus aid them in the preparation of assessments.[5]

However, research on the frequency of adoption, and the impact of ChatGPT on medical education has not been carried out extensively.[6] A study on ChatGPT conducted on medical professionals showed that 83% medical students had positive responses about its utility, whereas only 62% graduated physicians agreed on its usefulness.[7] Its potential to enhance knowledge acquisition, critical thinking and problem-solving skills is still under study. Along with frequency of use, another important factor is how students perceive the efficacy and value of ChatGPT as an educational tool.[8] This can provide valuable insights into its potential role and benefits in the learning process. Evaluation of its accuracy and reliability from the eyes of

[1,3]Department of Anatomy, Army Medical College, National University of Medical Sciences, [2,4-6]2nd Year MBBS Student, Army Medical College, National University of Medical Sciences, Rawalpindi, Pakistan.

Correspondence: Mehwish Abaid. **Email:** mehwish.abaid00@gmail.com

ORCID ID: 0000-0001-7624-1538

Submission complete: 12-10-2023 **First Revision received:** 09-01-2024

Acceptance: 28-09-2024 **Last Revision received:** 27-09-2024

medical students is another factor which needs further research.[4,5]

The current study was planned to develop a comprehensive understanding of the adoption and utilisation of ChatGPT in everyday academic practices by medical and dental students in an urban setting, and to examine the perceived usefulness of ChatGPT.

Subjects and Methods

The cross-sectional study was conducted at the Army Medical College, National University of Medical Sciences, Islamabad, Pakistan, from June 15 to July 31, 2023, and comprised medical and dental students in public and private medical and dental colleges of Rawalpindi and Islamabad.

After approval for the institutional ethics review committee, the minimum sample size of 307 was calculated using the World Health Organisation (WHO) calculator[9] with confidence interval 95%, margin of error 5.6% and population proportion 50.5%.[10] The sample was raised using non-probability convenience sampling technique. Random sampling was not used because most students of senior years were busy with their clinical rotations. The sample, as such, was raised from among the available students. Those included were students of Bachelor of Medicine and Bachelor of Surgery (MBBS) from first year till the final year of their 5-year course, and students of Bachelor in Dental Surgery (BDS) from the first year till the final year of their 4-year course. Those not willing to participate were excluded. Within the enrolled students, those who had used ChatGPT at least once for their academics before the study were included to check the impact of ChatGPT in their everyday academics, while those who had not used ChatGPT were also included to check the prevalence of ChatGPT usage.

After taking informed consent from the subjects, data was collected using a validated multi-structured questionnaire[10] which is provided that was sent to multiple public and private medical and dental colleges of Rawalpindi and Islamabad using the contacts of authors through WhatsApp groups. The limitations of the study were taken into consideration while constructing the questionnaire which is an indicator of the questionnaire validity. It was also printed and given to students for answering. Two different ways of data collection were employed; through printed questionnaire, and through Whatsapp groups.

The questionnaire consisted of three sections. The first section consisted of the student's demographics, followed by a closed-ended question to explore if they had heard about ChatGPT before the start of the study. If the participant had not used ChatGPT before, they were asked to fill the questionnaire regardless about how they would perceive the use of An Artificial Intelligence (AI) LLM model for educational purposes.

The next section was based on the technology acceptance model (TAM) framework[11] consisting of 12 items. Each question was evaluated on a 5-point Likert scale, with the response scores ranging 1-5. The last section consisted of 4 structured questions based upon how the participant used ChatGPT in their everyday academics. All the questions in this section were closed-ended.

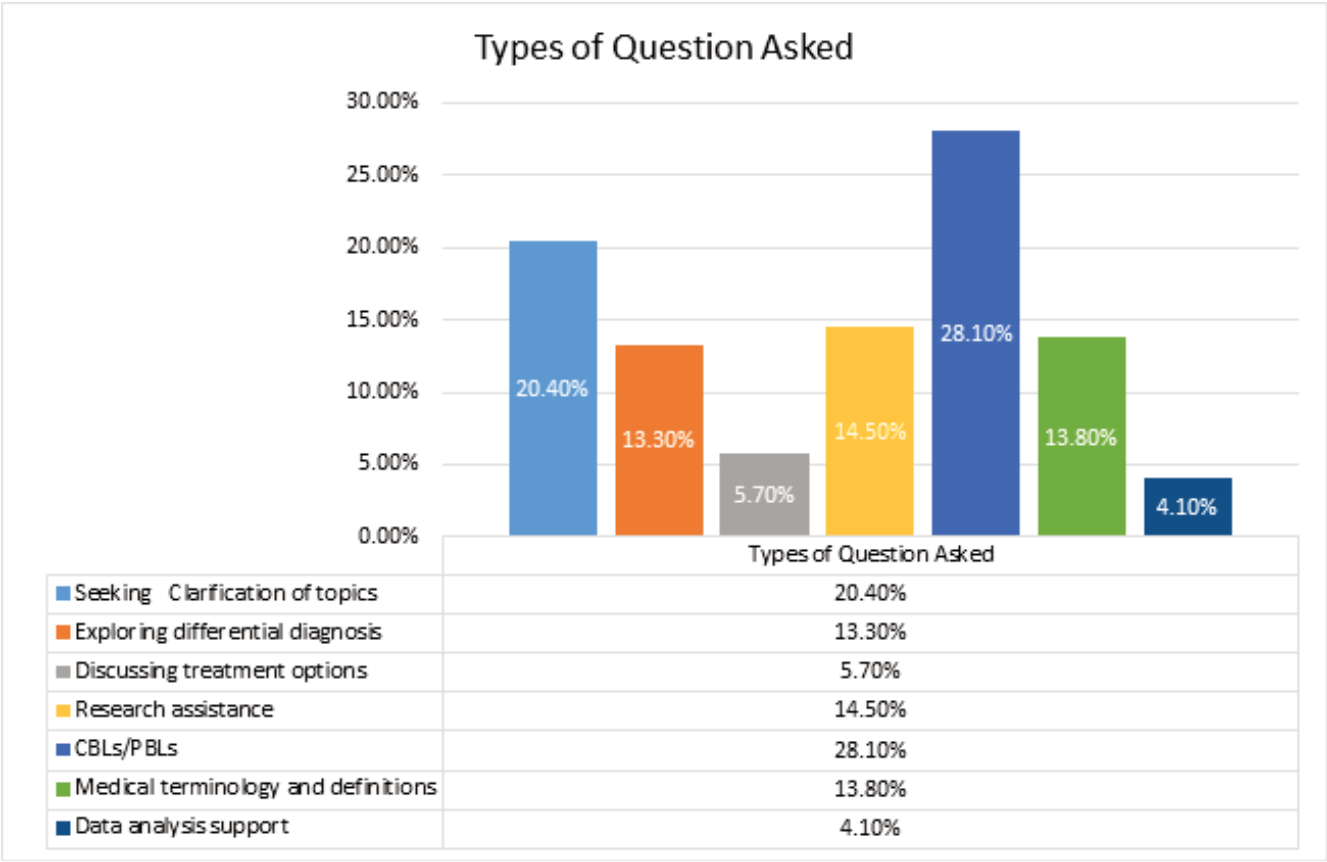
Data was compared for usage and non-usage, for usage among preclinical and clinical years, and for usage among students of public and private institutions. The operational definition of perceived usefulness was the respondents' belief in the technology's ability to enhance their productivity in performing a specific activity.[12] The perceived risk referred to the risks associated with the use of ChatGPT, including the potential for incorrect information, confidentiality issues and ethical concerns.[13] It was calculated by analysing the results of three questions which were asked in the questionnaire regarding the adverse effects of ChatGPT. They were based on Likert scale with 5 options (5 marks for strongly agree and 1 mark for strongly disagree). The mean was taken of the results of all three questions which was the final value for perceived risk. The perceived usefulness scale had a potential range of 5-25, with higher scores indicating a greater perceived usefulness of ChatGPT, and a score of 15 representing a neutral attitude.[10]

Data was analysed using SPSS 27. For quantitative data, mean ± standard deviation (SD) values were calculated, while frequencies and percentages were calculated for qualitative data. For the comparison of perceived usefulness and perceived risk with respect to gender, university type and field of study, independent T-test was used. $P<0.05$ was considered significant.

Results

Of the 315 respondents, 181(57.5%) were females, 134(42.5%) were males, 260(82.5%) were from public institutions, 55(17.5%) were from private, 259(82.20%) were medical students and 56(17.80%) were dental students. Overall, 192(61%) respondents had used ChatGPT for academics prior to the study.

Total responses from preclinical years were 227(72.1%) and from clinical years 88(27.9%). Year-wise, there were 135(42.9%) responses from the 1st year, 92(29.2%) from



	Types of Question Asked
Seeking Clarification of topics	20.40%
Exploring differential diagnosis	13.30%
Discussing treatment options	5.70%
Research assistance	14.50%
CBLs/PBLs	28.10%
Medical terminology and definitions	13.80%
Data analysis support	4.10%

Figure-1: Types of questions included in the study questionnaire about the usage of chatbot generative pretrained transformer (ChatGPT) in academics in medical and dental fields.

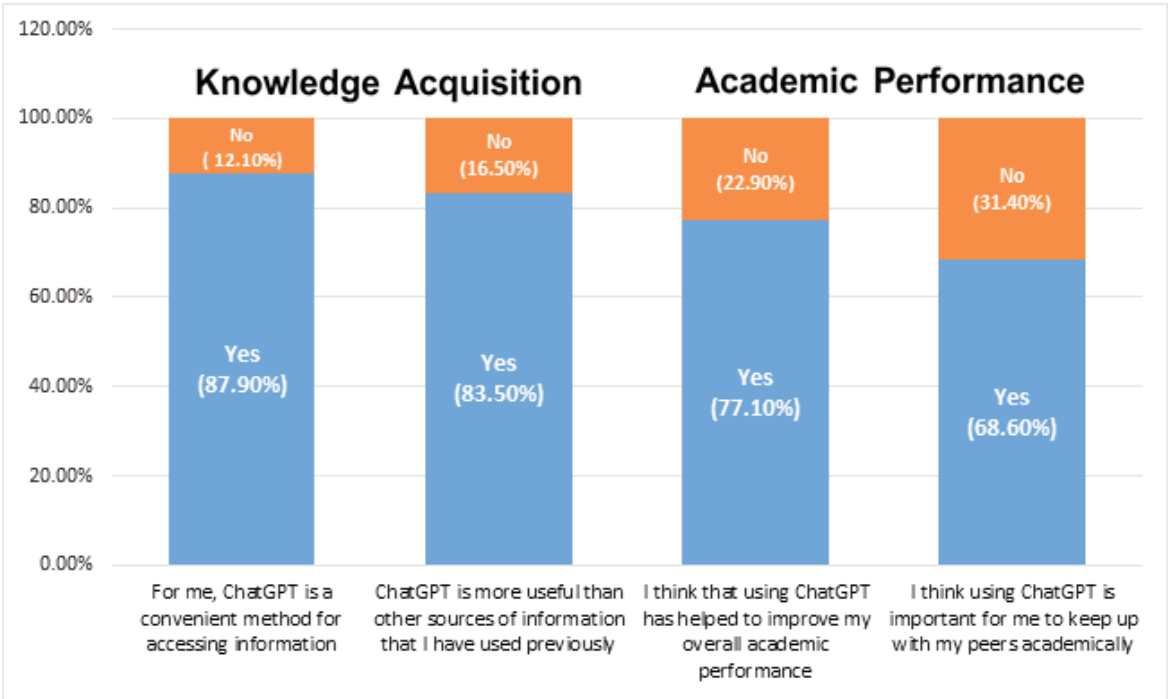


Figure-2: The response of students regarding knowledge acquisition with the use of chatbot generative pretrained transformer (ChatGPT) along with its comparison with other tools.

the 2nd year, 66(21%) from the 3rd year, 14(4.4%) from the 4th year and 8(2.5%) from the final year.

For students, the utility of ChatGPT was most for case based learning/problem based learning (CBL/PBL) and for seeking clarification of topics (Figure 1).

Fruitful acquisition of knowledge via ChatGPT was reported by 270 (85.7%) students, while 229(72.8%) students regarded it as a valuable resource for improving academic performance (Figure 2).

The mean perceived usefulness score was 17.93±5.08, indicating high perceived usefulness of ChatGPT.

There was a statistically significant difference with respect to gender and field of study, but the difference was not significant based on the type of institution (Table 1).

Table-1: Perceived usefulness and perceived risk of ChatGPT on gender, university type and field of study.

	Perceived Usefulness		Perceived Risk	
	Mean ± SD	p-value	Mean ± SD	p-value
Gender				
Male	16.25±3.97	<0.001*	8.00±2.42	0.116
Female	13.97±3.75		8.81±2.41	
University Type				
Public	15.13±3.89	0.308	8.51±2.39	0.371
Private	14.05±4.42		8.23±2.68	
Field of Study				
MBBS	15.34±3.93	<0.001*	8.35±2.36	0.002*
BDS	13.08±3.83		9.00±2.73	

*= statistically significant, ChatGPT: Chatbot generative pretrained transformer, SD: Standard deviation, MBBS: Bachelor of medicine, bachelor of surgery, BDS: Bachelor of dental surgery.

The mean perceived risk score was 8.46±2.44, indicating borderline perceived risk of ChatGPT. There was a significant difference with respect to the field of study (p=0.002), while no significant difference was seen based on gender (p=0.116) and the type of institution (p=0.371).

Discussion

ChatGPT is a tool that is now widely used in academic field. The current study focussed on its utility and advantages in undergraduate medical and dental academics. The study focussed on its usage among the students, and covered all aspects of studies starting from the beginning till the end of the medical school, which included not just the acquisition of knowledge, but also patient diagnosis and management. Along with the curriculum aspect, the research aspect was also covered, which included assistance in conducting the research, data analysis and interpretation. All these elements were also described by Garg et al.[14] who explained that ChatGPT helped in reviewing the manuscripts and editing them for any potential errors.

A study on specialists about their usage of ChatGPT in clinical scenarios elaborated that it was a quick and effective way to seek information and assistance regarding clinical case handling and management.[15]

Regarding making the differential diagnosis in a clinical setup, ChatGPT helps in making the top differentials in an emergency clinical setup when the chief complaints are mentioned to it. The condition that is to be followed properly is that complete information should be provided, and any previous medical conditions, if present, should be addressed.[16] Another study showed similar results, but also explained that the differential diagnosis made by physicians is far superior and more accurate compared to ChatGPT because of proper history-taking and physical examination. But there was >50% overlapping between the diagnosis the physicians made and the one suggested by ChatGPT.[17]

In medical research, ChatGPT provides valuable assistance in writing of papers, particularly the part of literature review, but a high percentage of plagiarism is observed in such cases. So, they must be verified by a human researcher. Hence, it supports the researcher in the early stages of medical writing. ChatGPT helps the users to minimise the systematic errors which they may be committing while conducting research and doing medical writing.[18]

In the study, majority of students considered it a useful knowledge acquisition tool, and said that it was much better than the resources that they had previously used. A study compared ChatGPT and Google search for the acquisition of knowledge, and ChatGPT turned out to be a better tool for the purpose as it integrated the relevant data from multiple sources.[19] It not only provided information on the required topic and field, but also assisted them in their tasks, assessed them and then provided positive or negative feedback. It could also make study plans according to the individual needs of a student when mentioned correctly in the chatbox which is a function not possible in the traditional classrooms because of shortage of time and resources.[20]

While most of the students in the current study deemed it a potential tool of knowledge acquisition, 12.10% did not. A study performed on the specialists showed that at times, the answers provided by ChatGPT were incorrect. This largely depended on the questions being asked in an improper way and without complete information. But even when the questions provided adequate information,

incorrect answers were sometimes noted by physicians. So, the answers must be counter-checked with trusted resources.[15] Another study assessed biomedical knowledge graphs with ChatGPT in which the graphs were found to be more reliable in providing knowledge and in the management of information.[21]

In the current study, majority of the students credited ChatGPT to be one of the reasons that they could perform better academically. This can be attributed to the reason that this technology improves students' reasoning, and engages them more in active learning.[22] ChatGPT also performs 'Personalised learning' for an individual student. In it, learning is aimed at catering to the individual needs of the student, and individual assessment is performed through which students can know about their shortcomings. Hence, it can be a major factor in improving academic performance as the study resolves around individual interests, abilities and needs.[23]

For the students who do not find a positive impact of ChatGPT on their academic performance, ChatGPT is only a tool. The role of teachers in the classroom cannot be entirely replaced by an AI tool. There should be integration of the role of teachers with the latest technology to achieve maximum efficiency and efficacy of both.[24]

Perceived usefulness was significant in both genders, and in both medical and dental fields of study in the current study, whereas it was non-significant in the type of institution. Perceived usefulness is one of the most important factors due to which an individual decides to opt for a particular tool, and its ease of usage determines its continued use.[25]

Regarding perceived risk, if the security is weak at any point, it can leak any organisation's information. Moreover, the data, which is an intellectual property, can also be provided by ChatGPT through some other sources, resulting in a major privacy breach.[26] Thus, the students in the current study stressed that information provided must be verified through authentic sources to avoid any kind of inaccurate and false knowledge. The need for expert verification is always there and cannot be achieved by ChatGPT.[27]

The current study has limitations as the findings were based on the self-perception of the participants which may change over time. As such, the findings may not be generalisable to the larger population of medical and dental students in other regions. Further research is needed to explore and address the perceptions and risks associated with ChatGPT usage in medical education.

Conclusion

The majority of the participants used ChatGPT for CBL/PBL and for seeking clarification on various topics. Overall, ChatGPT was regarded as a valuable tool for enhancing academic performance. Variations were observed in the perceived usefulness based on gender and field of study, while the perceived risk showed differences based on the field of study.

Disclaimer: None.

Conflict of Interest: None.

Source of Funding: None.

References

1. Mearian L. What are LLMs, and how are they used in generative AI? Computer world United states. 2023. [Online] [Cited 2023 July 11]. Available from: URL: https://www.computerworld.com/article/3697649/what-are-large-language-models-and-how-are-they-used-in-generative-ai.html
2. Ruby M. How ChatGPT Works: The Model Behind The Bot. Towards Data Science. [Online] [Cited 2023 June 20]. Available from: URL: https://towardsdatascience.com/how-chatgpt-works-the-models-behind-the-bot-1ce5fca96286
3. Company ETSA. What is a large language model (LLM)? [Online] [Cited 2023 July 13].Available from: URL: https://www.elastic.co/what-is/large-language-models.
4. Breeding T, Martinez B, Patel H, Nasef H, Arif H, Nakayama D, et al. The Utilization of ChatGPT in Reshaping Future Medical Education and Learning Perspectives: A Curse or a Blessing? Am Surg. 2024; 90:560-6. doi: 10.1177/00031348231180950.
5. Hosseini M, Gao CA, Liebovitz DM, Carvalho AM, Ahmad FS, Luo Y, et al. An exploratory survey about using ChatGPT in education, healthcare, and research. PLoS One. 2023; 18:e0292216.doi: 10.1371/journal.pone.0292216.
6. Vaishya R, Misra A, Vaish A, ChatGPT: Is this version good for healthcare and research? Diabetes & Metabolic Syndrome: Clin Res Rev. 2023; 17:102744. doi: 10.1016/j.dsx.2023.102744.
7. Tangadulrat P, Sono S, Tangtrakulwanich B. Using ChatGPT for Clinical Practice and Medical Education: Cross-Sectional Survey of Medical Students' and Physicians' Perceptions. JMIR Med Educ. 2023; 9:e50658.doi: 10.2196/50658.
8. Lo CK. What is the impact of ChatGPT on education? A rapid review of the literature. Educ Sci. 2023; 13: 410. doi:10.3390/educsci13040410
9. Organization WHO. WHO sample size calculator. World Health Organization. london:Departement of Mathematics Education, 2023; pp. 16100-16106
10. Sallam M, Salim N, Barakat M, Al-Mahzoum K, Al-Tammemi AB, Malaeb D, et al. Validation of a technology acceptance model-based scale TAME-ChatGPT on health students attitudes and usage of ChatGPT in Jordan.[Online] [Cited April 2023 17].Avaialble from: URL: DOI:10.2196/preprints.48254
11. Sallam M, Salim NA, Barakat M, Al-Mahzoum K, Ala'a B, Malaeb D, et al. Assessing health students' attitudes and usage of ChatGPT in Jordan: validation study. JMIR Med Educ. 2023; 9:e48254. doi: 10.2196/48254.
12. Yilmaz H, Maxutov S, Baitekov A, Balta N. Student Attitudes towards Chat GPT: A Technology Acceptance Model Survey. Int Educ Rev. 2023; 1:57-83. DOI:10.58693/ier.114

13. Hill C. Uncovering the risks of ChatGPT: a litigator's perspective. 2023. [Online] [Cited 2023 May 18]. Available from: URL: https://www.penningtonslaw.com/news-publications/latest-news/2023/uncovering-the-risks-of-chatgpt-a-litigator-s-perspective
14. Garg RK, Urs VL, Agrawal AA, Chaudhary SK, Paliwal V, Kar SK. Exploring the Role of Chat GPT in patient care (diagnosis and Treatment) and medical research: A Systematic Review. Health Promot Perspect. 2023; 13:183-91.doi: 10.34172/hpp.2023.22.
15. Johnson D, Goodman R, Patrinely J, Stone C, Zimmerman E, Donald R, et al. Assessing the accuracy and reliability of AI-generated medical responses: an evaluation of the Chat-GPT model. Res Sq. 2023; rs.3.rs-2566942.doi: 10.21203/rs.3.rs-2566942/v1.
16. Hirosawa T, Harada Y, Yokose M, Sakamoto T, Kawamura R, Shimizu T, et al. Diagnostic accuracy of differential-diagnosis lists generated by generative pretrained transformer 3 chatbot for clinical vignettes with common chief complaints: A pilot study. Int J Environ Res Public Health. 2023; 20:3378. doi: 10.3390/ijerph20043378.
17. Ten Berg H, van Bakel B, van de Wouw L, Jie KE, Schipper A, Jansen H, et al. ChatGPT and Generating a Differential Diagnosis Early in an Emergency Department Presentation. Ann Emerg Med. 2024; 83:83-6. doi: 10.1016/j.annemergmed.2023.08.003.
18. Ruksakulpiwat S, Kumar A, Ajibade A. Using ChatGPT in Medical Research: Current Status and Future Directions. J Multidiscip Healthc. 2023 Dec 31:1513-20. doi: 10.2147/JMDH.S413470.
19. Ayoub NF, Lee YJ, Grimm D, Divi V, Surgery N. Head-to-Head Comparison of ChatGPT Versus Google Search for Medical Knowledge Acquisition. Otolaryngol Head Neck Surg. 2024; 170:1484-91.doi: 10.1002/ohn.465
20. Khan RA, Jawaid M, Khan AR, Sajjad M. ChatGPT-Reshaping medical education and clinical management. Pak J Med Sci. 2023; 39:605-607.doi: 10.12669/pjms.39.2.7653.
21. Hou Y, Yeung J, Xu H, Su C, Wang F, Zhang R. From Answers to Insights: Unveiling the Strengths and Limitations of ChatGPT and Biomedical Knowledge Graphs. medRxiv. 2023; 6:23291208.doi: 10.1101/2023.06.09.23291208.
22. Mhlanga D. The Value of Open AI and Chat GPT for the Current Learning Environments and the Potential Future Uses.[Online] [Cited 2023 May 09]. Available from: URL: http://dx.doi.org/10.2139/ssrn.4439267
23. Fuchs K, Exploring the opportunities and challenges of NLP models in higher education: is Chat GPT a blessing or a curse? Front Educ. 2023; 8:1166682. doi.org/10.3389/feduc.2023.1166682
24. Ausat AM, Massang B, Efendi M, Nofirman N, Riady Y. Can chat GPT replace the role of the teacher in the classroom: A fundamental analysis. J Educ. 2023; 5:16100-6. doi.org/10.31004/joe.v5i4.2745
25. Liu G, Ma C. Measuring EFL learners' use of ChatGPT in informal digital learning of English based on the technology acceptance model. Innov Lang Learn Teach. 2023; 27:1-4. DOI:10.1080/17501229.2023.2240316
26. Cuomo J. Exploring the risks and alternatives of ChatGPT: Paving a path to trustworthy AI. IBM Blog. [Online] [Cited July 2023 27]. https://www.ibm.com/blog/exploring-the-risks-and-alternatives-of-chatgpt-paving-a-path-to-trustworthy-ai/.
27. Oviedo-Trespalacios O, Peden AE, Cole-Hunter T, Costantini A, Haghani M, Rod J, et al. The risks of using chatgpt to obtain common safety-related information and advice. Safety Sci. 2023; 167:106244. doi:10.1016/j.ssci.2023.106244

Authors' Contribution:
KQ: Concept, design, critical analysis and final approval.
SWB, AR, RT & ASI: Concept, design, data acquisition and analysis.
MA: Drafting, revision, final approval and agreement to be accountable for all aspects of the work.