OPEN ACCESS
Turkish Journal of Kinesiology

ORIGINAL RESEARCH

Does ChatGPT provide comprehensive and accurate information regarding the effects, types and programming of core exercises?

Erkan Erol, Halime Arıkan

Department of Physiotherapy and Rehabilitation, Faculty of Health Sciences, Tokat Gaziosmanpaşa University, Tokat, Türkiye.

Received:
July 16, 2024

Accepted:
September 25, 2024

Online Published:
September 30, 2024

Keywords:
Artificial intelligence, ChatGPT, exercise.

Abstract

The objective of this study is to assess the accuracy of ChatGPT's responses regarding core exercises. A total of 23 questions were asked to ChatGPT 3.5 about core exercises. Nine physiotherapists assessed the accuracy of the answers provided by ChatGPT for these questions using a 5-point Likert scale (5: strongly agree, 1: strongly disagree). The maximum possible score achievable through Likert scoring is 115, while the minimum score is 23. The answers of the artificial intelligence received an average of 3.93±0.46 (minimum: 3.48, maximum: 4.91) points. The lowest score obtained from the responses of ChatGPT was 3.22 ± 0.97 (question 21), whereas the highest score was 4.56 ± 0.53 (questions 12 and 18). Despite having some shortcomings, ChatGPT has generally provided satisfactory answers regarding core exercises. Artificial intelligence holds the potential to offer counselling to individuals, enhance their health outcomes, and support the work of professionals across the field of exercise science, including health and sports specialists.

Introduction

Artificial intelligence technology has made great progress recently. Artificial intelligence chatbots are computer programs equipped with the capability to comprehend human language and engage in detailed conversations with users by providing appropriate responses (Massey et al., 2023). ChatGPT is an artificial intelligence chatbot developed by OpenAI (Cotton et al., 2024). It is based on an architecture called GPT (Generative Pre-trained Transformer). This model is pre-trained with large amounts of text data and then made available to perform a variety of language understanding and natural language processing tasks (Daniel et al., 2020).

While Google provides users with websites containing relevant answers to their current questions, the information accessible through ChatGPT can indeed be found using the Google search engine. However, the distinction lies in the approach: Google presents users with a variety of options from different sources related to the searched topic, whereas ChatGPT generates responses based on educated guesses and offers users a single answer promptly (Olçar et al., 2023).

The core muscles, including the abdominals, paraspinals, gluteals, diaphragm, pelvic floor, and hip girdle muscles (Akuthota et al., 2008). These muscles are integral for stabilizing the spine, pelvis, and kinetic chain during functional movements. Additionally, core muscles provide protection to the spine against physical stress (Crisco et al., 1992).

Core stability is essential for preventing spinal instability and maintaining postural control. Studies have demonstrated that exercises aimed at enhancing core stability can effectively alleviate various health issues, particularly low back pain (Coulombe et al., 2017; Li et al., 2021; Yu & Park, 2013). Furthermore, core exercises are recommended to enhance athletic performance and prevent sports-related injuries (McGill, 2010; Reed et al., 2012). The core region serves as a foundation of "proximal stability for distal mobility" (Kibler et al., 2006).

Because of their numerous health benefits, core exercises are often incorporated into treatment programs by physiotherapists. As ChatGPT gains popularity, studies in the literature are emerging to explore its responses to various health conditions. A recent study assessed the reliability of ChatGPT as a

✉ E. Erol, e-mail: erkanerol@gmail.com
To Cite: Erol, E., & Arıkan, H. (2024). Does ChatGPT provide comprehensive and accurate information regarding the effects, types and programming of core exercises?*Turk J Kinesiol*, 10(3), 178-182. **DOI:** 10.31459/turkjkin.1516614

source of nutrition advice, indicating its potential for generating diet programs but recommending against its independent use without oversight from a nutritionist (Dergaa et al., 2024). Another study evaluated its suitability for prescribing resistance training, similarly suggesting promise but emphasizing the need for supervision by a specialist (Washif et al., 2024). However, the current number of studies is not yet sufficient to thoroughly evaluate the accuracy of artificial intelligence's responses to health-related inquiries. There is currently no study in the literature that evaluates ChatGPT's responses to inquiries about core exercises. The objective of this study is to assess the accuracy of ChatGPT's responses regarding core exercises.

The hypothesis of our study was to determine whether ChatGPT provides sufficient and accurate information regarding the effects, types, and programming of core exercises.

Methods

This study is a descriptive research project that involves the evaluation of information related to core stabilization by experienced physiotherapists in the field, using responses generated by the ChatGPT 3.5 version in Turkish. Since the study did not involve patients and solely compared computer-generated data, ethics committee approval was not deemed necessary.

The researchers identified eleven questions pertaining to core stabilization, with plans to commence the study in March 2024. An committee, comprising experienced physiotherapists, was established. Nine qualified physiotherapists, with experience ranging from six to eleven years, reviewed the eleven initial questions developed by the researchers. Given the importance of core stabilization across all areas of physiotherapy, these physiotherapists specialized in various rehabilitation fields, including orthopedic rehabilitation, sports rehabilitation, spine rehabilitation, pediatric rehabilitation, neurological rehabilitation, oncological rehabilitation, cardiopulmonary rehabilitation and geriatric rehabilitation. The team thoroughly examined the questions, making corrections and additions as needed. These questions were then presented to the ChatGPT version 3.5, which supports the Turkish language. Subsequently, the committee assessed the responses provided by ChatGPT version 3.5 for adequacy.

Furthermore, based on the recommendations of the committee, an additional twelve questions were posed to ChatGPT version 3.5. The committee also evaluated the responses of ChatGPT version 3.5 to these twelve questions. Consequently, a ChatGPT answer sheet, spanning a total of seventeen pages, was created.

The significance of the core region and exercises, the target audience benefiting from core exercises, considerations during core exercises, advantages of core exercises, potential drawbacks of core exercises, recommendations for a core exercise program, suggested duration and frequency of core exercises, methods for increasing intensity in core exercises, and questions regarding the evaluation of achievements through core exercises were formulated by the researcher through an extensive review of the literature (Hlaing et al., 2021; Sannasi et al., 2023; Vera-Garcia et al., 2020; Zemková & Zapletalová, 2021, 2022). In the second stage, the following questions recommended by the committee were asked: 1) The role of the physiotherapist in core exercises, 2) Decision-making on exercise intensity in the core exercise program, 3) Core exercises that can be preferred for different age groups, 4) Core exercises that can be preferred for different disease groups, 5) Equipment that can be used in core exercises, 6) Considerations for equipment usage in core exercises, 7) Determinants of equipment selection in core exercises, 8) The prioritization of core muscles, 9) Alternative exercises to core exercises, 10) Situations when core exercises are not suitable, 11) Stages of core stabilization training, and 12) Evaluation of gains achieved with core exercises using evidence-based methods (Table 1).

In the study, a booklet comprising twenty three questions and their corresponding answers, as determined by the researcher, was utilized. Nine physiotherapists assessed the accuracy of the answers provided by ChatGPT 3.5 for these questions using a 5-point Likert scale (5: strongly agree, 1: strongly disagree).

Physiotherapists evaluated the accuracy of ChatGPT 3.5's responses to a total of 23 questions and assessed its consistency with the committee opinions by scoring ChatGPT 3.5's answers. The maximum possible score achievable through Likert scoring is 115, while the minimum score is 23.

Data Analyses

SPSS 22 program was used for statistical analysis. The scores were reported as mean, standard deviation, minimum, and maximum.

Results

The answers of the artificial intelligence received an average of 3.93±0.46 (minimum: 3.48, maximum: 4.91) points. The lowest score obtained from the responses of ChatGPT was 3.22 ± 0.97 (question 21), whereas the highest score was 4.56 ± 0.53 (questions 12 and 18). The committee's ratings of the artificial intelligence's answers are shown in Table 2.

Table 1

Questions asked to ChatGPT 3.5 about core exercises.

1. What is the importance of the core area in the human body?
2. What is the importance of core exercises?
3. Who benefits from core exercises?
4. What should I pay attention to when doing core exercises?
5. What are the benefits of core exercises?
6. Can core exercises have adverse effects?
7. Do you recommend a core exercise program?
8. How long should I continue the core exercise program?
9. How often should I do core exercises?
10. How can I increase the intensity of core exercises?
11. With which tests can I evaluate the gains I have achieved with core exercises?
12. What is the role of the physiotherapist in core exercises?
13. How is exercise intensity decided when creating a core exercise program?
14. Which core exercises can be chosen for different age groups?
15. Which core exercises can be chosen according to different diseases?
16. What equipment should be used in core exercises?
17. What should be taken into consideration when using equipment in core exercises?
18. Which equipment should I choose for core exercises and on what basis?
19. Which core muscles are more important?
20. What exercises can be given as alternatives to core exercises?
21. In what situations are core exercises not suitable?
22. What are the stages of core stabilization training?
23. By what evidence-based methods can the gains obtained with core exercises be evaluated?

The initial 11 questions were formulated by the authors, while the subsequent 12 questions were developed in accordance with the committee's recommendations.

Table 2

Score distribution of the committee for ChatGPT 3.5's answers.

Questions	Minimum	Maximum	Mean ± SD	Questions	Minimum	Maximum	Mean ± SD
1	3	5	3.67 ± 0.87	13	3	5	3.67 ± 0.71
2	3	5	4.11 ± 0.78	14	4	5	4.11 ± 0.33
3	3	5	3.78 ± 0.97	15	3	5	3.67 ± 0.71
4	3	5	3.89 ± 0.93	16	4	5	4.44 ± 0.53
5	3	5	4.00 ± 0.87	17	3	5	4.33 ± 0.71
6	2	5	3.67 ± 1.00	18	4	5	4.56 ± 0.53
7	2	5	3.33 ± 1.12	19	3	5	4.22 ± 0.67
8	3	5	3.89 ± 0.78	20	3	4	3.67 ± 0.50
9	3	5	4.11 ± 0.60	21	2	5	3.22 ± 0.97
10	3	5	3.78 ± 0.83	22	3	5	3.89 ± 0.60
11	3	5	3.67 ± 0.87	23	3	5	4.22 ± 0.67
12	4	5	4.56 ± 0.53	Average	3.48	4.91	3.93 ± 0.46

SD: Standard deviation

Discussion

Artificial intelligence's guidance on health and exercise-related issues is one of the exciting events today. ChatGPT has the potential to contribute to health improvement by providing accurate guidance on core exercises. It can answer questions about core exercises, including considerations, benefits, potential adverse effects, equipment that can be used, types of exercises, exercise programming, and so on. Artificial intelligence holds the potential to offer counseling to individuals, enhance their health outcomes, and support the work of professionals across the field of exercise science, including health and sports specialists.

The results of committee opinions have indicated that ChatGPT generally delivers satisfactory answers to inquiries about core exercises. ChatGPT's responses to 23 questions received an average rating of 3.93±0.46 points from the committee. The highest scores were obtained for responses concerning the role of the physiotherapist (Q12: 4.56±0.53) and equipment selection (Q16: 4.44±0.53 and Q18: 4.56±0.53). The lowest scores were obtained from responses related to recommending exercise programs (Q7: 3.33±1.12) and situations where core exercises are not appropriate (Q21: 3.22±0.97). In their study, Olçar et al. examined ChatGPT's responses to hip prosthesis problems and reported that while ChatGPT's answers were generally correct, they expressed concern about its responses concerning personalized treatment and care. They emphasized that artificial intelligence should be regarded as a instrument to assist health professionals rather than a instrument to replace them (Dergaa et al., 2024; Olçar et al., 2023; Washif et al., 2024). In the current study, questions for which ChatGPT received high scores were those that could be answered in more general terms, whereas questions for which it received low scores were those requiring a more specific and individualized approach. Guidance from a health professional is crucial in individual exercise approaches. It was believed that artificial intelligence was not yet sufficiently competent for personalized exercise programs.

While ChatGPT may recommend core exercises, a notable limitation is its inability to monitor an individual's responses to the exercises and offer feedback accordingly. Variables such as exercise frequency, intensity, type, and number of repetitions should ideally be adjusted based on the individual's psychological responses. The inability of artificial intelligence to measure these factors and provide personalized feedback may hinder the potential for improving health outcomes (Borresen & Ian Lambert, 2009). Furthermore, artificial intelligence lacks the capability for patient evaluation, including assessing an individual's physical capacity, exercise history, and health status. Tailoring core exercise recommendations based on these factors would likely yield more effective results. A healthcare professional who can evaluate the individual and observe their responses to exercises will provide more personalized and safe exercise guidance (Dergaa et al., 2024).

Our questions were primarily open-ended questions based on interpretation, rather than questions with a single correct answer. While artificial intelligence provided more accurate answers to questions based on general knowledge, it received lower scores on questions that necessitated interpretation-based and individually specific answers. It was believed that artificial intelligence alone was far from being able to supervise individuals in performing core exercises, but it could provide satisfactory answers to guide them. Further studies are needed to expand the use of artificial intelligence in the field of health.

The exclusive use of the ChatGPT 3.5 version in the current study may be regarded as a limitation. It is possible that posing the same questions to the ChatGPT 4.0 version could yield different responses. Future research could consider simultaneously querying both versions to enhance comparative analysis.

Conclusion

Despite having some shortcomings, ChatGPT has generally provided satisfactory answers regarding core exercises. Artificial intelligence has the potential to counsel individuals and assist healthcare professionals in enhancing health. However, while ChatGPT demonstrates significant potential in providing exercise recommendations related to core stabilization, it is considered inadequate on its own due to its inability to monitor an individual's responses to exercises, provide feedback, conduct patient assessments—including evaluations of physical capacity, exercise history, and health status—and its lack of capability to visually assess individuals.

Authors' Contribution

Study Design: EE, HA; Data Collection: EE, HA; Statistical Analysis: EE; Manuscript Preparation: EE, HA.

Ethical Approval

Since the study did not involve patients and solely compared computer-generated data with expert information, ethics committee approval was not deemed necessary.

Funding

The authors declare that the study received no funding.

Conflict of Interest

The authors hereby declare that there was no conflict of interest in conducting this research.

References

Akuthota, V., Ferreiro, A., Moore, T., & Fredericson, M. (2008). Core stability exercise principles. *Curr Sports Med Rep*, 7(1), 39–44.

Borresen, J., & Ian Lambert, M. (2009). The quantification of training load, the training response and the effect on performance. *Sport Med,* 39(9), 779–795.

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innov Educ Teach Int,* 61(2), 228–239.

Coulombe, B. J., Games, K. E., Neil, E. R., & Eberman, L. E. (2017). Core Stability Exercise Versus General Exercise for Chronic Low Back Pain. *J Athl Train,* 52(1), 71–72.

Crisco, J. J., Panjabi, M. M., Yamamoto, I., & Oxland, T. R. (1992). Euler stability of the human ligamentous lumbar spine. Part II: Experiment. *Clin Biomech,* 7(1), 27–32.

Daniel, G., Cabot, J., Deruelle, L., & Derras, M. (2020). Xatkit: A Multimodal Low-Code Chatbot Development Framework. *IEEE Access,* 8, 15332–15346.

Dergaa, I., Ben Saad, H., El Omri, A., Glenn, J., Clark, C., Washif, J., Guelmami, N., Hammouda, O., Al-Horani, R., Reynoso-Sánchez, L., Romdhani, M., Paineiras-Domingos, L., Vancini, R., Taheri, M., Mataruna-Dos-Santos, L., Trabelsi, K., Chtourou, H., Zghibi, M., Eken, Ö., … Chamari, K. (2024). Using artificial intelligence for exercise prescription in personalised health promotion: A critical evaluation of OpenAI's GPT-4 model. *Biol Sport,* 41(2), 221–241.

Dergaa, I., Ben Saad, H., Ghouili, H., M Glenn, J., El Omri, A., Slim, I., Hasni, Y., Taheri, M., Ben Aissa, M., Guelmami, N., Al-Horani, R., Washif, J. A., Shoib, S., Mohammed Alyasiri, O., Jose Mataruna-Dos-Santos, L., Ferreira Alves, R., Ibrahim Ceylan, H., Swed, S., Alshahrani, N. Z., … Chamari, K. (2024). Evaluating the Applicability and Appropriateness of ChatGPT as a Source for Tailored Nutrition Advice: A Multi-Scenario Study. *New Asian J Med,* 2(1), 1–16.

Hlaing, S. S., Puntumetakul, R., Wanpen, S., & Saiklang, P. (2021). Updates on core stabilization exercise and strengthening exercise: A review article. *Asia-Pacific J Sci Technol,* 26(4), APST–26.

Kibler, W. Ben, Press, J., & Sciascia, A. (2006). The role of core stability in athletic function. *Sport Med,* 36(3), 189–198.

Li, X., Shen, J., Liang, J., Zhou, X., Yang, Y., Wang, D., Wang, S., Wang, L., Wang, H., & Du, Q. (2021). Effect of core-based exercise in people with scoliosis: A systematic review and meta-analysis. *Clin Rehabil,* 35(5), 669–680.

Massey, P. A., Montgomery, C., & Zhang, A. S. (2023). Comparison of ChatGPT–3.5, ChatGPT-4, and Orthopaedic Resident Performance on Orthopaedic Assessment Examinations. *J Am Acad Orthop Surg,* 31(23), 1173.

McGill, S. (2010). Core Training: Evidence Translating to Better Performance and Injury Prevention. *Strength Cond J,* 32(3), 33–46.

Olçar, H. A., Akbuğa, G. A., Ulutürk, S., Doğan, Ş. A., Korkmaz, M., & Korkmaz, İ. N. (2023). Evaluation of the ChatGPT-4 Artificial Intelligence Response to the Problems of Hip Replacement. *Bozok Med J,* 13(4), 118–122.

Reed, C. A., Ford, K. R., Myer, G. D., & Hewett, T. E. (2012). The Effects of Isolated and Integrated 'Core Stability' Training on Athletic Performance Measures. *Sport Med,* 42(8), 697–706.

Sannasi, R., Dakshinamurthy, A., Dommerholt, J., Desai, V., Kumar, A., & Sugavanam, T. (2023). Diaphragm and core stabilization exercises in low back pain: A narrative review. *J Bodyw Mov Ther,* 36, 221–227.

Vera-Garcia, F. J., Irles-Vidal, B., Prat-Luri, A., García-Vaquero, M. P., Barbado, D., & Juan-Recio, C. (2020). Progressions of core stabilization exercises based on postural control challenge assessment. *Eur J Appl Physiol,* 120, 567–577.

Washif, J., Pagaduan, J., James, C., Dergaa, I., & Beaven, C. (2024). Artificial intelligence in sport: Exploring the potential of using ChatGPT in resistance training prescription. *Biol Sport,* 41(2), 209–220.

Yu, S. H., & Park, S. D. (2013). The effects of core stability strength exercise on muscle activity and trunk impairment scale in stroke patients. *J Exerc Rehabil,* 9(3), 362–367.

Zemková, E., & Zapletalová, L. (2021). Back problems: Pros and cons of core strengthening exercises as a part of athlete training. *Int J Environ Res Public Health,* 18(10), 5400.

Zemková, E., & Zapletalová, L. (2022). The role of neuromuscular control of postural and core stability in functional movement and athlete performance. *Front Physiol,* 13, 796097.