The excellence Journal in Sciences Techniques Of Physical Activity And Sports
Volume: -11 N°01 -2026 P:249 -263
ISSN: 2507-7201 EISSN: 2602-6899 NDL: 787-2016
ASJP: https://www.asjp.cerist.dz/en/PresentationRevue/208

The Impact of a ChatGPT-Generated Training Program on Muscular Strength and Body Fat Percentage in Women

AMEUR CHAFA*[1], KETTAB KHALED[2]

[1] Institutes Science and Techniques of physical and Sports Activities, university of batna2 (ALGERIA), ameur.chafa@univ-batna2.dz

[2] Section Science and Techniques of physical and Sports Activities University Center Nour El Bachir El Bayadh, STAPS, (Algeria)., k.kettab@cu-elbayadh.dz

Received: 08/01 /2026 ***Accepted***: 04/06 /2026 ***Published***: 06/06 /2026

Abstract:

This study examined the effect of a ChatGPT-based training program on physical strength and body fat percentage in women. A randomized experimental design was applied to 20 female participants from an aerobics club in Batna Province, divided into experimental and control groups (n=10 each). The experimental group followed a six-week AI-generated program, while the control group maintained regular training. Body fat percentage and physical strength were assessed before and after the intervention. Results showed no statistically significant differences within or between groups ($p > 0.05$). Minor changes in the experimental group were not significant, suggesting limited effectiveness of short-term AI-generated training programs.

Key words : **Artificial Intelligence; ChatGPT; Training Program.**

الملخص:

هدفت هذه الدراسة إلى فحص أثر برنامج تدريبي قائم على CHATGPT على القوة البدنية ونسبة الدهون في الجسم لدى النساء. تم تطبيق المنهج التجريبي العشوائي على عينة مكونة من 20 مشاركة من نادٍ للأيروبيك بولاية باتنة، تم تقسيمهن إلى مجموعتين: تجريبية وضابطة (ن=10 لكل مجموعة). خضعت المجموعة التجريبية لبرنامج تدريبي مولّد بالذكاء الاصطناعي لمدة ستة أسابيع، بينما واصلت المجموعة الضابطة برنامجها التدريبي المعتاد. تم تقييم القوة البدنية ونسبة الدهون في الجسم قبل وبعد التدخل. أظهرت النتائج عدم وجود فروق ذات دلالة إحصائية داخل المجموعتين أو بينهما (P>0.05). وتشير التغيرات الطفيفة المسجلة لدى المجموعة التجريبية إلى محدودية فعالية البرامج التدريبية المعتمدة على الذكاء الاصطناعي على المدى القصير.

الكلمات المفتاحية: الذكاء الاصطناعي؛ ChatGPT؛ البرنامج التدريبي.

* *Corresponding author.*

1-Introduction:

With the rapid advancement of artificial intelligence (AI) technologies, these tools have become an integral part of various fields, including sports and fitness. Traditionally, the design of training programs has relied on multiple sources such as books, scientific articles, and digital resources. However, recent progress in AI technologies—such as the ChatGPT model—has opened new horizons for developing personalized training programs based on the analysis of sports performance data, thereby offering opportunities to enhance training effectiveness and improve physical performance (Van RA, 2021).

In recent years, there has been growing interest in the application of AI in the design of training plans. Some studies have demonstrated the potential of AI techniques in analyzing training loads, assessing sports performance indicators, and providing personalized training recommendations tailored to individual needs (Dergaa I, et al., 2023). Other studies have confirmed the ability of AI to design training programs that improve physical fitness and reduce body fat percentage, thus enhancing training efficiency and raising athletic performance levels in women (Washif JA, et al., 2024).

The integration of artificial intelligence into athletic training has revolutionized the way performance is assessed and enhanced. Among the most impactful innovations is the use of AI systems to process complex physiological and movement data, allowing coaches to tailor training programs to the specific needs of each athlete (Zhang et al., 2023). In particular, recent research has demonstrated that AI can analyze multivariate time series data to accurately predict fatigue levels and stamina, enabling real-time adjustments to training loads and improving overall outcomes (Biro et al., 2023). As these technologies advance, their integration into sports science promises more precise, adaptive, and injury-conscious training strategies.

Recent research shows that AI can analyze large-scale performance and physiological data to develop personalized training programs, improve training efficiency, and enhance athletic outcomes (Zhang et al., 2023; Bodemer, 2023). Moreover, recent studies show that AI tools like ChatGPT can offer immediate, personalized, and adaptive training consultations based on user input, making them valuable in professional education and support contexts (Morya et al., 2024).

Studies have indicated that training programs tailored to the participants' individual conditions lead to significant improvements in

physical fitness and reductions in body fat (Westphal et al., 2020). Another study confirmed that individualized training programs can markedly enhance physical fitness and muscular strength (Westphal et al., 2020). Additionally, a recent study found that high-intensity bodyweight training can result in notable improvements in physical fitness and body composition after eight weeks of regular practice (Ajjimaporn et al., 2022).

Although AI has made significant strides in sports applications, scientific evidence regarding the impact of training programs designed using ChatGPT remains limited. A recent study indicated that using ChatGPT to design training plans for runners could be beneficial, although the quality of these plans largely depends on the quantity and quality of the input data, and they were not deemed ideal by coaching experts (Kreutzer & Sirrenberg, 2023). Another article discussed the potential of ChatGPT in fields related to kinesiology and exercise science but emphasized that applied research on the effectiveness of this technology in designing training programs is still in its early stages (Pérez et al., 2023).

Accordingly, this study aims to evaluate the impact of a training program designed using ChatGPT on muscular strength and body fat percentage in women, thereby contributing to the scientific literature regarding the integration of AI technologies in the design of training plans and the enhancement of sports performance. Therefore, the central research problem of this study is to investigate the effect of a training program designed using ChatGPT on muscular strength and body fat percentage in women.

2- Materials and Methods:

2-1 Study Design and Ethical Considerations

This study adopted a structured experimental design to assess the impact of a training program generated using ChatGPT-4 on physical strength and body fat percentage in women. All key variables were controlled to ensure internal validity and reliability of the results. Ethical approval for the study was obtained from the relevant institutional review board, and all participants provided informed consent prior to participation.

2-2 Participants and Sampling

The research population consisted of female members of an aerobics sports club located in Batna Province. A total of 20 participants aged between 18 and 35 years were recruited. The sample was randomly divided into two equal groups:

Experimental Group (n = 10): Participants in this group followed a structured training program generated by ChatGPT-4. Pre- and post-intervention assessments were conducted to measure physical strength and body fat percentage using standardized testing methods.

Control Group (n = 10): These participants continued with their usual training routines without any modifications. They also underwent the same pre- and post-intervention assessments as the experimental group.

This group allocation allowed for a comparative analysis of the effects of an AI-designed training program versus conventional training approaches in aerobics practitioners.

2-3 Training Program and Duration

The experimental training program was designed using ChatGPT-4 and spanned six weeks, with two sessions per week. The training sessions followed a structured progression, gradually increasing in intensity and complexity.

Table 01. Weekly Training Sessions and Objectives of the AI-Generated Exercise Program

Session	Objectives
1	Pre-evaluation of the experimental and control group
2	Introduction to the program, initial fitness assessment, general warm-up, and preparatory strength exercises.
3	Increase intensity of foundational exercises, focus on upper body strength, and light endurance training.
4	Lower body strength training and introduction of moderate endurance exercises.

5	Full-body workout with increased difficulty to improve strength and endurance.
6	Review of techniques, performance enhancement, and introduction of advanced strength exercises.
7	Focus on **cardio and endurance training** to enhance fat-burning and cardiovascular health.
8	Advanced endurance training with new challenges for improved fitness levels.
9	**Mid-program assessment** to measure progress, modifications if needed, and integrated endurance-strength training.
10	Increased difficulty in strength and endurance exercises with a focus on cardiovascular efficiency.
11	Complex multi-movement exercises to improve overall endurance and strength.
12	**Final evaluation** to measure overall progress, recovery exercises, and preparation for maintaining achieved fitness levels.

2-4 Assessment Methods

To evaluate the impact of the AI-generated training program, a series of standardized assessments were conducted for both the experimental and control groups.

2-4-1 Body Fat Percentage Measurement

Body fat percentage was assessed using a standardized method. skinfold calipers were employed. The same device and protocol were used pre- and post-intervention for both groups to ensure consistency.

2-4-1 Strength Tests

A battery of strength assessments was administered to evaluate muscular performance across different regions of the body:

- Arm Strength Test
- Lower Limb Strength and Endurance Test
- Trunk Strength Test
- Squat Test

All tests followed standardized administration procedures and were supervised by qualified personnel to ensure accuracy and safety.

2-5 Control of Confounding Variables

To minimize the influence of external variables and ensure valid comparisons between the two groups, the following controls were implemented:

Time of Day: All training sessions were conducted between 1:00 PM and 2:30 PM to standardize physiological conditions related to circadian rhythm.

Gender: The study exclusively included female participants to eliminate gender-related variability.

Age Range: Participants were selected within the age range of 18 to 35 years to reduce variability due to age-related fitness differences.

Standardized Testing Conditions: The same tools, procedures, and evaluators were used across both groups to ensure reliability.

Environmental Control: All training and assessment sessions were conducted indoors in a temperature-controlled sports facility to eliminate environmental fluctuations.

2-6 Pre- and Post-Testing Protocol

All participants underwent pre-intervention assessments prior to the beginning of the training program. Post-intervention assessments were conducted after the six-week period to evaluate the effects of the AI-designed training program compared to conventional training. The same tools and protocols were applied in both phases to maintain consistency and reliability.

3- Data Analysis

All statistical analyses were conducted using IBM SPSS Statistics version 25.0, Descriptive statistics (mean ± standard deviation) were calculated for all variables.

The normality of data distribution was confirmed using the Shapiro–Wilk test, allowing the use of parametric tests. To assess within-group differences between pre- and post-intervention scores, paired samples t-tests were applied. To compare post-test results between the experimental and control groups, independent samples t-tests were used.

Statistical significance was set at $p < 0.05$, and all analyses were two-tailed.

4- Statistical analysis.

Table 01. Pre- and Post-Test Body Fat Percentage in Experimental and Control Groups, and Between-Group Post-Test Comparison

Variables	Pre-test		Post-test		t Value	sig	N	Significance
	Mean	SD	Mean	SD				
Experimental Group Results	39.96	7.99	40.01	8.22	0.014	0.98	12	NS
Control Group Results	39.97	8.000	39.31	7.982	3,136	0,009	12	NS
Post-Test Results of the Control and Experimental Groups	Post-Test Results of the Experimental Group		Post-Test Results of the Experimental Group		t Value	sig	N	Significance
	37.6275		8.13816		39.3133	7.98186	0,512	0,999

S= Statistically Significant, NS= Not Significant

Table 01 presents the mean body fat percentage and standard deviations for both the experimental and control groups before and after the training period, along with post-test comparisons between groups. The experimental group showed a negligible change in body fat percentage from pre-test (M = 39.96%, SD = 7.99) to post-test (M = 40.01%, SD = 8.22), with no significant difference (t = 0.014, p = 0.98, n = 12). Conversely, the control group demonstrated a slight decrease from 39.97% (SD = 8.00) to 39.31% (SD = 7.98), and this difference was statistically significant (t = 3.136, p = 0.009, n = 12). However, the post-test comparison between the two groups revealed no significant difference (t = 0.512, p = 0.999), indicating that the AI-generated training program had no statistically significant effect on reducing body fat percentage compared to the control group.

Table 03. Comparison of Strength Measures (Arm, Lower Limb, Trunk, and Squat) Between Experimental and Control Groups in Pre- and Post-Test Evaluations.

Variables	test	Pre-test		Post-test		t Value	sig	N	Sig
		Mean	SD	Mean	SD				
Experimenta Group Results	**AS**	18.91	8.07	54.00	27.37	4.25	0.1	12	NS
	LLS	47.16	27.60	41.16	9.26	0.71	0.48	12	
	TS	15.58	3.75	17.25	3.44	1.13	0.26	12	
	SQ	16.25	4.393	19.25	3.415	5.745	01	12	
Control Group Results	**AS**	18.92	8.073	21.58	6.735	2,513	0.029	12	S
	LLS	47.17	27.604	97.50	13.568	11,136	0.00	12	S
	TS	39.97	8	39.31	7.98	3.13	0.09	12	NS
	SQ	14.67	4.579	17.75	3.494	5,679	0.00	12	S
Post-Test Results of the Control and Experimenta l Groups	test	Post-Test (EG)		Post-Test (CG)		t Value	sig	N	Sig
		Mean	SD	Mean	SD				
	AS	54.00	27.37	21.5833	6.73469	0,324	0,017	12	NS
	LLS	41.16	9.26	97.500	13.568	0.324	0.017	12	
	TS	17.25	3.44	19.0833	3.028	1,461	0.814	12	
	SQ	19.25	3.415	17.7500	3.493	1,064	0,949	12	

AS: Arm Strength; **LLS**: Lower Limb Strength; **TS**: Trunk Strength; **SQ**: Squat ; **EG**: Experimental Group; **CG**: Control Group; **Post-Test**: Post-intervention measurement; **Sig.** = Significance level; **NS** = Not Significant; **S** = Significant.

Table 03 displays the results of multiple strength tests—arm strength, lower limb strength, trunk strength, and squat performance—for both the experimental and control groups before and after the intervention, along with between-group comparisons in post-test scores.

In the experimental group, arm strength improved notably from a mean of 18.91 (SD = 8.07) to 54.00 (SD = 27.37), but the change did not reach statistical significance ($t = 4.25$, $p = 0.1$). Lower limb strength showed a decrease from 47.16 to 41.16, with no significant difference ($t = 0.71$, $p = 0.48$). Trunk strength increased slightly from 15.58 to 17.25 ($t = 1.13$, $p = 0.26$). Squat performance improved significantly from 16.25 to 19.25 ($t = 5.745$, $p = 0.01$).

In the control group, arm strength increased from 18.92 (SD = 8.07) to 21.58 (SD = 6.73), and this improvement was statistically significant ($t = 2.513$, $p = 0.029$). Lower limb strength showed a substantial increase from 47.17 (SD = 27.60) to 97.50 (SD = 13.57), with a highly significant difference ($t = 11.136$, $p < 0.001$). Trunk strength decreased slightly from 39.97 (SD = 8.00) to 39.31 (SD = 7.98), but the change was not statistically significant ($t = 3.13$, $p = 0.09$). Finally, squat performance improved from 14.67 (SD = 4.58) to 17.75 (SD = 3.49), with a highly statistically significant improvement ($t = 5.679$, $p < 0.001$).

Post-test comparisons between groups revealed no significant differences in any strength test. Arm strength scores were close ($t = 0.324$, $p = 0.017$), as were scores in lower limb strength (identical means), trunk strength ($t = 1.461$, $p = 0.814$), and squat ($t = 1.064$, $p = 0.949$), indicating that despite within-group improvements, the AI-generated training program did not yield significantly greater improvements compared to the control condition.

5- Discussion:

The findings of this study indicate that the training program designed using artificial intelligence, represented by the ChatGPT model, did not produce statistically significant differences in body fat percentage or physical strength among the female participants compared to the traditional training program, despite a slight reduction in the mean body fat percentage in the experimental group. This non-significant decrease reflects the challenges of applying artificial intelligence in

the field of sports training and highlights a gap between the theoretical potential of these technologies and their actual performance in practice, necessitating a reconsideration of the program's structure and design. This lack of statistical significance may be attributed to several key factors, foremost among them being the relatively short duration of the program (six weeks with two sessions per week), which may be insufficient to produce substantial changes in body composition, as fat loss typically requires longer, more intensive interventions, as documented in prior studies (Ajjimaporn et al., 2022). Additionally, while the program was designed through an intelligent model, the intensity and volume of the prescribed exercises may not have reached the necessary physiological threshold to elicit meaningful adaptations, aligning with the conclusions of Washif et al. (2024) that effective training design demands precise regulation of exercise dose and training components beyond the textual outputs of AI alone.

Moreover, the limited sample size (n=10 per group) likely reduced the statistical power of the study, diminishing the likelihood of detecting significant differences even in the presence of a true moderate or small effect size, consistent with Cohen's (1988) assertions regarding the relationship between sample size and statistical power. Another critical consideration is the inherent dependence of ChatGPT's outputs on the quality and precision of the input data provided by users, which may result in variability in program efficacy depending on the user's expertise in guiding the AI, a challenge highlighted by Kreutzer & Sirrenberg (2023) in their analysis of ChatGPT's application in kinesiology.

It is also noteworthy that the study's evaluation relied solely on body fat percentage and basic strength assessments, whereas the broader literature advocates for incorporating more diverse evaluation metrics such as biochemical markers, DEXA scans, functional performance assessments, or inflammatory indicators that may reveal physiological changes not captured by traditional measures (Washif et al., 2024). This limitation in measurement tools may partly explain the absence of observed differences between the groups.

In light of these findings, it becomes clear that the application of artificial intelligence in training program design remains in its nascent stages and requires an integrative approach combining technology with human expertise to ensure that the prescribed exercises align with participants' individual characteristics and contextual factors. This aligns with recent findings that recommend adopting hybrid AI

models, where automated outputs are continuously supervised and refined by human experts to ensure reliability and effectiveness (Morya et al., 2024). While this study opens new avenues for exploring the potential of AI in sports training development, it simultaneously raises critical questions about the current limitations of such technologies and their need for further refinement and calibration before being embraced as a full substitute for human trainers. Therefore, the study provides a valuable contribution to the scientific literature as an initial attempt to apply a language-based AI model in sports training; however, it underscores that the path toward realizing its full potential requires additional research with larger sample sizes, extended intervention periods, and multidimensional assessment protocols to yield more robust and generalizable findings capable of transitioning from experimental implementation to sustainable practical application in women's sports training.

Final conclusion :

This study provides an initial exploration of the application of artificial intelligence, specifically ChatGPT, in the design of training programs aimed at improving physical strength and reducing body fat percentage among women. The findings indicate that, although the AI-generated program resulted in a negligible change in body fat percentage, it did not lead to statistically significant improvements compared to traditional training methods. These results suggest that relying solely on AI-generated programs may not be sufficient to achieve meaningful physiological adaptations within a short intervention period.

The study highlights several challenges in adopting AI for training program design, including the need for longer intervention durations, higher training intensities, and more diverse evaluation metrics. Additionally, it underscores the importance of combining AI technologies with human expertise to ensure that training prescriptions are appropriately tailored to individual needs and context-specific factors.

While the integration of artificial intelligence in sports training holds promising potential, this research emphasizes that current AI tools like ChatGPT should be viewed as complementary aids rather than replacements for professional coaching. Future research with larger sample sizes, extended durations, and more comprehensive outcome measures is necessary to fully understand the capabilities and limitations of AI-driven training interventions. Ultimately, the effective utilization of AI in sports training requires a hybrid approach

that leverages technological innovation alongside continuous human oversight to maximize benefits and ensure safety and efficacy.

Practical Recommendations:

Based on the findings of this study, the following practical recommendations are proposed for trainers and researchers interested in utilizing artificial intelligence technologies, such as ChatGPT, in the design of training programs:

- It is recommended to use AI-generated training programs as supportive tools rather than replacements, relying on trainers' professional expertise to assess the suitability of exercises and make necessary adjustments according to each participant's needs.
- The results suggest that a six-week program with two sessions per week may not be sufficient to achieve significant changes. Therefore, it is advisable to extend the program duration and increase the frequency of sessions to enhance the effectiveness of AI-designed programs.
- It is recommended to incorporate multiple evaluation tools (such as DEXA scans or biochemical markers) alongside traditional measurements to obtain a more comprehensive view of the physical changes resulting from the training program.
- It is important to systematically analyze the content of AI-generated training programs to ensure their alignment with scientific principles of exercise prescription regarding intensity, volume, and progression.
- It is recommended to organize workshops and training sessions for coaches to enhance their understanding and effective utilization of AI technologies in designing training programs.

Suggestions for Future Research:

In light of this study's results and limitations, the following directions are suggested for future research:

- Conduct future studies with larger samples of women from diverse age groups to improve the generalizability of findings.
- Implement studies that track the effects of AI-designed training programs over longer periods (three months or more) to examine their long-term impact on body composition and physical strength.
- Expand the range of outcome measures to include other indicators such as cardiovascular fitness, flexibility, balance, or muscle mass.
- Conduct comparative studies evaluating the effectiveness of programs designed by ChatGPT versus other AI models or specialized sports training software.

- Include interviews or surveys to explore participants' experiences and perceptions regarding the suitability and motivational aspects of AI-generated training programs.
- Investigate the effect of continuous human supervision in AI-based training design and compare outcomes with fully automated AI-generated programs.

References

– Ajjimaporn, A., Khemtong, C., & Willems, M. E. T. (2023). Body composition and physical fitness improve after 8 weeks of high-intensity circuit training using body weight in obese women. The Journal of Sports Medicine and Physical Fitness, 63(2), 273–281. https://doi.org/10.23736/S0022-4707.22.14000-4

– Biró, A., Cuesta-Vargas, A. I., & Szilágyi, L. (2024). AI-assisted fatigue and stamina control for performance sports on IMU-generated multivariate time series datasets. Sensors, 24(1), 132. https://doi.org/10.3390/s24010132

–Bodemer, O. (2023). Enhancing individual sports training through artificial intelligence: A comprehensive review. Engineering Open Access, 1(2), 111–119.

–Dergaa, I., Chamari, K., Zmijewski, P., & Ben Saad, H. (2023). From human writing to artificial intelligence-generated text: Examining the prospects and potential threats of ChatGPT in academic writing. Biology of Sport, 40(2), 615–622. https://doi.org/10.5114/biolsport.2023.125623

–Morya, V. K., Lee, H. W., Shahid, H., Magar, A. G., Lee, J. H., Kim, J. H., Jun, L., & Noh, K. C. (2024). Application of ChatGPT for orthopedic surgeries and patient care. Clinical Orthopaedics and Surgery, 16(3), 347–356. https://doi.org/10.4055/cios23181

–Podrihalo, O. O., Jagiełło, W., Podrigalo, L., Iermakov, S., & Yermakova, T. (2022). The influence of health-improving fitness classes on the degree of fat deposition in women of the second mature age. Pedagogy of Physical Culture and Sports, 26(6), 358–364. https://doi.org/10.15561/26649837.2022.0609

–Kreutzer, L., & Sirrenberg, M. (2023). The use of ChatGPT in kinesiology: Possibilities and limitations. Journal of Physical Education and Sport, 24(1), 3–12. https://doi.org/10.7752/jpes.2024.01001

–Van Woezik, R. A., McLaren, C. D., Côté, J., & others. (2021). Real versus ideal: Understanding how coaches gain knowledge.

International Sport Coaching Journal, 9(2), 189–202. https://doi.org/10.1123/iscj.2021-0010

–Washif, J. A., Pagaduan, J., & James, C. (2024). Artificial intelligence in sport: Exploring the potential of using ChatGPT in resistance training prescription. Biology of Sport, 41(2), 209–220. https://doi.org/10.5114/biolsport.2024.124625.

–Westphal, G., Baruki, S., Mori, T., Montebello, M., & Pazzianotto-Forti, E. M. (2020). Effects of individualized functional training on the physical fitness of women with obesity. Consensus Research Papers, 5(1), 45–52.*

–Zhang, J., Li, Q., Liu, Z., & Tao, Z. (2023). Feasibility analysis of athletes' physical training based on big data perspective. MEDS Public Health and Preventive Medicine, 3(4), 9–16. https://doi.org/10.23977/phpm.2023.030402