Korean Journal of Nuclear Medicine Technology, Volume 30, Number 1, Article 6, May 2026.
Korean J Nucl Med Technol 2026;30(1):6. https://doi.org/10.12972/kjnmt.2026.30.1.6

pISSN 1229-9901
eISSN 2982-8406

Original article

User Experience and Learning Enhancement of a Nuclear Medicine Specialized GPTs Model

Hoon-Hee Park, Ph.D., RT[1*], Jin-Woo Jung, RT[2], Jooyoung Lee, Ph.D., RT[1]

[1]Department of Radiological Science, Shingu College

[2]Department of Radiation Oncology, SoonChunHyang University Hospital Seoul

[*]**Corresponding Author: Hoon-Hee Park,** PhD. Department of Radiological Science, Shingu College. 377 Gwangmyeong-ro, Jungwon-gu, Seongnam, 13174, Republic of Korea. Tel: +82-10-8944-5497, E-mail: hzpark@shingu.ac.kr

Abstract

Purpose: This study aimed to develop a Generative Pre-trained Transformer-based model specialized in nuclear medicine education and patient guidance and to evaluate its effectiveness as a supplementary learning tool **Materials and Methods:** The specialized system was designed to provide tailored interactions for students and the general public through role-specific prompts and structured explanations based on nuclear medicine textbooks. Thirty-two participants experienced the system and assessed its performance across four domains: user experience, suitability, information quality, and reliability, using a 5-point Likert scale. **Results:** All evaluation domains achieved mean scores of approximately 4.0, indicating overall positive responses. Participants highly rated the system for its structured explanations (mean 4.15±0.71), interactive follow-up questions (mean 4.34±0.64), and efficient information retrieval (mean 4.03±0.72). However, visual formatting (mean 3.90±0.80) and the risk of receiving incorrect information (positive response rate 62.5%) were identified as areas requiring further improvement. **Conclusion:** Domain-specific GPT models can effectively supplement traditional textbooks by enhancing understanding and supporting self-directed learning in nuclear medicine.

Keywords: GPT (Generative Pre-trained Transformer), UX (User Experience), Learning Enhancement

Introduction

Recent advances in AI (artificial intelligence), particularly LLMs (large language models) and GPT series models, have introduced new possibilities across medical education and the broader field of healthcare [1-3]. These models, trained on vast amounts of text data, are capable of understanding context and generating natural language, enabling diverse applications such as educational content creation, learning support, and knowledge exploration. In the field of radiology, Radiology-GPT has demonstrated the potential of specialized models in research, education, and clinical support, while studies applying iterative optimization frameworks have reported improvements in the accuracy and consistency of report summarization [1-2]. Furthermore, research on the educational application of ChatGPT has emphasized the need to move beyond simple performance evaluation to consider UX and learning effectiveness [3].

In Korea, research on domain-specific language models in healthcare education is also emerging. For instance, a study focusing on the national nursing licensure examination reported an average ChatGPT accuracy of 49.5%, with performance exceeding 60.0% in some subjects, suggesting its potential as a supplementary learning tool [4]. More broadly, a recent meta-analysis demonstrated that conversational AI can positively influence not only knowledge acquisition but also learning attitudes and higher-order thinking skills [5].

Received April 06, 2026 **Revised** April 20, 2026 **Accepted** April 20, 2026

Nuclear medicine is a specialized field that utilizes radiopharmaceuticals to visualize and quantitatively evaluate physiological and pathological processes. PET (Positron Emission Tomography), in particular, provides functional information such as metabolism, blood flow, and receptor binding, playing a critical role in early diagnosis and treatment response assessment. However, the complexity of radiopharmaceutical administration and examination procedures imposes a high cognitive burden on learners and the general public. Insufficient understanding may lead to reduced learning efficiency or unnecessary anxiety regarding the examination process. Therefore, there is an increasing need for intuitive, repeatable learning support tools and information delivery systems. While future perspectives in nuclear medicine highlight the promise of AI, they also point out accompanying ethical and technical challenges [6]. Research on a nuclear medicine-specific conversational language model has confirmed its applicability for self-directed learning and clinical simulation, though studies on DRG (Diagnosis-Related Group) classification and question-answering accuracy report varying degrees of success in real clinical workflows [7].

Despite these advancements, many existing studies still primarily focus on model performance metrics or technical feasibility. A systematic review covering the broader medical field highlighted the important role of conversational language models in education and information dissemination [8]. Additionally, a comparative study based on national examinations showed that GPT-4 achieved significantly higher accuracy than GPT-3.5 [9]. To address the limitations of purely performance-based metrics, a proposed evaluation framework for medical language models emphasized the critical need for comprehensive UX and safety validation [10]. Another recent systematic review comprehensively analyzed the impact of AI-based conversational systems in healthcare professional education and assessment, further underscoring the importance of evaluating UX and reliability [11].

Therefore, this study aims to develop a conversational language model-based learning support system for nuclear medicine education and public information provision, evaluating its application among both learners and general users. This study assesses UX, suitability, information quality, and reliability through both quantitative and qualitative analyses. By providing empirical evidence on whether a domain-specific conversational system can improve learning efficiency and information delivery in nuclear medicine, this study seeks to suggest future directions for the development and implementation of medical education tools.

Materials and Methods

1. Study Design and Objectives

This study is a quantitative descriptive study that developed GPTs (Generative Pre-trained Transformer; customized GPT models) specialized for nuclear medicine education and guidance, and evaluated UX and satisfaction by providing them to actual users. The purpose of this study is to verify the educational and practical applicability of GPTs, and UX was collected and analyzed in a multidimensional manner.

2. Development Process of Nuclear Medicine GPTs

The developed GPTs were designed to support nuclear medicine education and guidance, and were structured to provide tailored interactions based on user type (students vs. general users). At the start of the conversation, users could select either "student" or "general user," after which different recommended prompts were presented accordingly. The overall interaction was organized in a structured Q&A format, focusing on examination purpose, procedures, precautions, and underlying principles(Fig. 1).

Fig. 1. Configuration of role-specific prompts for different user groups

The training and response content of the model were constructed based on nuclear medicine textbooks used at the university level. The content focused on key elements such as examination purpose, methods, precautions, and covered major nuclear medicine procedures included in the academic curriculum. In the initial development stage, a UX flowchart was established to design the interaction sequence as follows: "start screen → user role selection → role-based recommended questions → user input → AI response → follow-up question suggestions → additional information or termination." Based on this structure, prompt engineering principles were applied to differentiate the level of questions and explanations according to user roles.

For students, the system emphasized academic topics such as examination principles, equipment QC (Quality Control) and procedures. For general users, it guided comprehension-oriented questions, including definitions, preparation, safety, and potential side effects of examinations. The system was restricted to providing textbook-based explanations only, and medical judgments such as diagnosis, treatment, and interpretation of results were not permitted. In cases of uncertainty or insufficient evidence, a conservative response was fixed: "Accurate information could not be confirmed. Please consult a medical professional."

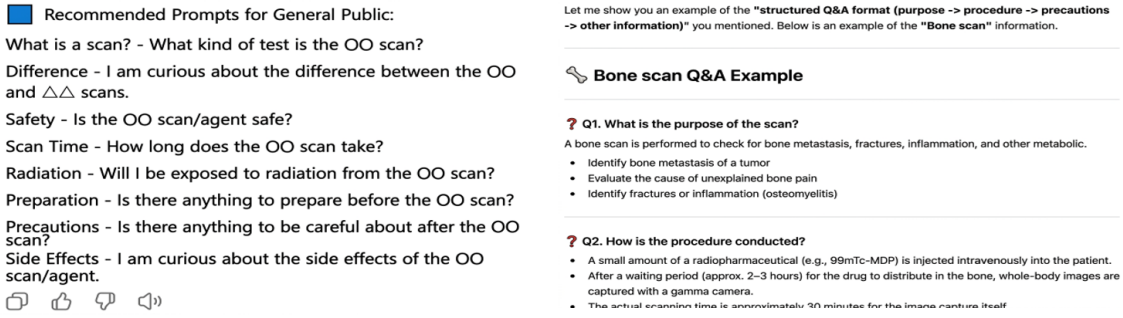

Fig. 2. Operational interface and response example of the nuclear medicine GPT model for general users.

3. Survey Questionnaire Design

A total of 30 questionnaire items were developed to assess UX. Each item was rated on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). The evaluation domains included four categories: △ UX △user suitability △information quality △reliability.

4. Study participants

This study was conducted after providing participants with sufficient information about the research purpose and procedures and obtaining

their voluntary consent. All surveys were conducted anonymously, and personal information protection and research ethics were strictly maintained to ensure that the collected data were used solely for research purposes.

The study participants consisted of students and general users who experienced the GPTs. Respondent characteristics were classified based on their level of nuclear medicine knowledge (ranging from no knowledge to high knowledge) and the device used (PC or mobile). However, due to structural limitations of the survey instrument, group-level analysis distinguishing between students and general users could not be performed.

The survey was distributed through an online platform Naver Form (Naver, Seongnam, Korea), and participants responded anonymously after reviewing the study information.

Results

This study conducted a survey of 32 participants to evaluate the UX and suitability of GPTs specialized in nuclear medicine. The questionnaire consisted of four domains: UX, user suitability, information quality, and reliability, with each item rated on a 5-point Likert scale.

For each item, the mean, standard deviation, positive response rate (scores of 4–5), and negative response rate (scores of 1–2) were calculated.

Participants were categorized based on characteristics such as their level of nuclear medicine knowledge and the device used.

Table 1. Results of the distribution of respondent characteristics

Category	n	%
Not at all knowledgeable	7	21.9
Slightly knowledgeable	8	25.0
Moderately knowledgeable	11	34.4
Very knowledgeable	6	18.7
Extremely knowledgeable	0	0.0

n = Number of respondents; % = Percentage.

A total of 32 respondents participated in the survey. Regarding the level of nuclear medicine knowledge, 7 participants (21.9%) reported "Not at all knowledgeable," 8 (25.0%) reported "Slightly knowledgeable," 11 (34.4%) reported "moderatey knowledge," and 6 (18.7%) reported "very knowledge," with no respondents indicating "extremely knowledgeable knowledge" (Table 1).

Table 2. Results of the user experience evaluation

Category	M	SD	%
Easy to understand	3.96	0.58	81.2
Intuitive to use	3.93	0.65	75.0
More efficient	4.03	0.72	81.2
Well organized	4.15	0.71	87.5
Easy-to-read format	3.90	0.80	68.7
Fast enough No delay	4.25	0.61	90.6
Starter buttons helpful	3.93	0.65	75.0
No confusing steps	3.78	0.89	65.6
Accurate guidance	4.06	0.82	81.2
Useful follow-ups	4.34	0.64	90.6
Helpful examples	4.34	0.68	93.7

M = Mean; SD = Standard Deviation; % = Positive Response Rate

In the UX domain, overall evaluations were positive. The item evaluating "Helpful examples" showed the highest score, with a mean of 4.34 and a positive response rate of 93.7%, indicating that the provided examples and explanations were effective in enhancing learning comprehension. This was followed by "Useful follow-ups" (mean 4.34, positive response rate 90.6%) and "Fast enough No delay" (mean 4.25, positive response rate 90.6%), demonstrating that the interactivity and responsiveness of the GPTs contributed to maintaining learning engagement.

Additionally, items such as "Well organized" (87.5%), "Accurate guidance" (81.2%), "More efficient" (81.2%), and "Easy to understand" (81.2%) also received generally positive evaluations, reflecting satisfaction with the systematic organization, logical learning sequence, and efficiency compared to traditional materials. In contrast, "Easy-to-read format" (68.7%) and "No confusing steps" (65.6%) received relatively lower scores, suggesting the need for improvement in information visualization and process simplification (Table 2).

Table 3. Results of the user suitability evaluation

Category	M	SD	%
User-type fit	4.15	0.75	78.1
Purpose fit	4.21	0.73	87.5
Level of detail	3.96	0.84	68.7
Understanding improved	4.21	0.69	84.3
More efficient than traditional	4.06	0.82	81.2

M = Mean; SD = Standard Deviation; % = Positive Response Rate

In the user suitability domain, overall evaluations were positive. The items "Purpose fit" and "Understanding improved" showed the highest results, with mean scores of 4.21 and positive response rates of 87.5% and 84.3%, respectively, indicating that the GPTs were effective in achieving learning objectives and enhancing comprehension.

The item "More efficient than traditional" methods (e.g., textbooks, lectures) also received a relatively high positive response rate of 81.2%. In contrast, "Level of detail" had the lowest score, with a mean of 3.96 and a positive response rate of 68.7%, suggesting limitations in delivering sufficiently detailed and personalized information. Additionally, "User-type fit" showed a somewhat lower positive response rate of 78.1%, indicating a need for further optimization in tailoring explanations for specific groups (student vs. general user)(Table 3).

Table 4. Evaluation of information quality

Category	M	SD	%
Direct relevance	3.93	0.99	65.6
Step-by-step clarity	4.25	0.66	87.5

M = Mean; SD = Standard Deviation; % = Positive Response Rate

In the information quality domain, variations were observed across items. The item evaluating "Step-by-step clarity" showed the highest score, with a mean of 4.25 and a positive response rate of 87.5%, indicating that the GPTs have strengths in structuring and explaining complex concepts step-by-step to aid understanding.

In contrast, the item evaluating "Direct relevance" received a relatively lower evaluation, with a mean of 3.93 and a positive response rate of 65.6%, suggesting that some respondents perceived the inclusion of unnecessary or non-essential information in the responses rather than being directly relevant to the topic (Table 4).

Table 5. Results of the reliability evaluation

Category	M	SD	%
Evidence-based	3.80	1.07	75.0
Will refer in future	3.94	0.62	90.6
Low misinformation risk	4.00	0.98	62.5
Recommend to others	4.09	0.75	78.1
Apply to other topics	3.78	0.89	65.6

M = Mean; SD = Standard Deviation; % = Positive Response Rate

In the reliability domain, overall scores were somewhat lower compared to other domains. The item "Will refer in future" showed the highest positive response rate at 90.6%, indicating a strong intention among respondents for practical use. In contrast, "Low misinformation risk " had the lowest positive response rate at 62.5%, suggesting that concerns about potential inaccuracies still remain.

Additionally, "Recommend to others" received a relatively positive response rate of 78.1%, while "Evidence-based" recorded 75.0%. On the other hand "Apply to other topics" showed a comparatively lower response rate of 65.6%(Table 5).

Discussion

The nuclear medicine GPTs developed in this study demonstrated their potential as a supplementary learning tool by providing real-time question-and-answer interactions, thereby enhancing learners' understanding and improving learning efficiency. Conventional nuclear medicine textbooks and lecture materials often present a high level of difficulty due to specialized terminology and complex procedures. In contrast, the GPTs enabled learners to identify and address gaps in their understanding in real time through interactive explanations, offering an effective learning environment, particularly for beginners in nuclear medicine.

The model incorporated a user-adaptive branching system to deliver different levels of information to general users and students. For general users, it provided simplified explanations of nuclear medicine concepts to improve health literacy, while for students, it delivered more advanced and specialized knowledge to maximize educational utility. Previous studies have reported that the use of ChatGPT not only facilitates knowledge acquisition but also positively influences learning perception and higher-order thinking skills; the personalized information delivery function of this study is aligned with such improvements in learning attitudes and cognitive development.

Despite these strengths, several limitations should be acknowledged. First, the study was limited by a small sample size. The survey was conducted with a relatively small number of participants, most of whom belonged to a specific educational setting, making it difficult to generalize the findings. Future studies should include a broader range of participants across different regions, institutions, and levels of expertise.

Second, there were limitations in the scope and diversity of the data. Although the GPTs were designed based on nuclear medicine textbooks and fundamental learning materials, they did not sufficiently incorporate real clinical cases or the latest research findings. As a result, some responses lacked depth or did not fully meet learners' expectations.

Third, methodological constraints should be considered. This study primarily relied on survey data to assess UX and satisfaction. However, it did not evaluate actual academic performance improvement or long-term learning outcomes. Future research should incorporate additional indicators such as academic achievement, performance changes, and conceptual understanding.

Fourth, technical limitations were also identified. The GPTs were implemented using an open-source–based language model, and some inconsistencies in responses and expression styles were observed. In particular, there were instances where contextual understanding of technical terminology was insufficient or where explanations were overly simplified.

Fifth, due to the limited study period, long-term follow-up was not conducted. It remains unclear whether the positive responses observed in the short term will be sustained over time. This aspect should be addressed in future longitudinal studies.

Conclusion

This study developed GPTs with the aim of enhancing learners' understanding and promoting self-directed learning in the context of nuclear medicine education, and evaluated UX through a survey (n = 32). The questionnaire consisted of four domains—UX, user suitability, information quality, and reliability—and was measured using a 5-point Likert scale. The results showed that the mean scores across all four domains were approximately 4.0, indicating an overall positive trend, with positive response rates (scores of 4–5) exceeding 70.0% for most items.

The GPTs developed in this study demonstrated strengths in improving learning comprehension, encouraging follow-up exploration, maintaining response flow, and providing structured explanations. However, areas for improvement were identified in visual presentation, incorporation of up-to-date information, and communication of reliability. These findings support the potential of GPTs to complement the one-directional nature of traditional textbooks and lectures by enabling continuous, interactive, and personalized learning.

In addition, this program explored both educational and clinical applicability. It can be utilized as a supplementary learning tool for students and as a means to enhance public understanding of nuclear medicine for general users. Furthermore, it presents potential for expansion into clinical settings as a tool for patient education and for supporting communication between healthcare providers and patients.

Such technological approaches may serve as a foundation for improving accessibility in the field of nuclear medicine within a rapidly evolving digital environment. By addressing the identified areas for improvement and further refining the model, it is expected that this system can make a meaningful contribution to enhancing patients' understanding of diagnostic procedures in real clinical practice.

Conflict of Interest

The authors declared no conflicts of interest.

Funding

This research received no external funding.

References

1. Liu X, Zhang Y, Chen H, Wang Z, Li L, Smith A, et al. Radiology-GPT: A large language model for radiology. Radiology. 2025;310(1):151-65
2. Ma C, Wu Z, Wang J, Xu S, Wei Y, Zeng F, et al. An iterative optimizing framework for radiology report summarization with ChatGPT. Front Artif Intell. 2024;5(8):163-75.
3. Chan KY, Yuen TH, Co M. Using ChatGPT for medical education: the technical perspective. BMC Med Educ. 2025;25(1):201.
4. Jang SM. Performance evaluation of ChatGPT on the national examinations. J Korean Acad Soc Nurs Educ. 2024;30(2):121-7.
5. Fatima A, Shafique MA, Alam K, Ahmed TKF, Mustafa MS ChatGPT in medicine: A cross-disciplinary systematic review. Medicine (Baltimore). 2024;103(15):e37845.
6. Hirata K, Matsui Y, Yamada A, Fujioka T, Yanagawa M, Nakaura T, et al.. Generative AI and large language models in nuclear medicine: current status and future prospects. Ann Nucl Med. 2025;39(1):1-12.
7. Papageorgiou N, Diamantopoulos A, Anastasopoulos N, Katsanos K, Kitrou P, Spiliopoulos S, et al. The role of large language models in improving DRG assignment and clinical decision support in radiology and nuclear medicine. Appl Sci. 2025;15(1):1-15.
8. Vachatimanont S, Kingpetch K. Exploring the capabilities and limitations of large language models in nuclear medicine knowledge with primary focus on GPT-3.5, GPT-4 and Google Bard J Med Artif Intell. 2024;7:1-12.

9. Liu M, Okuhara T, Chang XY, Shirabe R, Nishiie Y, Okada H, et al. Performance of ChatGPT across different versions in medical licensing examinations: a systematic review. JMIR Med In form. 2024;12:e53215.
10. Tam TYC, Sivarajkumar S, Kapoor S, Stolyar AV, Polanska K, McCarthy KR, et al. A framework for human evaluation of large language models in healthcare. npj Digit Med. 2024;7(1):1-12.
11. Feigerlova E, Drouin A, Robert C, Lecerf M, Goueslard K, Verdonck L, et al. Impact of artificial intelligence and ChatGPT in training and assessment in health professions education: a systematic review. BMC Med Educ. 2025;25(1):1-18.
12. Yoo K. Nuclear medicine technology. Seoul: Academia; 2018. p. 1-492.
13. Cherry SR, Sorenson JA, Phelps ME. Physics in nuclear medicine. 4th ed. Philadelphia: Saunders; 2012. p. 1-544.
14. Ahn SM, Kim SH, Kim JS, Kim HS, Ryu CJ, Park HH, et al. Nuclear medicine science. 3rd ed. Seoul: Daehak Seorim; 2024. p. 1-556.
15. Kang KW, Kim SE, Lee DS, Chung JK. Koh Chang-soon nuclear medicine. 4th ed. Seoul: Koryo Medicine; 2019. p. 1-1176.