Article

Performance of ChatGPT on Registered Nurse License Exam in Taiwan: A Descriptive Study

Huiman Huang [*]

Article

Performance of ChatGPT on Registered Nurse License Exam in Taiwan: A Descriptive Study

Huiman Huang

Tzu Chi University of Science and Technology; hhuiman@ems.tcust.edu.tw; hhuiman@ems.tcust.edu.tw; Tel.: 886-3-8572158 ex. 2780

Abstract: AI (Artificial Intelligence) chatbots has been widely applied. AI (Artificial Intelligence) chatbots has been widely applied. ChatGPT could enhance individual learning capabilities and clinical reasoning skills, facilitate students' understanding of complex concepts in healthcare education: (1) Background: There is currently less emphasis on its application in nursing education. The application of ChatGPT in nursing education needs to be verified; (2) Methods: A descriptive study is used to analyze the scores of ChatGPT on the registered nurse license exam(RNLE) and compare to the pass rate of candidates in Taiwan; (3) Results: The average score of four exams were around 51.6 to 63.75 by ChatGPT, and pass the RNLE in 2022 1st and 2023 2nd. However, ChatGPT may generate misleading or inaccurate explanations, or it could lead to hallucinations; (4) Conclusions: ChatGPT may have potentially to assist with nursing education. Although AI can simulate human behavior in context, it still cannot replace human creativity, critical thinking, and clinical reasoning.

Keywords: artificial intelligence; ChatGPT; registered nurse license; nursing graduate

1. Introduction

The rapid advancement of technology brings forth new concepts and innovations, causing various impacts in our daily lives. AI (Artificial Intelligence) chatbots has been widely applied [1,2]. ChatGPT (Chat Generative Pre-trained Transformer) is the one of AI chatbots, which could be programmed to generate different perspectives or suggest relevant resources from the prompt [3–5]. Similar to the learning technology of neural networks, ChatGPT underwent training using a vast textual database comprising books, articles, websites, and various other sources of textual information [3–5]. By utilizing an AI chatbot for step-by-step guidance, interactive learning, and feedback, it is possible to enhance the skills and knowledge of students [6–8]. ChatGPT could enhance individual learning capabilities, and clinical reasoning skills, and facilitate students' understanding of complex concepts in healthcare education [6–8]. The generative AI-based applications have the potential to revolutionize virtual learning of students [9]. Despite the potential benefits of ChatGPT as an educational tool, the complete impact it has on education remains uncertain and necessitates additional investigation [5].

The applications of ChatGPT would be as an interactive medical education tool to support learning [1,3]. However, the application of ChatGPT still has its limitations, such as the potential for biased output, the accuracy of the information, and the potential impact on student's critical thinking and communication skills etc. [7,10,11]. The impact of ChatGPT on nursing education has not yet been determined [12,13]. Despite its potential to enhance efficiency, there is currently less emphasis on its application in nursing education [12,14].

The registered nurse license exam (RNLE) is a significant concern for nursing students in Taiwan. It serves as a crucial assessment tool to evaluate the knowledge level of individuals aspiring to become registered nurses. The RNLE is a standardized testing program that comprehensively covers various topics within the nurses' fund of knowledge. To ensure the validity and reliability of exam is a vital issue. It is imperative to thoroughly assess both the potential benefits and risks associated with ChatGPT in order to anticipate and prepare for the future of nursing education. The purpose of this study was to understand the performance of ChatGPT on the registered nurse license exam (RNLE) in Taiwan.

2. Background

ChatGPT first emerged in November 2022. It is a system developed by OpenAI, based on a large language model (LLM), and possesses the ability to simulate human language processing [3–5]. AI chatbots could generate human-like responses based on the input text it received [11,15,16]. ChatGPT provides personalized interactions and responses within a few seconds [17]. ChatGPT empowers the customization of educational materials and evaluation instruments, allowing for personalized adaptations that cater to the unique requirements and learning preferences of individual students [18]. The performance of ChatGPT in complex and intellectually demanding tasks in the medical field still requires further empirical research and evidence-based support [11,19].

The application of ChatGPT has its limitations. ChatGPT would be trained on inadequate datasets, which could lead to the generation of biased or misleading results [11,20]. Ethical issues should be considered [21]. Using AI chatbots could potentially lead to issues of plagiarism, improper referencing, or bias [4,22,23]; the problems of accuracy of the information, copyright, and AI hallucination etc. [4,10,11]. AI hallucination could potentially lead to hallucinations, causing users to rely on measures that make it difficult to ascertain the accuracy of certain content produced by ChatGPT [15,24,25]. Additionally, due to ChatGPT's knowledge cutoff being in 2021, it might not be equipped to handle or provide information on more recent data or developments [11,15]. All these opinions showed it needs to apply evidence-based research to understand the potential shortcomings of implementing ChatGPT in healthcare education [11].

Without any professional training or reinforcement, ChatGPT demonstrated impressive performance in various medical evaluation tasks [2,26], such as showcased its natural language processing capabilities in medical question-and-answer tasks [1,27], reaching a score of 60.2% in the US Medical Licensing Examination (USMLE), 57.7% in Multiple-choice medical exam questions (MedMCQA), and 78.2% in the medical reading comprehension dataset (PubMedQA) [26,28]. ChatGPT's medical knowledge and interpretation in the field of medicine were not yet sufficient for practical application [23]. It is essential to ensure that nursing students are well-informed about the potential risks associated with AI tools.

ChatGPT has constructs the key trends and development in nursing [29]. However, the knowledge development should base on the theory of Bloom to scaffold learning process, the theory consisted of six major categories: knowledge, comprehension, application, analysis, synthesis, and evaluation [30]. An example of Yale New Haven Hospital (YNHH), which was an early adopter of the Rothman Index, a tool that reflects patient acuity and risk, now finds its performance positively influenced by nursing assessment data, indicating significant potential for nurses to impact patient care [31–33]. AI software design could simulate human behavior to a certain extent, but it still fell short when it came to replacing human creativity, critical thinking abilities, providing emotional and personal support, assessment, and nursing care plans [4,29].

The RNLE is a vital exam for nursing students, assessing their knowledge and determining their eligibility as registered nurses in Taiwan. It serves as the benchmark for their professional qualification following nursing education. The Taiwan RNLE consists of a total of five subjects, including "basic medical science (BMS)", "fundamentals of nursing and nursing administration (FNNA)", "medical and surgical nursing (MSN)", "maternal and pediatric nursing (MPN)", and "psychiatric and community health nursing (PCHN)" [34]. In the first administration of the RNLE exam in 2022, the "basic medical science (BMS)" consisted of 80 items. However, following feedback from schools, the number of questions trimmed down from 80 to 50 in the exam of 2022 2nd, as well as in 2023 [34,35]. The other subjects each consist of 80 multiple-choice questions, totaling 400 items in 2022 1st, and then decreased to 370 items since the exam of 2022 2nd. Each subject has a maximum score of 100. The overall examination result is determined by computing the average score across all subjects. To pass, a candidate must attain a minimum total average score of 60. Nevertheless, if a candidate receives a score of zero in any subject, they will not qualify for a passing grade [34].

ChatGPT possesses the capability to offer immediate feedback and corrections to inquiries, thereby aiding students in comprehending intricate concepts that may prove challenging to grasp [36]. However, the utilization of ChatGPT in nursing education carries inherent limitations and risks.

It is imperative to ascertain the effects of ChatGPT on nursing education, as underscored by O'Connor et al. (2022) [12] and Seney et al. (2023) [13]. To proactively anticipate and prepare for the future of nursing education, a comprehensive evaluation of the prospective advantages and risks linked to ChatGPT is essential. Thus, the purpose of this study was: (1) to understand the scores achieved by ChatGPT on the Registered Nurse License Exam (RNLE) and compare the pass rate with nursing graduates in Taiwan; (2) to evaluate the limitations and advantage of using ChatGPT.

3. Materials and Methods

This study utilized a descriptive research approach to evaluate ChatGPT's performance on the Registered Nurse Licensure Examination (RNLE) in Taiwan, comparing it with the test scores of nursing graduates. Additionally, randomly chosen items from each examination employed to conduct a more in-depth evaluation of ChatGPT's explanations and analyses. The registered nurse license exam is held for nursing graduate twice a year in Taiwan as usually. This study would include the exams from 2022 to 2023 of RNLE. The exclusion criteria were RNLE after 2021, due to ChatGPT's knowledge cutoff being in 2021 [11,15]. All questions (items) were prompted by research assistants and encoding by researcher.

The assessment of ChatGPT's performance in the RNLE was carried out employing the GPT-3.5 model, which encompasses diverse elements, including input sourcing, encoding methods, and statistical analysis (Figure 1). The random numbers were used to evaluate explanation of RNLE by Random - Number Generator of Microsoft V 2.0.4.2.
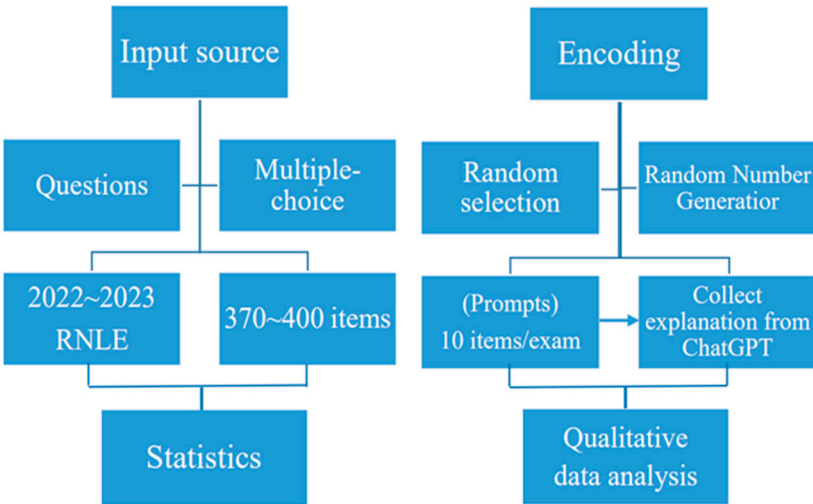

Figure 1. The framework for evaluating ChatGPT's performance on the RNLE.

3.1. Input source

Over the course of four exams spanning from 2022 to 2023, each consisting of five subjects—BMS, FNNA, MSN, MPN, and PCHN—a total of 370 to 400 items were included. The input source encompassed all questions from RNLE in 2022 to 2023, which were then formatted into two sections: one comprising question and the other featuring multiple-choice single answers. Examples include "The patient is suspected of anaphylactic shock, and the medical order is as follows: adrenalin (epinephrine) 1mg/mL/amp 0.3mg IM stat. Which of the following routes of administration is incorrect?" "Which of the following statements about nursing instructions for medication administration in myasthenia gravis patients is correct?"; or "Miss Zhao has been diagnosed with schizophrenia. Which of the following statements about schizophrenia is correct?" "Here's an example of a multiple-choice question with single answers: (A) Discussing the use of asthma medication with a pharmacist, (B) Performing urinary catheterization as instructed by a physician, (C) Instructing a primipara on breastfeeding, (D) Administering antibiotics every 6 hours [37]".

3.2. Encoding

Qualitative data was derived from ChatGPT's responses to RNLE items, encompassing answers, explanations, and narratives. The total number of items included 400 in the exam of 2022 1st, 370 in the exam of 2022 2nd and 2023, arranged sequentially from 1 to 400 or 1 to 370. Based on Random - Number Generator of Microsoft V 2.0.4.2., each exam was randomly selected ten items (questions) within five subjects—BMS, FNNA, MSN, MPN, and PCHN. Provide ChatGPT with randomly sampled items as prompts, research encoded and analyzed the responses generated from ChatGPT. The responses included answer and explanation. An example of explanation like "Mr. Wang, who has undergone a colon resection surgery and just returned to the patient ward from the recovery room, has four health problems identified by the nursing staff according to his needs. In terms of priority for handling, which one is ranked last? (A) Ineffective airway clearance; (B) Acute pain; (C) Lack of knowledge about home care; and (D) Changes in peripheral tissue perfusion [37]. The answer is (C) While important, it holds a lower priority compared to the preceding issues. Once the patient's physiological condition stabilizes, the nursing team can provide more education and guidance [38].

3.3. Statistics and analysis

This study used SPSS for Window 26.0 for data analysis. Descriptive statistics were used to analyze the scores achieved by ChatGPT on the RNLE and compare them to the pass rate of nursing graduates in Taiwan from 2022 to 2023. The qualitative data would be encoded and examined the limitations and risks of using OpenAI's ChatGPT to answer nursing questions. As this study does not involve human or animal subjects, but to the analysis of the results of the RNLE in Taiwan, ethical approval and informed consent were not required.

4. Results

This study was to understand the performance of ChatGPT on the registered nurse license exam (RNLE) and evaluate the limitations and advantage using ChatGPT. According to Taiwanese Ministry of Examination (2023) [34,35], a passing grade requires a total average score of at least 60. The pass rate for the examination of RNLE in 2022 1st and 2023 1st was relatively low, standing at 18.17% and 20.42%, compared to the other exams were 45.31% and 45.73% (Table 1). The number of candidates in these exams also was totally different, the second exam was more candidates than first exam in 2022 and 2023. The number of candidates participating in the RNLE examination in Taiwan exhibits a significant disparity across the four administrations (p = .014), particularly with a notably lower turnout observed in the first examination each year compared to the second. Moreover, the pass percentage for the RNLE in Taiwan also demonstrates a conspicuous discrepancy (p = .024).

Table 1. The performance of candidates on the RNLE in 2022~2023

RNLE	2022 1st	2022 2nd	2023 1st	2023 2nd	p value
Candidates in Taiwan	6,873	14,699	7,476	13,990	.014*
Number of passes	1,031	6,018	1,267	5,758	.083
Pass percentage (%)	18.17	45.31	20.42	45.73	.024*

*p value< .05

4.1. The performance of ChatGPT on RNLE

ChatGPT underwent testing with four exam questions of 370~400 questions (items) in 2022~2023. The scores were between 37.50~78. In the four exams, the lowest score was recorded in the exam of 2022 2nd FNNA examination, with a score of 37.50. Additionally, the exam of MSN in 2023 1st exam yielded a score of 45.25, and the exam of PCHN in 2023 1st resulted in a score of 42.50. The highest

scores were achieved in the exam of BMS in 2023 1st, with a score of 78, as well as in the exam of PCHN in 2022 1st and BMS in 2023 2nd, both of which earned a score of 70. There were significant differences in the scores for all subjects across the four exams from 2022 to 2023(Table 2). Correct percentage were around 54.05~63.75%. The correct percentage were 63.75%, 58.37%, 54.05%, and 60%. ChatGPT performed well in 1st 2022 and 2nd 2023 exams, with percentage of 63.75 and 60.

Table 2. The RNLE scores of ChatGPT

RNLE Items / Subjects*	2022 1st		2022 2nd		2023 1st		2023 2nd		p value
	Correct/ Incorrect	Scores	Correct/ Incorrect	Scores	Correct/ Incorrect	Scores	Correct/ Incorrect	Scores	
BMS**	53/27	66.25	24/26	48.00	39/11	78.00	35/15	70.00	.002***
FNNA	49/31	61.25	30/50	37.50	42/38	52.50	51/29	63.75	.003***
MSN	52/48	65.00	40/40	50.00	37/43	45.25	41/39	51.25	.001***
MPN	45/35	56.25	47/33	58.75	43/37	53.75	47/33	58.75	.000***
PCHN	56/24	70.00	51/29	63.75	39/41	42.50	48/32	60.00	.002***
Correct percentage	63.75%		58.37%		54.05%		60%		
Mean score (SD)	63.75±5.22		51.6±10.16		54.4±4.02		60.75±6.88		.000***

Note. *BMS= basic medical science; FNNA= fundamentals of nursing and nursing administration; MSN= medical and surgical nursing; MPN= maternal and pediatric nursing; PCHN= psychiatric and community health nursing. ** BMS had decreased items from 80 to 50 in 2022 2nd (Ministry of Examination, 2023). *** p value< .05

To speculate the performance of ChatGPT on RNLE in four exams, the average scores for various subjects were dramatically different in 2022 to 2023, however, the subject of MPN showed smaller variations (Figure 2). ChatGPT displayed variations in performance across different subject exams in the RNLE, with its most robust performance observed in the 2022 1st compared to the other exams. With ChatGPT obtained mean scores ranging 51.6±10.16 to 63.75±5.22. ChatGPT required a total average score of 63.75±5.22 and 60.75±6.88 in the exam of 2022 1st and 2023 2nd and passed the examination. On the other hand, compared to the scores of the other two exams, the score was below 60, which was below the standard and failed.
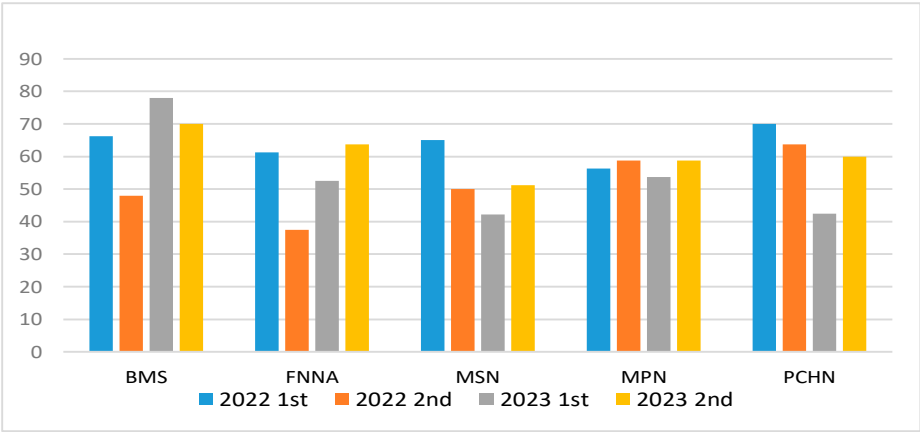
Note. BMS= basic medical science; FNNA= fundamentals of nursing and nursing administration; MSN= medical and surgical nursing; MPN= maternal and pediatric nursing; PCHN= psychiatric and community health nursing. [37]

Figure 2 The performance of ChatGPT on RNLE

4.2. To evaluate the limitations and advantage of using ChatGPT

A total of forty items selected by Random - Number Generator of Microsoft V 2.0.4.2. to examine the answers and explanations of RNLE by ChatGPT. ChatGPT answered a total of 27 questions

correctly with accurate explanations, but it got 13 questions wrong and provided incorrect explanations for them. ChatGPT could answer the most questions directly, such as "Question 1" and "Question 2" displayed (Table 3). After reading the question, ChatGPT would did not provide a correct answer but instead explained the feasibility of each answer, the explanations provided for each question are all accurate, the correct answer rationale cannot be discerned, such as "Question 3". It was no definitive answer to this question (Table 3) [38].

Table 3 The ChatGPT prompts and answers on the RNLE (Example 1)

Question 1	Which of the following supplies nourishment to the lungs through the bronchial arteries? **Answer: (C)** (A) Pulmonary artery (B) Pulmonary vein (C) Thoracic aorta (D) Internal thoracic artery [37].
ChatGPT	**The answer is (C).** The thoracic aorta is responsible for transporting oxygenated blood from the heart to the lungs and supplying the bronchial arteries to nourish lung tissues [38].
Question 2	Mrs. Wang, G1P0, who has had normal prenatal check-ups in the past, presents at 32 weeks of gestation with a complaint of reduced fetal movements and is feeling very concerned. What would be the most appropriate action at this time? **Answer: (C)** (A) Instruct the patient to return home and record daily fetal movements. (B) Explain that this is a common occurrence at this stage and reassure her not to worry. (C) Schedule a non-stress test. (D) Immediately arrange for an induction of labor [37].
ChatGPT	**The answer is (C).** This test can help assess the well-being of the fetus by monitoring its heart rate and response to its own movements, providing valuable information about the baby's condition [38].
Question 3	Regarding the description of the Study of Osteoporotic Fractures frailty index (SOF), which of the following statements is incorrect? **Answer: (B)** (A) There are three assessment items. (B) Weight loss: A weight loss of more than 5 kilograms compared to one year ago. (C) Lower extremity function: Inability to rise from a chair without using hands five times. (D) Meeting two or more criteria qualifies as frailty [37].
ChatGPT	**No answer.** (A), (B), (C), and (D) all correct. Based on the description, individuals meeting two or more of these three assessment criteria are classified as frail. In summary, according to the provided description, there are no errors. In summary, based on the descriptions provided in the data, there is no definitive answer to this question [38].

In this current study, the initial response from ChatGPT and the subsequent answer was totally different. ChatGPT would provide the incorrect answer and information, for example the "Question 4" (Table 4), which indicated "During the fourth stage of labor, it is customary for the healthcare team to administer a vitamin K1 injection to the newborn as a preventive measure against hemorrhagic disorders [38]." The "Question 5" (Table 4) indicated that black stools are typically caused by upper gastrointestinal bleeding, and rectal bleeding, iron supplementation, and overconsumption of pork liver generally do not result in black stools. All of explanations were incorrect based on the concept and nursing knowledge.

Table 4 The ChatGPT prompts and answers on the RNLE (Example 2)

Question 4	Ms. Chen, who has been in labor for 12 hours, has given birth to a healthy newborn. Regarding the nursing interventions performed by the nurse during the fourth stage of labor, which of the following is not appropriate? **Answer: (D)** (A) Encouraging the mother to have eye-to-eye contact with the baby. (B) Wrapping the baby in a blanket to keep warm and encouraging touching the baby. (C) Delaying the administration of vitamin K1 and initiating skin-to-skin contact first. (D) Placing the newborn in an incubator to allow the mother to rest [37].
ChatGPT	**The answer is (C).** During the fourth stage of labor, it is customary for the healthcare team to administer a vitamin K1 injection to the newborn as a preventive measure against hemorrhagic disorders. Placing the newborn in an incubator to allow the mother to rest (D) is a reasonable arrangement to ensure that the mother has an opportunity to rest and recover postpartum [38].
Question 5	Which of the following situations is less likely to result in black stools in a patient? **Answer: (D)** (A) Duodenal ulcer (B) Rectal bleeding (C) Iron supplementation (D) Overconsumption of pork liver (C) Iron supplementation [37].
ChatGPT	**The answer is (C).** black stools are typically caused by upper gastrointestinal bleeding, and rectal bleeding, iron supplementation, and overconsumption of pork liver generally do not result in black stools. Food undergoes normal digestion in the stomach and intestines, which does not result in black stools [38].

Table 5 The ChatGPT prompts and answers on the RSLE (example 3)

Question 6	Which of the following statements about home-based care services is less appropriate? **The answer is (D).** (A) Nursing personnel can independently establish home nursing facilities. (B) Home care can provide care for nasogastric tubes, tracheostomies, and urinary catheters. (C) Home rehabilitation services may be provided when deemed necessary based on an assessment by care management specialists. (D) Home pharmaceutical services are not within the scope of services [37].
ChatGPT-1	**The answer is (A).** Nursing personnel who wish to independently establish home care facilities must adhere to specific laws and regulations, typically requiring certification and licensing by government authorities. Nursing personnel cannot arbitrarily establish home care facilities but must instead comply with relevant laws and regulations, subject to regulatory oversight [38].
ChatGPT-2	**The answer is (D).** Home-based care services can include medication management and assistance, so stating that home pharmaceutical services are not within the scope of services is less appropriate [38].

5. Discussion

In this study, we provide new and impressed evidence about performance of ChatGPT on RNLE. The finding can be organized two parts: (1) the performance of ChatGPT on RNLE and compared to

pass rate of ChatGPT with nursing graduates; and (2) to evaluate the limitations and advantage of using ChatGPT.

5.1. The performance of ChatGPT on RNLE

The present study employed a multiple-choice test. ChatGPT could respond promptly within seconds while also providing answers and information during the interaction, which was same as the study of Ahn (2023) [9] and Ali et al. (2023) [17]. Results showed that ChatGPT had a basic level of nursing knowledge, corresponding to a success rate over 54.05%. However, to compare the four exams, which displayed that ChatGPT had lower ratio of pass percentage, and mean scores was lower than 50 in the exam of 2022 2nd and 2023 1st. A total average score was 60, if any subject receives a score of zero, it will not qualify for a passing grade [34,35].

ChatGPT obtained a total average score of 63.75±5.22 in exam of 2022 1st and 60.75±6.88 in exam 2023 2nd, which all passed the RNLE. The higher average score of candidates in Taiwan might be due to their prior learning of related knowledge before test, as Huh (2023) [23] mentioned. In study by Huh (2023) [23] indicated that ChatGPT correctly answered 60.8%, which was lower than medical students' 90.8%. ChatGPT was able to achieve at or near the passing threshold of 60% accuracy in study by Kung et al. (2023) [28], achieving a level of a third-year medical student, or a first-year resident [1,27]. Furthermore, this current study scores of BMS around 48~78, which was similar to study by Talan & Kalinkara (2023) [2], ChatGPT exhibited exceptional performance on the anatomy exams.

The pass rate for the examination of RNLE in 2022 and 2023 was relatively low, standing at 18.17% and 20.42%, possibly related to students' graduation timing and their level of preparation for re-examination based on Taiwanese culture. RNLE candidate includes those who graduated in the same year as the exam and those who failed the previous examination. Due to students graduating in June every year and subsequently participating in the second RNLE, the number of candidates for the first exam in 2022 and 2023 was noticeably lower than that for the second examination.

5.2. To evaluate the limitations and advantage of using ChatGPT

ChatGPT could answer the most questions directly. As the study by Kung, et al. (2023) [28] indicated that ChatGPT exhibits a remarkable level of concordance and insight in its explanations. ChatGPT is capable of delivering responses within a short timeframe [2]. In this current study, ChatGPT might provide wrong explanations or information, which was corresponded to study by Huh (2023) [23], ChatGPT's medical knowledge and interpretation in the field of medicine unreliable were still inadequate. Ahn (2023) [9] indicated that ChatGPT could search for incorrect knowledge and skills via web surfing, which might be due to inaccurate information and unreliable sources. Based on the findings of this study, ChatGPT had the potential to generate misleading or inaccurate explanations, although it would rephrase the question clearly and accurately when the question is ambiguous or incomprehensible [2]. As the exam progresses, it could lead to hallucinations, especially when it becomes difficult to ascertain the accuracy of certain content produced by ChatGPT [15,24].

The initial response from ChatGPT and the subsequent answers was inconsistent in this study. ChatGPT may confusing or misunderstand about complicated scenarios, which would lead to inconsistent explanations and answers from ChatGPT. It might be due to restricting knowledge before 2021 [11,15]. Moreover, the model's generation diversity, sensitivity to changes in contextual information, and the inherent uncertainty in the machine learning process [11]. ChatGPT still unable to provide corresponding measures and nursing plans based on patient needs [2,29]. ChatGPT did not allow processing of questions based on images [17], which did not be verify in this study. Although AI could simulate human behavior in context, it still could not replace human creativity, critical thinking, and clinical reasoning. Additionally, Sallam (2023) [11] argued excluding different language, such as English records, which might induce selection bias. In the current study, the items of RNLE were to prompt to ChatGPT in Chinese, which might have certain limitations.

6. Limitation

In the current study, researchers only evaluated the performance of ChatGPT using three RNLE assessments conducted after 2021. Additionally, the items used to assess ChatGPT were in Chinese, which could potentially introduce selection bias. These factors may impose certain limitations on the study. It could be displayed some different responses between ChatGPT 3.5 and 4.0.

7. Conclusions

This current study was to understand the performance of ChatGPT on RNLE. ChatGPT displayed a high level of concordance and insight in exams. ChatGPT may have potentially to assist with nursing education. Although ChatGPT provides information and knowledge in seconds as this current study had showed, it cannot replace the essential elements of nursing care, such as clinical reasoning and clinical judgment, especially in complex scenarios. The findings suggest that ChatGPT's performance on the RNLE has the potential to enhance knowledge and analysis. However, further investigation is needed to assess its effectiveness in promoting comprehension, application, synthesis, and evaluation, as per Bloom's taxonomy. Using ChatGPT for nursing students to practice or explain nursing care knowledge should be approached with caution, particularly due to hallucination misleading the answer, or regarding issues like plagiarism in article writing and homework assignments. We believe ChatGPT would reach a maturity level and impact clinical nursing care. To anticipate and prepare for the future of nursing education, it is essential to thoroughly assess both the potential benefits and risks associated with ChatGPT. In the future, it is recommended to integrate ChatGPT into different nursing courses, to assess its limitations and effectiveness by a variety of tools and methods, and to evaluate the response in different clinical situations and nursing care.

References

1.   Gilson A, Safranek CW, Huang T, Socrates V, Chi L, Taylor RA, Chartash D. How Does ChatGPT Perform on the United States Medical Licensing Examination? The Implications of Large Language Models for Medical Education and Knowledge Assessment. JMIR Med Educ. 2023 Feb 8,9, e45312. https://doi.org/10.2196/45312

2.   Talan T., & Kalinkara Y. The role of artificial intelligence in higher education: ChatGPT assessment for anatomy course. J Manag Inf Syst Comput Sci. 2023, 7(1), 33-40. https://doi.org/10.33461/uybisbbd.1244777

3.   Miao H, Ahn H. Impact of ChatGPT on Interdisciplinary Nursing Education and Research. Asian Pac Isl Nurs J. 2023 Apr 24, 7, e48136. https://doi.org/10.2196/48136

4.   O'Connor S., & ChatGPT. Open artificial intelligence platforms in nursing education: Tools for academic progress or abuse? Nurse Educ Pract. 2023 Jan, 66, 103537. https://doi.org/10.1016/j.nepr.2022.103537

5.   Qadir J. Engineering education in the era of ChatGPT: Promise and pitfalls of generative AI for education. TechRxiv. Preprint. 2022. https://doi.org/10.36227/techrxiv.21789434.v1(accessed on 20 Aug. 2023)

6. Grunhut J, Wyatt AT, Marques O. Educating Future Physicians in Artificial Intelligence (AI): An Integrative Review and Proposed Changes. J Med Educ Curric Dev. 2021 Sep 6, 8, 23821205211036836. https://doi.org/10.1177/23821205211036836

7. Sallam M., Salim N., Barakat M. & Al-Tammemi A. ChatGPT applications in medical, dental, pharmacy, and public health education: A descriptive study highlighting the advantages and limitations. Narra J. 2023, 3(1), e103. https://doi.org/10.52225/narra.v3i1.103

8. Sallam M, Salim NA, Al-Tammemi AB, Barakat M, Fayyad D, Hallit S, Harapan H, Hallit R, Mahafzah A. ChatGPT Output Regarding Compulsory Vaccination and COVID-19 Vaccine Conspiracy: A Descriptive Study at the Outset of a Paradigm Shift in Online Search for Information. Cureus. 2023 Feb 15, 15(2), e35029. https://doi.org/10.7759/cureus.35029

9. Ali K, Barhom N, Tamimi F, Duggal M. ChatGPT-A double-edged sword for healthcare education? Implications for assessments of dental students. Eur J Dent Educ. 2023 Aug 7. https://doi.org/10.1111/eje.12937.

10. Khan RA, Jawaid M, Khan AR, Sajjad M. ChatGPT - Reshaping medical education and clinical management. Pak J Med Sci. 2023 Mar-Apr, 39(2), 605-607. https://doi.org/10.12669/pjms.39.2.7653

11. Sallam M. ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns. Healthcare (Basel). 2023 Mar 19, 11(6), 887. https://doi.org/10.3390/healthcare11060887

12. O'Connor S, Yan Y, Thilo FJS, Felzmann H, Dowding D, Lee JJ. Artificial intelligence in nursing and midwifery: A systematic review. J Clin Nurs. 2023 Jul, 32(13-14), 2951-2968. https://doi.org/10.1111/jocn.16478

13. Seney V, Desroches ML, Schuler MS. Using ChatGPT to Teach Enhanced Clinical Judgment in Nursing Education. Nurse Educ. 2023 May-Jun 01, 48(3), 124. https://doi.org/10.1097/NNE.0000000000001383

14. Kikuchi R. Application of artificial intelligence technology in nursing studies: A systematic review. 2021. HIMSS Foundation and the HIMSS Nursing Informatics Community. https://www.himss.org/resources/application-artificial-intelligence-technology-nursing-studies-systematic-review (accessed on 10 Sep. 2023)

15. Cascella M, Montomoli J, Bellini V, Bignami E. Evaluating the Feasibility of ChatGPT in Healthcare: An Analysis of Multiple Clinical and Research Scenarios. J Med Syst. 2023 Mar 4, 47(1), 33. https://doi.org/10.1007/s10916-023-01925-4

16. OpenAI. GPT-4 Technical Report. 2023. https://cdn.openai.com/papers/gpt-4.pdf

17. Ahn C. Exploring ChatGPT for information of cardiopulmonary resuscitation. Resuscitation. 2023 Apr, 185, 109729. https://doi.org/10.1016/j.resuscitation.2023.109729

18. Bhutoria A. Personalized education and artificial intelligence in the United States, China, and India: A systematic review using a Human-In-The-Loop model. Comput Educ: Artif Intell. 2022; 3:100068 https://doi.org/10.1016/j.caeai.2022.100068

19. Hutson M. Could AI help you to write your next paper? Nature. 2022 Nov, 611(7934), 192-193. https://doi.org/10.1038/d41586-022-03479-w

20. O'Connor S. Artificial intelligence and predictive analytics in nursing education. Nurse Educ Pract. 2021 Oct, 56, 103224. https://doi.org/10.1016/j.nepr.2021.103224

21. Else H. Abstracts written by ChatGPT fool scientists. Nature. 2023 Jan, 613(7944), 423. https://doi.org/10.1038/d41586-023-00056-7

22. Benoit J. ChatGPT for clinical vignette generation, revision, and evaluation. medRxiv - Medical Education. 2023 Feb 08. https://doi.org/10.1101/2023.02.04.23285478 (accessed on 5 Sep. 2023)

23. Huh S. Are ChatGPT's knowledge and interpretation ability comparable to those of medical students in Korea for taking a parasitology examination?: a descriptive study. J Educ Eval Health Prof. 2023, 20, 1. https://doi.org/10.3352/jeehp.2023.20.1

24. Eysenbach G. The Role of ChatGPT, Generative Language Models, and Artificial Intelligence in Medical Education: A Conversation with ChatGPT and a Call for Papers. JMIR Med Educ 2023, 9, e46885. https://doi.org/10.2196/46885

25. Wikipedia. Hallucination (artificial intelligence). 2023. https://en.wikipedia.org/wiki/Hallucination_(artificial_intelligence)

26. Li'evin V., Hother C.E., & Winther O. Can large language models reason about medical questions? arXiv - CS - Artif Intell. 2022. https://arxiv.org/pdf/2207.08143.pdf

27. Antaki F, Touma S, Milad D, El-Khoury J, Duval R. Evaluating the Performance of ChatGPT in Ophthalmology: An Analysis of Its Successes and Shortcomings. Ophthalmol Sci. 2023 May 5, 3(4), 100324. https://doi.org/10.1016/j.xops.2023.100324

28. Kung TH, Cheatham M, Medenilla A, Sillos C, De Leon L, Elepaño C, Madriaga M, Aggabao R, Diaz-Candido G, Maningo J, Tseng V. Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. PLOS Digit Health. 2023 Feb 9, 2(2), e0000198. https://doi.org/10.1371/journal.pdig.0000198

29. Gunawan J. Exploring the future of nursing: Insights from the ChatGPT model. Belitung Nurs J. 2023 Feb 12, 9(1), 1-5. https://doi.org/10.33546/bnj.2551.eCollection 2023

30. Mahmud MM, Yaacob Y, Ramachandiran CR, Wong SC, Ismail O. Theories into Practices: Bloom's Taxonomy, Comprehensive Learning Theories (CLT) and E-Assessments. In Proceeding Book of 1st International Conference on Educational Assessment and Policy. Vol. 2. Center for Educational Assessment. 2019. 22-27p. https://doi.org/10.26499/iceap.v2i1.91

31. Finlay GD, Rothman MJ, Smith RA. Measuring the modified early warning score and the Rothman index: advantages of utilizing the electronic medical record in an early warning system. J Hosp Med. 2014 Feb, 9(2), 116-9. https://doi.org/10.1002/jhm.2132

32. Robert N. How artificial intelligence is changing nursing. Nurs Manage. 2019 Sep, 50(9), 30-39. https://doi.org/10.1097/01.NUMA.0000578988.56622.21

33. The Rothman Index. Model development and scientific validation.2023. www.perahealth.com/the-rothman-index/model-development-and-scientific-validation.

34. Ministry of Examination. Regulations on Professional and Technical Examinations.https://wwwc.moex.gov.tw/main/ExamLaws/wfrmExamLaws.aspx?kind=3&menu_id=320&laws_id=110 (accessed on 11 Sep. 2023)

35. Ministry of Examination. Statistics of the Number of Applicants and Acceptance/Pass Rates for the Second Professional and Technical Examinations for Nursing License in the Year. https://wwwc.moex.gov.tw/main/ExamReport/wHandStatisticsFile.ashx?file_id=2122 (accessed on 11 Sep. 2023)

36. Cotton D., Cotton P. & Shipway R. Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. Innov Educ and Teach Int. 2023 Mar. https://doi.org/10.1080/14703297.2023.2190148

37. Ministry of Examination. Exam question query platform. 2023. https://wwwq.moex.gov.tw/exam/wFrmExamQandASearch.aspx (accessed on 11 Sep. 2023)

38. OpenAI. ChatGPT (Mar 14 version) [Large language model]. 2023. https://chat.openai.com/chat (accessed on 02 Sep. 2023)