



This is an Accepted Manuscript version of the article published originally by Elsevier accepted for publication in the journal:

*International Journal of Nursing Studies*

This version may differ from the original in pagination and typographic details. When using, please cite the original.

AUTHOR(S)

Hobensack, M., von Gerich, H., Vyas, P., Withall, J., Peltonen, L.-M., Block, L. J., Davies, S., Chan, R., Van Bulck, L., Cho, H., Paquin, R., Mitchell, J., Topaz, M., & Song, J.

TITLE

A rapid review on current and potential uses of large language models in nursing

YEAR

2024

DOI

10.1016/j.ijnurstu.2024.104753

CITATION

Hobensack, M., von Gerich, H., Vyas, P., Withall, J., Peltonen, L.-M., Block, L. J., Davies, S., Chan, R., Van Bulck, L., Cho, H., Paquin, R., Mitchell, J., Topaz, M., & Song, J. (2024). A rapid review on current and potential uses of large language models in nursing. *International Journal of Nursing Studies*, 154, 104753.

<https://doi.org/10.1016/j.ijnurstu.2024.104753>

VERSION

Accepted Manuscript

LICENSE

© 2024 This version is published under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives (CC BY-NC-ND) License, which permits use and distribution in any medium, provided the original work is properly cited, and no modifications or adaptations are made. <https://creativecommons.org/licenses/by-nc-nd/4.0/>

**Title:** A Rapid Review on Current and Potential Uses of Large Language Models in Nursing

**Mollie Hobensack [Corresponding Author]**

Clinical Program Manager – Quality & Informatics  
Department of Geriatrics and Palliative Care  
Icahn School of Medicine at Mount Sinai  
One Gustave L. Levy Place Box 1070  
New York City, NY, 10029, USA  
mx.hobensack@gmail.com

**Hanna Von Gerich**

Department of Nursing Science, University of Turku, Finland

**Pankaj Vyas**

College of Nursing, University of Arizona  
Tucson, AZ, USA

**Jennifer Withall**

Department of Biomedical Informatics, Columbia University  
New York, NY, USA

**Laura-Maria Peltonen,**

Department of Nursing Science, University of Turku and Research Services  
Turku University Hospital, Finland

**Lorraine J. Block**

School of Nursing, University of British Columbia  
Vancouver, Canada

**Shauna Davies**

Faculty of Nursing, University of Regina  
Regina, Canada

**Ryan Chan**

Arthur Labatt Family School of Nursing, Western University  
London, ON, Canada

**Liesbet Van Bulck**

Department of Public Health and Primary Care, KU Leuven – University of Leuven  
Leuven, Belgium

**Hwayoung Cho**

College of Nursing, University of Florida  
Gainesville, FL, USA

**Robert Paquin**

Faculty of Nursing, Midwifery, and Palliative Care, King's College London  
London, UK

**James Mitchell**

Department of Biomedical Informatics, University of Colorado School of Medicine  
Denver, CO, USA

**Maxim Topaz**

Columbia University School of Nursing  
Data Science Institute, Columbia University  
VNS Health  
New York, NY, USA

**Jiyoun Song**

Department of Biobehavioral Health Sciences, University of Pennsylvania School of Nursing  
Philadelphia, Pennsylvania, USA

**[Abstract]** 358/400

**[Word Count]** 3,369/7,000

**[Tables]** 1

**[Figures]** 2

**[Supplementary Materials]** 2

**[References]** 63

**[Keywords]** Rapid Review, Nursing Informatics, Large Language Models, Generative AI,  
ChatGPT

**What is already known:**

- The applications of large language models are growing at an exponential rate in the health sector, and this is leading to evolving perspectives on how they will impact health and care.
- Existing reviews on the use of large language models in healthcare have only focused on the medical perspective, leaving out the nursing perspective.

**What this paper adds:**

- This is the first rapid review to summarize the current and potential uses of large language models in nursing.
- Most of the current literature is editorials or commentaries; future studies need to focus on conducting empirical and qualitative research to understand how to utilize large language models to improve nursing practice, education, or research.

**Abstract**

**Background:** The application of large language models across commercial and consumer contexts has grown exponentially in recent years. However, a gap exists in the literature on how large language models can support nursing practice, education, and research. This study aimed to synthesize the existing literature on current and potential uses of large language models across the nursing profession.

**Methods:** A rapid review of the literature, guided by Cochrane rapid review methodology and PRISMA reporting standards, was conducted. An expert health librarian assisted in developing broad inclusion criteria to account for the emerging nature of literature related to large language models. Three electronic databases (i.e., PubMed, CINAHL, and Embase) were searched to identify relevant literature in August 2023. Articles that discussed the development, use, and application of large language models within nursing were included for analysis.

**Results:** The literature search identified a total of 2,028 articles that met the inclusion criteria. After systematically reviewing abstracts, titles, and full texts, 30 articles were included in the final analysis. Nearly all (93%; n = 28) of the included articles used ChatGPT as an example, and subsequently discussed the use and value of large language models in nursing education (47%; n = 14), clinical practice (40%; n = 12), and research (10%; n = 3). While the most common assessment of large language models was conducted by human evaluation (26.7%; n

= 8), this analysis also identified common limitations of large language models in nursing, including lack of systematic evaluation, as well as other ethical and legal considerations.

**Discussion:** This is the first review to summarize contemporary literature on current and potential uses of large language models in nursing practice, education, and research. Although there are significant opportunities to apply large language models, the use and adoption of these models within nursing has elicited a series of challenges, such as ethical issues related to bias, misuse, and plagiarism.

**Conclusion:** Given the relative novelty of large language models, ongoing efforts to develop and implement meaningful assessments, evaluations, standards, and guidelines for applying large language models in nursing are recommended to ensure appropriate, accurate, and safe use. Future research along with clinical and educational partnerships are needed to enhance understanding and application of large language models in nursing and healthcare.

## Introduction

The applications of large language models are growing at an exponential rate, and this is leading to evolving perspectives on how it will revolutionize science<sup>1</sup>, education<sup>2</sup>, and healthcare<sup>3</sup>. Large language models are a type of generative artificial intelligence designed to collect, process, translate, and respond to human language<sup>4</sup>. They function by using advanced computational methods (e.g., neural network architecture) to learn from large volumes of narrative data using natural language processing. While the building blocks of large language models have previously existed, the release of ChatGPT in November 2022 marked a pivotal transition in which large language models became accessible to the public, allowing widespread usability and experimentation in everyday applications<sup>4,5</sup>.

There is great excitement about using large language models in healthcare<sup>4,6</sup>. Several areas have been identified where large language models are poised to positively impact clinicians, such as facilitating patient communication, off-loading administrative tasks, and improving clinical decision-making<sup>7</sup>. However, integrating large language models into healthcare applications has also brought trepidation related to misinformation, privacy, and biases<sup>4</sup>. There is no shortage of individual perspectives on how large language models will shape healthcare;

however, the plethora of literature has made it challenging to synthesize and summarize consensus in the rapidly emerging literature.

Few review studies have sought to summarize the literature, and among these studies they primarily focused on the broader healthcare context including feasibility and application of large language models<sup>4-6</sup>. Prior research has described the ever-growing use of artificial intelligence technologies within the nursing profession<sup>8</sup> and the value of the nursing perspective to improve innovation, quality care, and health equity<sup>9</sup>. There are many perspectives on how large language models can support nurses, such as creating nursing care plans, synthesizing nursing research articles, and creating individualized educational learning pathways within nursing schools. However, no prior study has summarized the overarching themes of how large language models could best serve nursing practice, education, and research and what risks are associated with the applications of such technologies. Therefore, this work aimed to synthesize the existing literature on current and potential uses of large language models across the nursing profession to assess this gap.

## **Methods**

### *Study criteria and synthesis*

Rapid reviews are defined as a form of synthesis that accelerates the traditional systematic review to help disseminate literature in a resource-efficient manner amidst the rapidly evolving literature<sup>10</sup>. This rapid review aimed to explore and summarize the current literature on using large language models in nursing. This study was guided by the Cochrane Rapid Review guidelines<sup>10</sup> and an a-priori protocol<sup>11</sup> published on September 12, 2023 in Open Science Framework (<https://osf.io/n8xf2/>). In consultation with a health librarian, a search strategy was created for the following three databases in August 2023: PubMed, CINAHL, and Embase. The search strategy for each database is available in **Supplemental Material A**. Concepts of interest in the search query included terms related to nursing and large language models.

Given the recent emergence of large language models, the criteria for inclusion were broad. All article types such as conference proceedings, grey literature, and editorials were included. There was no restriction on publication date, but articles did need to be published in English. Full inclusion and exclusion criteria are included in **Supplemental Material B**. The screening process was facilitated using Covidence (Veritas Health Innovation, Melbourne, Australia) and visualized in a Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram<sup>12</sup>. Each title and abstract were independently reviewed by at least two co-authors and cross-checked for screening. Conflicts were resolved by a third co-author not involved in the original screening (MH).

A structured data extraction form was developed based on the format recommended by the *Cochrane Handbook for Systematic Reviews of Interventions*<sup>13</sup> and from author consensus on relevant information to extract. Once the review team finalized the data extraction form, it was set up in Covidence, an online software tool designed to facilitate the process of conducting a systematic review (<https://www.covidence.org>). Data extraction was organized into the following categories: 1) Study characteristics, 2) Large language model characteristics, and 3) Application of large language model in nursing. Two independent reviewers extracted data from each study and entered it into Covidence. A third reviewer resolved the differences. The extracted data pertaining to study characteristics encompassed the authors' affiliated country, their names,

publication year, study type (e.g., commentary, research article), study design (if applicable), and the article's aim. Similarly, data pertaining to the large language model characteristics comprised its type, description, the type, and quality of data used for training, and the tools utilized for model evaluation (if applicable). Lastly, the data concerning the application of the large language model in nursing involved the audience and setting of the application, the methodology of its application, and the study's limitations.

## Results

An overview of the review process is visualized in **Figure 1**. The initial search yielded 2,028 articles. Following a series of iterative screenings, 30 studies were eligible for extraction and final evaluation.

### *Study characteristics*

The characteristics of the included studies are described in **Table 1**. Most articles were written by authors from North America (33.3%; n = 10)<sup>14–23</sup> and Asia (33.3%; n = 10)<sup>24–33</sup>. The remaining articles were authored in Europe (26.7%; n = 8)<sup>34–41</sup>, Oceania (3.3%; n = 1)<sup>42</sup>, and South America (3.3%; n = 1)<sup>43</sup>. Two articles included ChatGPT as a co-author<sup>37,41</sup>. All articles were written in 2023, except for one article from 2012<sup>35</sup> and another from 2022<sup>38</sup>. The majority of the articles (57%; n = 17) were editorials or letters to the editor<sup>16,17,19–21,23,24,26,28–30,34,36,37,39,42,43</sup>. Seven articles were research articles that reported on the results of an original study<sup>25,31,33,35,38,40,41</sup>, of which only one was a qualitative study<sup>25</sup>. Regarding the main aims, many articles (n = 14) focused on describing or investigating the value of large language models for nursing education<sup>14–16,20,22,23,27,31–34,37,42,43</sup>, such as the performance of ChatGPT in passing a nursing exam. A total of 12 articles described or investigated the use of large language models in nursing practice<sup>14,21,25,26,28,29,35,36,38–41</sup>, and three articles focused on potential applications for nursing research<sup>17,19,36</sup>.

### *Large Language Model Characteristics*

Nearly all (n = 28; 93%) studies discussed the large language model, ChatGPT, within the nursing domain. ChatGPT was described as a generative artificial intelligence-based predictive chatbot utilizing deep learning<sup>14,25–27,32–34</sup> developed using reinforcement learning techniques<sup>37</sup>. The described functionalities of ChatGPT included the generation of natural language text<sup>14,40,42</sup> mimicking human-like responses and carrying on conversations with users<sup>15,17,19,24–26,28,34,36,37,41</sup> with the ability to analyze and understand user inputs<sup>16,24,25,37,39</sup> and create images<sup>18</sup>.

Generally, discussions regarding the training data of ChatGPT described it as a compilation of textual data<sup>15,26,34,36</sup> derived from the internet<sup>16,18,28</sup>. However, the content or source of the training data, along with the quality of the data, were not further elaborated. The first of the two non-ChatGPT large language models referenced in this review included a natural language generation system named BT-Nurse, described as a system used in generating summaries from electronic health record data<sup>35</sup>. The second described their large language models as combining long short-term memory neural networks with local model-agnostic explanations used in keyword extraction<sup>38</sup>. Both large language models utilized hospital data without further description on the data quality.

Among the studies that detailed their evaluation process, the focus was primarily on user satisfaction and trust<sup>25,35,36,40,43</sup>, performance assessment<sup>31,34,35,38,40,41</sup>, and model explainability<sup>38</sup>. Nearly all articles that contained an assessment of ChatGPT were conducted by

human evaluations<sup>23,25,31,34,36,40,41,43</sup>. The one exception used ChatGPT to provide an assessment of itself<sup>17</sup>.

### *Application of the Large Language Model in Nursing*

The included studies collectively illustrate the diversity and expanding role of large language models in nursing, ranging from educational advancements to clinical support and research applications. One article discussed the potential benefits of using large language models to conduct literature reviews<sup>16</sup>. By automating the creation of literature, large language models could reduce screening workload and help facilitate faster dissemination of information<sup>17,35</sup>. Other articles discussed utilizing large language models such as ChatGPT to assist nurses by offloading administrative tasks and improving communication. Examples include using large language models in roles such as customer service agents and assisting with appointment scheduling and management<sup>37</sup>.

Over half of the studies discussed the impact large language models could have on nursing practice (53%; n = 16)<sup>14-17,22,24,25,28-30,32,35,37,39-41</sup>, particularly as it relates to patient care and education<sup>24</sup>. For example, ChatGPT could aid in patient education through translating medical terminology into lay language, thereby improving communication with patients<sup>26,36</sup>. In clinical practice, these models could improve patient outcomes by automating the writing of medical summaries and clinical notes<sup>17</sup>. Large language models could also extend their utility in diagnostic assistance, enhancing efficiency, personalizing care, and providing multilingual support, thus contributing to patient care and nursing education<sup>24,25</sup>. Specifically, ChatGPT could support patient engagement by providing reliable information about medical conditions and treatments, thereby improving decision-making<sup>17,25,29,30,40</sup> and generating nursing care plans<sup>41</sup>. Academic practice was another area where large language models were discussed (n = 15) as an educational aid<sup>14-16,18,20,23,24,26,27,31,34,36,37,42,43</sup>. In the educational setting, large language models could provide personalized learning experiences, offering tutoring and homework help to students, thereby enhancing their understanding of complex concepts<sup>37</sup>. Additionally, they possess the capacity to enhance curriculum integration and development<sup>16</sup>, by facilitating the creation of educational content, curriculum blocks, and case studies<sup>16,42</sup>, thereby contributing to the development of critical thinking skills<sup>15,26,36</sup>. One study even demonstrated the efficacy of a large language model in completing the National Licensing Examination for nurses<sup>31</sup>. Importantly, students' ethical use of ChatGPT, particularly concerning academic integrity and the responsibilities of nursing students and faculty, is still under scrutiny<sup>18,27,34</sup>.<sup>20</sup>

### *Limitations of Using Large Language Models*

The included studies reported several concerns and challenges regarding the adoption of large language models in nursing. Ethical issues such as bias, plagiarism, data privacy, security risks, and legal concerns are prominent<sup>14,16,17,24,37</sup>. The accuracy of information provided by large language models is a critical issue, with risks of inaccurate or incorrect information, and a lack of originality<sup>14,17,25</sup>. Additionally, there is apprehension about the erosion of critical thinking skills, over-reliance on technology, and the inability of artificial intelligence to replicate human empathy and clinical judgment, which are central to nursing<sup>14,15,26</sup>. Barriers to effective integration arise from the necessity for educators to differentiate between the work produced by large language models and that of students. Anticipated challenges include ensuring the accuracy and originality of content generated by artificial intelligence, along with concerns about preserving academic integrity<sup>16,17,42</sup>. Therefore, while large language models present significant advantages

in nursing education and practice, they also necessitate a balanced approach to harness their benefits effectively while mitigating potential risks and ethical dilemmas.

## **Discussion**

This review aimed to synthesize the existing literature on current and potential uses of large language models across the nursing profession. It is unsurprising that the results indicated a recent surge in literature in 2023, coinciding with the release of ChatGPT in November 2022. This release introduced an artificial intelligence tool powered by a large language model that was comprehensible, valuable, and accessible to nurses, including those without specialized knowledge in informatics<sup>44</sup>. Artificial intelligence technologies that are easily accessible and user friendly such as ChatGPT, enable nurses at all levels to make use of the newest artificial intelligence technologies to create innovative solutions to improve workflow and care. While ChatGPT was by far the most frequently cited large language models in this review literature, stakeholders must adopt a forward-thinking approach to formulate policies that are adaptable to the anticipated proliferation of future large language models and generative artificial intelligence tools within nursing. The expansion of literature and policies surrounding the ethical use of large language models in healthcare are rapidly emerging, describing various perspectives and use cases<sup>3,45,46</sup>. Among other review studies conducted, which tended to focus on healthcare broadly<sup>6,7</sup> or medicine<sup>3-5</sup>, this review is the first to collate the literature describing the consideration and utilization of large language models from the nursing perspective.

Findings from this study highlight that large language models were primarily evaluated and addressed using human evaluation (e.g., satisfaction, trust, explainability)<sup>47-49</sup>. While the performance assessment of large language models by human experts using professional benchmarks is important, quantitative measures should be developed, applied, and described in evaluating large language model outputs. For instance, evaluating and validating the content, quality, and sources of training data within large language models through rigorous methods can help to identify potential biases in the training data. It is crucial to demonstrate the data quality and evaluation in specific contexts to ensure the output of large language models is appropriate and accurate. Thus, future studies are needed to explore the standardization of evaluation metrics and guidelines that rigorously assess and transparently report on the functioning and bias of the large language model training data, especially as it pertains to ChatGPT and the release of alternative large language models.

Guidelines on best practices and responsible use of large language models in healthcare have struggled to keep pace with the rapidly expanding adoption of these technologies. The use of non-transparent, "black box" systems in patient care poses significant challenges. Healthcare clinicians are required to understand the underlying workings of each system, its recommendations, and how to manage patient care safely, ethically, and responsibly, adding to their existing responsibilities<sup>50</sup>. Developing instructions for clinicians on what large language models are intended to do and not do is vital as large language models are introduced into clinical practice. Thus, nursing partnerships are essential to ensure accountability, patient safety, and trust in large language models regarding healthcare decisions and how to critically assess the validity and accuracy of each large language model's output.

Healthcare organizational leaders are being asked to inform the development of policy to guide large institutions on the safe and appropriate use of large language models regarding patient care. For example, some experts provided guidance that includes being discreet and not

inputting specific patient information into the generative artificial intelligence software, ensuring there are processes in place to verify the data that the large language models has produced, watching for biases, and being critical of the information presented (e.g., is it outdated?)<sup>51</sup>. Nurses must be aware of the ethical, privacy, and legal implications of using large language model in clinical decision-making and care delivery. Questions should be readily discussed among nurses, such as what happens if a clinical decision, informed by ChatGPT, turns out to contribute to a treatment error? How does one mitigate the risk of generative artificial intelligence patient-error and prevent technology-mediated adverse events? Who is liable for decisions guided by large language models? What should be documented about the use of large language models in clinical practice?

While ChatGPT's inception was in the United States (US), this study highlights the global impact large language models have had on healthcare research and policy. Canada Health Infoway has recently summarized key themes to guide healthcare leaders through questions of artificial intelligence use and implications, highlighting the ethical, privacy and legal considerations, opportunities, and responsibilities<sup>52</sup>. These themes include defining what is considered an automated decision-making system; the need for explainability, which is done in such a way that it is understandable for the intended audience; the need for transparency, which includes disclosure when artificial intelligence systems are used in diagnostic support; the right to contest decisions made by artificial intelligence, which has important implications in healthcare; and the need for the de-identification of data and preserving the privacy and dignity of people. These themes resonate with what was found in this study, noting alignment with several authors<sup>14,16,17,24,25,37</sup> as well as with other researchers calling for explainable and responsible artificial intelligence in related fields, such as engineering, mathematics, and computer science<sup>53,54</sup>.

This alignment also extends to broader policy discussions, as organizations in Canada, the US, and the European Union have been rapidly working on. For example, in the US, Senators have introduced the Algorithmic Accountability Act<sup>55</sup>, which would outline the requirements for transparency in automated decision systems that use computation as the basis for judgments and decisions. If implemented, this could mean that nurses would have software systems, such as electronic health records, expected to display (or make accessible) the precise data points and processes that led to the automated clinical decision support action or recommendation. In Great Britain, large language models are regulated as medical devices by the Medical Device Regulation 2002<sup>56,57</sup>. The European Union will regulate large language models based on the Artificial Intelligence Act which assigns "different rules for different risk levels"<sup>58</sup>. As large language models continue to evolve, future studies are encouraged to examine regulations and policies nationally and internationally to assess how transparency is incorporated into each regulation.

Nurses are experiencing high burden levels, leading to high turnover rates and workforce shortages<sup>59</sup>. Nursing leaderships should be cognizant on how the integration of additional artificial intelligence tools could increase burden on nurses (e.g., additional responsibility to understand large language model) if not intentionally implemented. This review generated many suggestions on how large language models can help offload burdens related to documentation (e.g., automating health data summaries) and administrative tasks (e.g., scheduling appointments). In addition, large language models have the potential to create tailored patient

educational materials aligned with a patient's health literacy. Given the lack of qualitative research identified in this review, future qualitative research can be utilized to thoroughly explore and identify target areas to implement large language models to help optimize workflows and reduce workforce burden for nurses. Exploring the use of qualitative research would be ideal for increasing understanding of how to use large language models in clinical practice by exploring salient factors (e.g., barriers and facilitators) with respect to specific aspects of large language models across nursing at the practice, research, and education. Without hesitation, this type of technology can be anticipated to actively change and impact the nursing profession. While there are many benefits to large language models in aiding in efficient dissemination and creating tailored personalized learning plans and case studies, evidence of these impacts presented in this paper also raises important challenges to consider, such as academic integrity and plagiarism, patient care, ethics, and critical thinking<sup>16,17,42</sup>. In the US, the National Academies of Science, Engineering, and Medicine held a symposium to discuss how artificial intelligence will impact the education of health professionals<sup>60</sup>. Conversations within this symposium surfaced the importance of educators strategically identifying the purpose of an assessment or assignment and if artificial intelligence aids or hinders a learner's ability to gain a specific competency (e.g., critical thinking, recall). As of today, many universities that teach nursing curricula have responded to the use of large language models with policies that seek to guide and support innovation while also allowing instructors to make decisions regarding its permitted or non-permitted use<sup>61</sup>. Nursing students and educators had to quickly adapt and learn new methods of seeking and utilizing information with large language models such as ChatGPT. Some publishers have also established guidelines<sup>62</sup> on how to acknowledge the use of a large language models in their writing process, clarifying that such acknowledgement won't affect acceptance<sup>63,64</sup>. The ethical guidelines for the utilization and citation of large language models are constantly evolving, prompting new discussions on whether these models should be considered as co-authors or merely as tools in academic work<sup>37,62</sup>.

Given the fast-growing body of literature, studies published after August 2023 or outside the three databases may have been missed in this review. This study also recognizes that the pace of international policies and perspectives is rapidly changing and may have evolved following the publication of this review paper. This paper helps to bring to light the importance of considering the unique role that large language models will play in the nursing profession as the conversation and application of large language models continue to transpire.

## **Conclusion**

This rapid review is the first to describe the existing literature on current and potential uses of large language models across the nursing profession. Most included articles were editorials or letters to the editors, indicating the need for more empirical and qualitative research in this domain. Almost all studies described the large language model ChatGPT. Nurses in all areas of practice, research, education, and policy are expected to be impacted by the use and application of large language models. Partnerships with nurses are recommended as large language models are integrated into research, practice, and education. While large language models present significant advantages in the nursing profession, they also necessitate a balanced approach to harness their benefits effectively while mitigating potential risks and ethical dilemmas.

## **Conflicts of Interest**

There are no conflicts of interest to disclose.

### **Declaration of Generative Artificial Intelligence**

During the preparation of this work the author(s) used ChatGPT for editorial support. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

### **Tables**

Table 1. Study characteristics of the included studies

<b>Author(s)</b>	<b>Year Published</b>	<b>Country</b>	<b>Type of record</b>	<b>Design type (if study)</b>	<b>Aim of the manuscript</b>
Abdulai & Hung, 2023	2023	Canada	Commentary	N/A	To describe and reflect on potential benefits, limitations, and risks of using ChatGPT in nursing education and practice
Ahmed, 2023	2023	Iraq	Editorial/Letter to the editor	N/A	To dive into the ramifications of ChatGPT within the nursing domain and accentuate its capacity and constraints to transform the discipline
Alanzi, 2023	2023	Saudi Arabia	Research article	Qualitative study	To investigate the impact of ChatGPT on teleconsultants in managing their operations and services

Allen & Woodnutt, 2023	2023	United Kingdom	Editorial/Letter to the editor	N/A	To discuss if ChatGPT can pass a nursing exam
Berşe et al., 2023	2023	Turkey	Editorial/Letter to the editor	N/A	To examine the potential benefits and limitations of using artificial intelligence (AI) chatbots in nursing practice, with a particular focus on ChatGPT
Castonguay et al., 2023	2023	Canada	Reflection article	N/A	To examine ChatGPT through its technological history, types of use, potential uses in nursing education, and a call to address the professional development need of nursing faculties to better plan and address generative artificial intelligence understanding needs
Choi et al., 2023	2023	China	Opinion, Perspective, commentary	N/A	To discuss application of ChatGPT in nursing education
Draganic, 2023	2023	United States	Editorial/Letter to editor	N/A	To discuss opportunities and challenges in NP education

Fraser, 2023	2023	Canada	Editorial/Letter to editor	N/A	To discuss the use of AI-generated content in nursing literature
Frith, 2023	2023	United States	Opinion, Perspective, Commentary	N/A	To discuss the disruptive characteristics of ChatGPT as an educational technology in nursing
Gunawan, 2023	2023	Indonesia	Editorial/Letter to Editor	N/A	To provide a comprehensive overview of the future of nursing through the lens of the ChatGPT model
Hunter et al., 2012	2012	United Kingdom	Research article, case study	Non-randomized experimental study	To determine whether and how a computer system could automatically generate helpful natural language nursing shift summaries solely from an electronic patient record system, in a neonatal intensive care unit (NICU)
Irwin et al., 2023	2023	Australia	Editorial/Letter to Editor	N/A	To discuss the implications of the age of artificial intelligence for nursing and midwifery practice and education
Lyon, 2023	2023	United States	Editorial/Letter to Editor	N/A	To discuss the potential utility and concerns about Large

					Language Model Chatbots for oncology nursing authors
Miao & Ahn, 2023	2023	United States	Editorial/Letter to Editor	N/A	To explore the background, development, and potential linkages to nursing education by ChatGPT
Moons & Van Bulck, 2023	2023	Belgium	Editorial/Letter to Editor	N/A	To inform readers about ChatGPT and the potential use of ChatGPT for cardiovascular clinic and research
Nashwan & Abujaber, 2023	2023	Qatar	Editorial/Letter to Editor	N/A	To explore the potential applications of artificial intelligence in the nursing process, with a focus on patient data assessment and interpretation, communication with patients and families, identifying gaps in care plans, and ongoing professional development
O'Connor & ChatGPT, 2023	2023	United Kingdom	Editorial/Letter to Editor	N/A	To provide the pros and cons of using artificial intelligence in nursing education
Odom-Forren, 2023	2023	United States	Editorial/Letter to Editor	N/A	To provide the pros and cons of artificial intelligence for perianesthesia nurses

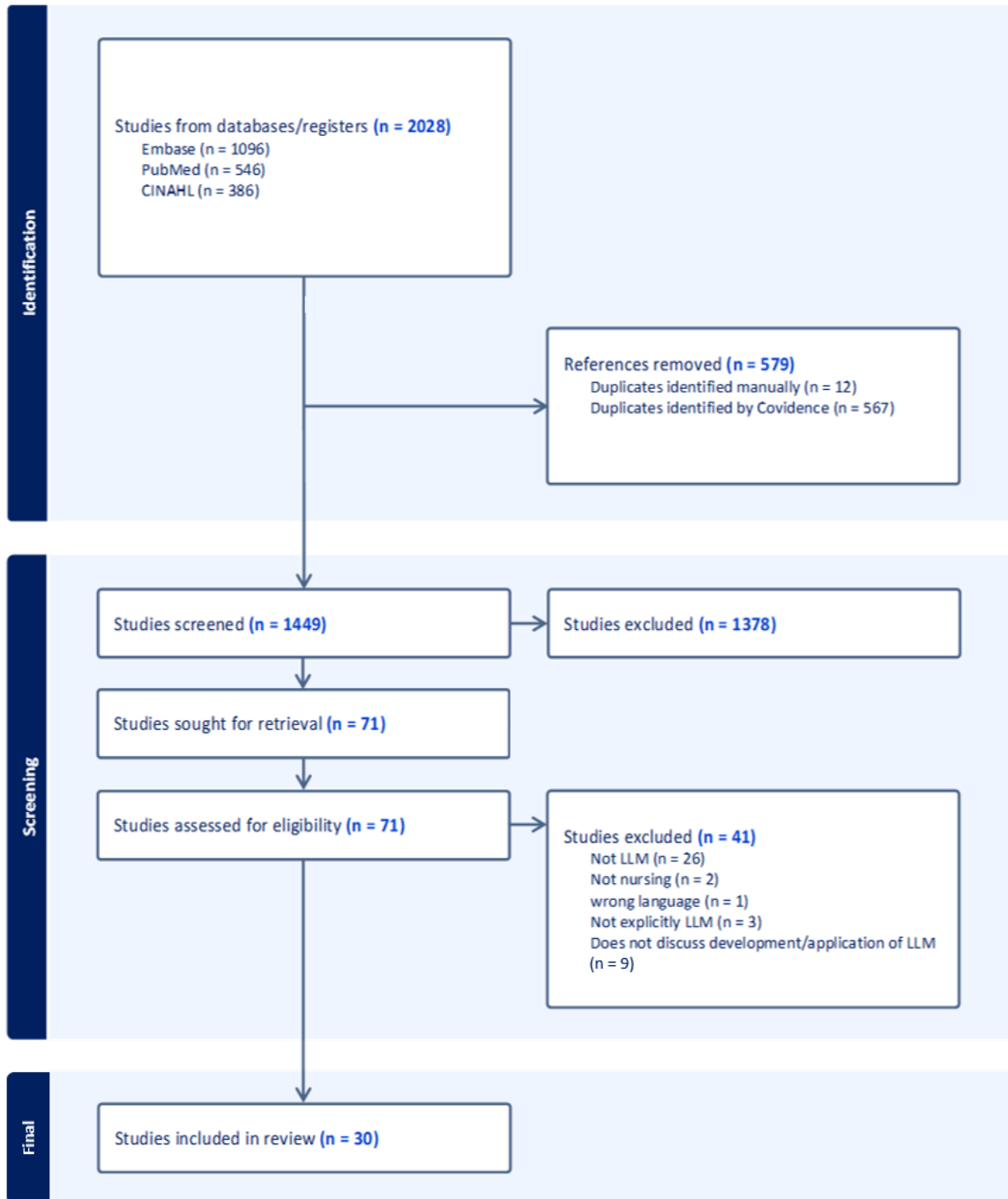
Ray & Majumder, 2023	2023	India	Editorial/Letter to Editor	N/A	To provide a response to a previously published paper on the use of ChatGPT for nursing
Reunamo et al., 2022	2022	Finland	Research article	Observational study	To extract keywords and phrases that provide an intuitive overview of the content in multiple nursing entries in electronic health records written during individual patients' care episodes
Scerri & Morin, 2023	2023	Malta	Editorial/Letter to Editor	N/A	To examine the response of ChatGPT when considering its use in nursing practice
Sun & Hoelscher, 2023	2023	United States	Perspective, Brief/report	N/A	To discuss recommendations of the use of ChatGPT in Nursing Education
Taira et al., 2023	2023	Japan	Research article	Observational study	To evaluate the accuracy of ChatGPT on the Japanese National Nurse Examination

Tam et al., 2023	2023	Singapore , Vietnam	Perspective, Commentary	N/A	To discuss the challenges and implications of artificial intelligence powered chatbot (Artificial Intelligence Chatbots) in nursing education
Thakur et al., 2023	2023	Canada	Editorial/Letter to Editor	N/A	To provide an opinion on the potential value in using ChatGPT to help curriculum developers in and ultimately learners in nursing education design
Van Bulck & Moons, 2023	2023	Belgium	Research article	Observational study	To evaluate ChatGPT responses to patient questions in terms of trustworthiness, value and danger
Vitorino & Júnior, 2023	2023	Brazil	Editorial/Letter to Editor	N/A	To discuss the use of ChatGPT in nursing education
Woodnutt et al., 2023	2023	United Kingdom	Research article, Case study, Commentary	Observational study	To use ChatGPT to generate a mental health nursing care plan for a fictitious patient and evaluate the quality of it using expert opinion in the form of NICE guidelines and the authors' own experiences in this care context

Zong et al., 2023	2023	China	Research article	Observational study	To quantitatively evaluate the performance of ChatGPT on three types of national medical examinations in China
-------------------	------	-------	------------------	---------------------	--

## Figures

Figure 1. PRISMA Flowchart



## Supplementary Materials

Supplementary Material A: Search Query

Database: MEDLINE (via PubMed)

("Machine Learning"[Mesh] OR "Neural Networks, Computer"[Mesh] OR "Natural Language Processing"[Mesh] OR "large language model\*"[Title/Abstract] OR ChatGPT[Title/Abstract] OR "Generative Pre-trained Transformer"[Title/Abstract] OR GPT[Title/Abstract] OR "Google Bard"[Title/Abstract] OR LLM[Title/Abstract] OR OpenAI[Title/Abstract] OR "Natural Language Generat\*"[Title/Abstract] OR NLG[Title/Abstract]) AND (nurs\*[Text Word] OR RN[Title/Abstract] OR ARNP[Title/Abstract] OR BSN[Title/Abstract])

Results: 546

**Database: CINAHL**

((MH "Nurses+") OR nurs\* OR RN OR ARNP OR BSN) AND ((MH "Machine Learning+") OR (MH "Neural Networks (Computer)") OR (MH "Natural Language Processing") OR "large language model\*" OR chatgpt OR "generative pre-trained transformer" OR gpt OR "google bard" OR llm OR openai OR "natural language generat\*" OR nlg)

Results: 386

**Database: Embase**

('machine learning'/de OR 'artificial neural network'/de OR 'feature learning (machine learning)'/exp OR 'supervised machine learning'/exp OR 'unsupervised machine learning'/exp OR 'support vector machine'/exp OR 'natural language processing'/exp OR 'chatgpt:ti,ab' OR 'large language model\*:ti,ab OR 'generative pre-trained transformer':ti,ab OR gpt:ti,ab OR 'google bard':ti,ab OR llm:ti,ab OR openai:ti,ab OR 'natural language generat\*:ti,ab OR nlg:ti,ab) AND ('nursing'/exp OR 'nurse'/exp OR nurs\*:ti,ab OR rn:ti,ab OR arnp:ti,ab OR bsn:ti,ab)

Results: 1,096

Supplementary Material B: Inclusion and Exclusion Criteria

Inclusion Criteria	Exclusion Criteria
<ul style="list-style-type: none"> <li>● All articles discussing the development or application of large language models (Thirunavukarasu et al, 2023)</li> <li>● Articles that focus on large language models</li> <li>● Articles with nurses as the primary stakeholder group</li> <li>● Large language models to include: ChatGPT, LLaMA, MT-NLG, Language Model for Dialogue Applications (LaMDA), Anthropic-LM, Pathways Language Model</li> </ul>	<ul style="list-style-type: none"> <li>● Studies that focus on artificial intelligence broadly</li> <li>● Studies that focus on BERT models</li> </ul>

(PaLM) and Open Pretrained Transformer (OPT), BART (Google)	
---	--

## References

1. Birhane A, Kasirzadeh A, Leslie D, Wachter S. Science in the age of large language models. *Nature Reviews Physics* 2023 5:5. 2023;5(5):277-280. doi:10.1038/s42254-023-00581-4
2. Kasneci E, Sessler K, Küchemann S, et al. ChatGPT for good? On opportunities and challenges of large language models for education. *Learn Individ Differ*. 2023;103:102274. doi:10.1016/J.LINDIF.2023.102274
3. Thirunavukarasu AJ, Ting DSJ, Elangovan K, Gutierrez L, Tan TF, Ting DSW. Large language models in medicine. *Nature Medicine* 2023 29:8. 2023;29(8):1930-1940. doi:10.1038/s41591-023-02448-8
4. Clusmann J, Kolbinger FR, Muti HS, et al. The future landscape of large language models in medicine. *Communications Medicine* 2023 3:1. 2023;3(1):1-8. doi:10.1038/s43856-023-00370-1
5. Li H, Moon JT, Purkayastha S, Celi LA, Trivedi H, Gichoya JW. Ethics of large language models in medicine and medical research. *Lancet Digit Health*. 2023;5(6):e333-e335. doi:10.1016/S2589-7500(23)00083-3
6. Sallam M. ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns. *Healthcare (Basel)*. 2023;11(6). doi:10.3390/HEALTHCARE11060887
7. Yu P, Xu H, Hu X, Deng C. Leveraging Generative AI and Large Language Models: A Comprehensive Roadmap for Healthcare Integration. *Healthcare (Basel)*. 2023;11(20). doi:10.3390/HEALTHCARE11202776
8. von Gerich H, Moen H, Block LJ, et al. Artificial Intelligence -based technologies in nursing: A scoping literature review of the evidence. *Int J Nurs Stud*. 2022;127:104153. doi:10.1016/J.IJNURSTU.2021.104153

9. Tiase VL, Hobensack M. Enabling Techquity in Nursing Practice: Informatics, Technology, and Innovation. *The Future of Nursing 2020-2030: Global Applications to Advance Health Equity*. Published online 2023:131-144. doi:10.1007/978-3-031-29746-5\_10
10. Cochrane Rapid Review (RR) Definition.
11. Garritty C, Hamel C, Hersi M, et al. Assessing how information is packaged in rapid reviews for policy-makers and other stakeholders: A cross-sectional study. *Health Res Policy Syst*. 2020;18(1):1-17. doi:10.1186/S12961-020-00624-7/TABLES/3
12. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *The BMJ*. 2021;372. doi:10.1136/bmj.n71
13. Li T, Higgins JP, Deeks JJ. Chapter 5: Collecting data . In: *Cochrane Handbook for Systematic Reviews of Interventions*. Vol 6. 4th ed. ; 2023. Accessed December 28, 2023. <https://training.cochrane.org/handbook/current/chapter-05>
14. Abdulai AF, Hung L. Will ChatGPT undermine ethical values in nursing education, research, and practice? *Nurs Inq*. 2023;30(3):e12556. doi:10.1111/NIN.12556
15. Castonguay A, Farthing P, Davies S, et al. Revolutionizing nursing education through Ai integration: A reflection on the disruptive impact of ChatGPT. *Nurse Educ Today*. 2023;129:105916. doi:10.1016/J.NEDT.2023.105916
16. Draganic K. Artificial intelligence: Opportunities and challenges in NP education. *Nurse Pract*. 2023;48(4):6. doi:10.1097/01.NPR.0000000000000023
17. Fraser D. Who Wrote This Editorial? The Use of Artificial Intelligence-Generated Content in Nursing Literature. *Neonatal Netw*. 2023;42(3):115-117. doi:10.1891/NN-2023-0027
18. Frith KH. ChatGPT: Disruptive Educational Technology. *Nurs Educ Perspect*. 2023;44(3):198-199. doi:10.1097/01.NEP.0000000000001129
19. Lyon D. Artificial Intelligence for Oncology Nursing Authors: Potential Utility and Concerns About Large Language Model Chatbots. *Oncol Nurs Forum*. 2023;50(3):276-277. doi:10.1188/23.ONF.276-277
20. Miao H, Ahn H. Impact of ChatGPT on Interdisciplinary Nursing Education and Research. *Asian Pac Isl Nurs J*. 2023;7:48136. doi:10.2196/48136
21. Odom-Forren J. The Role of ChatGPT in Perianesthesia Nursing. *Journal of Perianesthesia Nursing*. 2023;38(2):176-177. doi:10.1016/j.jopan.2023.02.006
22. Sun GH, Hoelscher SH. The ChatGPT Storm and What Faculty Can Do. *Nurse Educ*. 2023;48(3):119-124. doi:10.1097/NNE.0000000000001390
23. Thakur A, Parikh D, Thakur A. ChatGPT in nursing education: Is there a role for curriculum development? - ClinicalKey for Nursing. *Teaching and Learning in Nursing*. 2023;18(3):450-451. Accessed December 30, 2023. <https://www.clinicalkey.com/nursing/#!/content/journal/1-s2.0-S1557308723000574>
24. Ahmed SK. The Impact of ChatGPT on the Nursing Profession: Revolutionizing Patient Care and Education. *Ann Biomed Eng*. 2023;51(11):2351-2352. doi:10.1007/S10439-023-03262-6
25. Alanzi TM. Impact of ChatGPT on Teleconsultants in Healthcare: Perceptions of Healthcare Experts in Saudi Arabia. *J Multidiscip Healthc*. 2023;16:2309. doi:10.2147/JMDH.S419847

26. Berşe S, Akça K, Dirgar E, Kaplan Serin E. The Role and Potential Contributions of the Artificial Intelligence Language Model ChatGPT. *Ann Biomed Eng*. Published online 2023. doi:10.1007/S10439-023-03296-W
27. Choi EPH, Lee JJ, Ho MH, Kwok JYY, Lok KYW. Chatting or cheating? The impacts of ChatGPT and other artificial intelligence language models on nurse education. *Nurse Educ Today*. 2023;125. doi:10.1016/J.NEDT.2023.105796
28. Gunawan J. Exploring the future of nursing: Insights from the ChatGPT model. *Belitung Nurs J*. 2023;9(1):1. doi:10.33546/BNJ.2551
29. Nashwan AJ, Abujaber AA. Harnessing Large Language Models in Nursing Care Planning: Opportunities, Challenges, and Ethical Considerations. *Cureus*. 2023;15(6). doi:10.7759/CUREUS.40542
30. Ray PP, Majumder P. Dr. ChatGPT in Cardiovascular Nursing: A Deeper Dive into Trustworthiness, Value, and Potential Risks. *European journal of cardiovascular nursing*. Published online May 17, 2023. doi:10.1093/EURJCN/ZVAD047
31. Taira K, Itaya T, Hanada A. Performance of the Large Language Model ChatGPT on the National Nurse Examinations in Japan: Evaluation Study. *JMIR Nurs*. 2023;6(1). doi:10.2196/47305
32. Tam W, Huynh T, Tang A, Luong S, Khatri Y, Zhou W. Nursing education in the age of artificial intelligence powered Chatbots (AI-Chatbots): Are we ready yet? *Nurse Educ Today*. 2023;129. doi:10.1016/J.NEDT.2023.105917
33. Zong H, Li J, Wu E, Wu R, Shen B. Performance of ChatGPT on Chinese National Medical Licensing Examinations: A Five-Year Examination Evaluation Study for Physicians, Pharmacists and Nurses. *medRxiv*. Published online July 9, 2023:2023.07.09.23292415. doi:10.1101/2023.07.09.23292415
34. Allen C, Woodnutt S. Can ChatGPT pass a nursing exam? *Int J Nurs Stud*. 2023;145. doi:10.1016/J.IJNURSTU.2023.104522
35. Hunter J, Freer Y, Gatt A, Reiter E, Sripada S, Sykes C. Automatic generation of natural language nursing shift summaries in neonatal intensive care: BT-Nurse. *Artif Intell Med*. 2012;56(3):157-172. doi:10.1016/J.ARTMED.2012.09.002
36. Moons P, Van Bulck L. ChatGPT: can artificial intelligence language models be of value for cardiovascular nurses and allied health professionals. *European journal of cardiovascular nursing*. 2023;22(7):E55-E59. doi:10.1093/EURJCN/ZVAD022
37. O'Connor S, ChatGPT. Open artificial intelligence platforms in nursing education: Tools for academic progress or abuse? *Nurse Educ Pract*. 2023;66. doi:10.1016/J.NEPR.2022.103537
38. Reunamo A, Peltonen LM, Mustonen R, et al. Text Classification Model Explainability for Keyword Extraction - Towards Keyword-Based Summarization of Nursing Care Episodes. *Stud Health Technol Inform*. 2022;290:632-636. doi:10.3233/SHTI220154
39. Scerri A, Morin KH. Using chatbots like ChatGPT to support nursing practice. *J Clin Nurs*. 2023;32(15-16):4211-4213. doi:10.1111/JOCN.16677
40. Van Bulck L, Moons P. What if your patient switches from Dr. Google to Dr. ChatGPT? A vignette-based survey of the trustworthiness, value and danger of ChatGPT-generated responses to health questions. *European journal of cardiovascular nursing*. Published online April 24, 2023. doi:10.1093/EURJCN/ZVAD038

41. Woodnutt S, Allen C, Snowden J, et al. Could artificial intelligence write mental health nursing care plans? *J Psychiatr Ment Health Nurs*. Published online 2023. doi:10.1111/JPM.12965
42. Irwin P, Jones D, Fealy S. What is ChatGPT and what do we do with it? Implications of the age of AI for nursing and midwifery practice and education: An editorial. *Nurse Educ Today*. 2023;127. doi:10.1016/J.NEDT.2023.105835
43. Vitorino LM, Júnior GHY. ChatGPT and the teaching of contemporary nursing: And now professor? *J Clin Nurs*. 2023;32(21-22):7921-7922. doi:10.1111/JOCN.16706
44. De Angelis L, Baglivo F, Arzilli G, et al. ChatGPT and the rise of large language models: the new AI-driven infodemic threat in public health. *Front Public Health*. 2023;11. doi:10.3389/FPUBH.2023.1166120/FULL
45. Gottlieb S, Silvis L. How to Safely Integrate Large Language Models Into Health Care. *JAMA Health Forum*. 2023;4(9):e233909-e233909. doi:10.1001/JAMAHEALTHFORUM.2023.3909
46. Generative AI & LLMs in Health & Medicine - National Academy of Medicine. Accessed December 28, 2023. <https://nam.edu/event/generative-ai-llms-in-health-medicine/>
47. Ibrahim H, Liu X, Denniston AK. Reporting guidelines for artificial intelligence in healthcare research. *Clin Exp Ophthalmol*. 2021;49(5):470-476. doi:10.1111/CEO.13943
48. WHO outlines considerations for regulation of artificial intelligence for health. Accessed December 28, 2023. <https://www.who.int/news/item/19-10-2023-who-outlines-considerations-for-regulation-of-artificial-intelligence-for-health>
49. Crossnohere NL, Elsaid M, Paskett J, Bose-Brill S, Bridges JFP. Guidelines for Artificial Intelligence in Medicine: Literature Review and Content Analysis of Frameworks. *J Med Internet Res*. 2022;24(8). doi:10.2196/36823
50. Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable artificial intelligence in health care. *Lancet Digit Health*. 2021;3(11):e745-e750. doi:10.1016/S2589-7500(21)00208-9
51. MacLean Ainsley. 'Making Friends' With ChatGPT: A Guide To Best Practices For AI Interactions. Published 2023. Accessed December 28, 2023. <https://www.forbes.com/sites/forbestechcouncil/2023/03/27/making-friends-with-chatgpt-a-guide-to-best-practices-for-ai-interactions/?sh=745994051e55>
52. Health Infoway C. Toolkit for Implementers of Artificial Intelligence in Health Care (Version 2.0). Published online 2023.
53. Jin W, Li X, Fatehi M, Hamarneh G. Guidelines and evaluation of clinical explainable AI in medical image analysis. *Med Image Anal*. 2023;84. doi:10.1016/J.MEDIA.2022.102684
54. Barredo Arrieta A, Díaz-Rodríguez N, Del Ser J, et al. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*. 2020;58:82-115. doi:10.1016/J.INFFUS.2019.12.012
55. Clarke YD. *H.R.6580 - 117th Congress (2021-2022): Algorithmic Accountability Act of 2022*. House of Representatives; 2022. Accessed December 30, 2023. <https://www.congress.gov/bill/117th-congress/house-bill/6580/text>
56. The Medical Devices Regulations 2002. Accessed December 30, 2023. <https://www.legislation.gov.uk/uksi/2002/618/contents/made>

57. Large Language Models and software as a medical device - MedRegs. Accessed December 30, 2023. <https://medregs.blog.gov.uk/2023/03/03/large-language-models-and-software-as-a-medical-device/>
58. EU AI Act: first regulation on artificial intelligence . European Parliament. Published 2023. Accessed January 14, 2024. <https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>
59. Kurtzman ET, Ghazal L V., Girouard S, et al. Nursing Workforce Challenges in the Postpandemic World. *J Nurs Regul.* 2022;13(2):49. doi:10.1016/S2155-8256(22)00061-8
60. National Academies of Sciences E and M. Artificial Intelligence in Health Professions Education: Proceedings of a Workshop. *Artificial Intelligence in Health Professions Education.* Published online August 14, 2023. doi:10.17226/27174
61. ChatGPT and Other Generative AI Tools: Thinking about ChatGPT. The University of British Columbia. Published 2023. Accessed December 30, 2023. <https://academicintegrity.ubc.ca/chatgpt-faq/#:~:text=If%20the%20use%20of%20ChatGPT,accessing%20websites%20or%20other%20online>
62. Goto A, Katanoda K. Should We Acknowledge ChatGPT as an Author? *J Epidemiol.* 2023;33(7):333. doi:10.2188/JEA.JE20230078
63. McAdoo T. How to cite ChatGPT. APA Style. doi:10.1037/0000165-000
64. How do I cite generative AI in MLA style. MLA Style Center. Published 2023. Accessed December 30, 2023. <https://style.mla.org/citing-generative-ai/>