ChatGPT, a language model developed by OpenAI, has the potential to play a role in public health. With its ability to generate human-like text based on large amounts of data, ChatGPT has the potential to support individuals and communities in making informed decisions about their health (Panch et al. Lancet Digit Health 1:e13–e14, 2019; Baclic et al. Canada Commun Dis Rep 46.6:161, 2020). However, as with any technology, there are limitations and challenges to consider when using ChatGPT in public health. In this overview, we will examine the potential uses of ChatGPT in public health, as well as the advantages and disadvantages of its use.

ChatGPT is an artificial intelligence (AI)-based conversational large language model (LLM). The potential applications of LLMs in health care education, research, and practice could be promising if the associated valid concerns are proactively examined and addressed. The current systematic review aimed to investigate the utility of ChatGPT in health care education, research, and practice and to highlight its potential limitations. Using the PRIMSA guidelines, a systematic search was conducted to retrieve English records in PubMed/MEDLINE and Google Scholar (published research or preprints) that examined ChatGPT in the context of health care education, research, or practice. A total of 60 records were eligible for inclusion. Benefits of ChatGPT were cited in 51/60 (85.0%) records and included: (1) improved scientific writing and enhancing research equity and versatility; (2) utility in health care research (efficient analysis of datasets, code generation, literature reviews, saving time to focus on experimental design, and drug discovery and development); (3) benefits in health care practice (streamlining the workflow, cost saving, documentation, personalized medicine, and improved health literacy); and (4) benefits in health care education including improved personalized learning and the focus on critical thinking and problem-based learning. Concerns regarding ChatGPT use were stated in 58/60 (96.7%) records including ethical, copyright, transparency, and legal issues, the risk of bias, plagiarism, lack of originality, inaccurate content with risk of hallucination, limited knowledge, incorrect citations, cybersecurity issues, and risk of infodemics. The promising applications of ChatGPT can induce paradigm shifts in health care education, research, and practice. However, the embrace of this AI chatbot should be conducted with extreme caution considering its potential limitations. As it currently stands, ChatGPT does not qualify to be listed as an author in scientific articles unless the ICMJE/COPE guidelines are revised or amended. An initiative involving all stakeholders in health care education, research, and practice is urgently needed. This will help to set a code of ethics to guide the responsible use of ChatGPT among other LLMs in health care and

. Introduction

Artificial intelligence (AI) can be defined as the multidisciplinary approach of computer science and linguistics that aspires to create machines capable of performing tasks that normally require human intelligence [1]. These tasks include the ability to learn, adapt, rationalize, understand, and to fathom abstract concepts as well as the reactivity to complex human attributes such as attention, emotion, creativity, etc. [2].

The history of AI as a scientific discipline can be traced back to the mid-XX century at the Dartmouth Summer Research Project on AI [3]. This was followed by the development of machine learning (ML) algorithms that allow decision-making or predictions based on the patterns in large datasets [4]. Subsequently, the development of neural networks (brain-mimicking algorithms), genetic algorithms (finding optimal solutions for complex problems by application of evolutionary principles), and other advanced techniques followed [5].

Launched in November 2022, “ChatGPT” is an AI-based large language model (LLM) trained on massive text datasets in multiple languages with the ability to generate human-like responses to text input [6]. Developed by OpenAI (OpenAI, L.L.C., San Francisco, CA, USA), ChatGPT etymology is related to being a chatbot (a program able to understand and generate responses using a text-based interface) and is based on the generative pre-trained transformer (GPT) architecture [6,7]. The GPT architecture utilizes a neural network to process natural language, thus generating responses based on the context of input text [7]. The superiority of ChatGPT compared to its GPT-based predecessors can be linked to its ability to respond to multiple languages generating refined and highly sophisticated responses based on advanced modeling [6,7].

In the scientific community and academia, ChatGPT has received mixed responses reflecting the history of controversy regarding the benefits vs. risks of advanced AI technologies [8,9,10]. On one hand, ChatGPT, among other LLMs, can be beneficial in conversational and writing tasks, assisting to increase the efficiency and accuracy of the required output [11]. On the other hand, concerns have been raised in relation to possible bias based on the datasets used in ChatGPT training, which can limit its capabilities and could result in factual inaccuracies, but alarmingly appear to be scientifically plausible (a phenomenon termed hallucination) [11]. Additionally, security concerns and the potential of cyber-attacks with the spread of misinformation utilizing LLMs should also be considered [11].

The innate resistance of the human mind to any change is a well-described phenomenon and can be understandable from evolutionary and social psychology perspectives [12]. Therefore, the concerns and debate that arose instantaneously following the widespread release of ChatGPT appear to be understandable. The attention that ChatGPT received involved several disciplines. In education, for example, ChatGPT release could mark the end of essays as assignments [13]. In health care practice and academic writing, factual inaccuracies, ethical issues, and the fear of misuse including the spread of misinformation should be considered [14,15,16].

The versatility of human intelligence (HI) compared to AI is related to its biological evolutionary history, adaptability, creativity, the ability of emotional intelligence, and the ability to understand complex abstract concepts [2]. However, HI-AI cooperation can be beneficial if an accurate and reliable output of AI is ensured. The promising utility of AI in health care has been outlined previously with possible benefits in personalized medicine, drug discovery, and the analysis of large datasets aside from the potential applications to improve diagnosis and clinical decisions [17,18]. Additionally, the utility of AI chatbots in health care education is an interesting area to probe. This is related to the massive information and various concepts that health care students are required to grasp [19]. However, all of these applications should be considered cautiously considering the valid concerns, risks, and categorical failures experienced and cited in the context of LLM applications. Specifically, Borji comprehensively highlighted the caveats of ChatGPT use that included, but were not limited to, the generation of inaccurate content, the risk of bias and discrimination, lack of transparency and reliability, cybersecurity concerns, ethical consequences, and societal implications [20].

Therefore, the aim of the current review was to explore the future perspectives of ChatGPT as a prime example of LLMs in health care education, academic/scientific writing, health care research, and health care practice based on the existing evidence. Importantly, the current review objectives extended to involve the identification of potential limitations and concerns that could be associated with the application of ChatGPT in the aforementioned areas in health care settings.

2. Materials and Methods

2.1. Search Strategy and Inclusion Criteria

The current systematic review was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRIMSA) guidelines [**21**]. The information sources included PubMed/MEDLINE and Google Scholar.

The eligibility criteria involved any type of published scientific research or preprints (article, review, communication, editorial, opinion, etc.) addressing ChatGPT that fell under the following categories: (1) health care practice/research; (2) health care education; and (3) academic writing.

The exclusion criteria included: (1) non-English records; (2) records addressing ChatGPT in subjects other than those mentioned in the eligibility criteria; and (3) articles from non-academic sources (e.g., newspapers, internet websites, magazines, etc.).

The exact PubMed/MEDLINE search strategy, which concluded on 16 February 2023, was as follows: (ChatGPT) AND ((“2022/11/30” [Date–Publication]: “3000” [Date–Publication])), which yielded 42 records.

The search on Google Scholar was conducted using Publish or Perish (Version 8) [**22**]. The search term was “ChatGPT” for the years 2022–2023, and the Google Scholar search yielded 238 records and concluded on 16 February 2023.

2.2. Summary of the Record Screening Approach

The records retrieved following the PubMed/MEDLINE and Google Scholar searches were imported to EndNote v.20 for Windows (Thomson ResearchSoft, Stanford, CA, USA), which yielded a total of 280 records.

Next, screening of the title/abstract was conducted for each record with the exclusion of duplicate records (n = 40), followed by the exclusion of records published in languages other than English (n = 32). Additionally, the records that fell outside the scope of the review (records that examined ChatGPT in a context outside health care education, health care practice, or scientific research/academic writing) were excluded (n = 80). Moreover, the records published in non-academic sources (e.g., newspapers, magazines, Internet websites, blogs, etc.) were excluded (n = 18).

Afterward, full screening of the remaining records (n = 110) was carried out with the exclusion of an additional 41 records that fell outside the scope of the current review. An additional nine records were excluded due to the inability to access the full text of these records, being subscription-based. This yielded a total of 60 records eligible for inclusion in the current review.

2.3. Summary of the Descriptive Search for ChatGPT Benefits and Risks in the Included Records

Each of the included records was searched specifically for the following: (1) type of record (preprint, published research article, opinion, commentary, editorial, review, etc.); (2) the listed benefits/applications of ChatGPT in health care education, health care practice, or scientific research/academic writing; (3) the listed risks/concerns of ChatGPT in health care education, health care practice, or scientific research/academic writing; and (4) the main conclusions and recommendations regarding ChatGPT in health care education, health care practice, or scientific research/academic writing.

Categorization of the benefits/applications of ChatGPT was as follows: (1) educational benefits in health care education (e.g., generation of realistic and variable clinical vignettes,

customized clinical cases with immediate feedback based on the student's needs, enhanced communications skills); (2) benefits in academic/scientific writing (e.g., text generation, summarization, translation, and literature review in scientific research); (3) benefits in scientific research (e.g., efficient analysis of large datasets, drug discovery, identification of potential drug targets, generation of codes in scientific research); (4) benefits in health care practice (e.g., improvements in personalized medicine, diagnosis, treatment, lifestyle recommendations based on personalized traits, documentation/generation of reports); and (5) being a freely available package.

Categorization of the risks/concerns of ChatGPT was as follows: (1) ethical issues (e.g., risk of bias, discrimination based on the quality of training data, plagiarism); (2) hallucination (the generation of scientifically incorrect content that sounds plausible); (3) transparency issues (black box application); (4) risk of declining need for human expertise with subsequent psychologic, economic and social issues; (5) over-detailed, redundant, excessive content; (6) concerns about data privacy for medical information; (7) risk of declining clinical skills, critical thinking and problem-solving abilities; (8) legal issues (e.g., copyright issues, authorship status); (9) interpretability issues; (10) referencing issues; (11) risk of academic fraud in research; (12) incorrect content; and (13) infodemic risk.