tutor
VOL. 24 N. 1, 2024: 46 - 77
DOI: 10.13135/1971-8551/11403

Teachers and students' perception of Artificial Intelligence in Medical Education: A meta-ethnographic review

Consorti Fabrizio[1], D'Oria Marika[2]

[1] Gruppo Qualità e Innovazione nella Didattica, Università Sapienza di Roma, Rome (IT) fabrizio.consorti@uniroma1.it
[2] Center of Excellence in Oncology, Ospedale Isola Tiberina – Gemelli Isola, Rome (IT) - marika.doria.fw@fbf-isola.it

ABSTRACT

Introduction. This review synthesises the results of qualitative studies on the use of AI in medical education

Methods and tools. The review is based on a systematic review of qualitative studies, that were synthesised with the meta-ethnographic method.

Results. Eighteen selected articles resulted in five themes: convenience of AI, resemblance to reality of AI, critical understanding of users of AI, dialogic and social nature of AI, multidisciplinarity as an environment of design.

Discussion. The use of AI in education is still in its preliminary stage. This review support some paths of research.

Keywords:
artificial intelligence, medical education, simulation, robotics, diffusion of innovation

ABSTRACT

Introduzione. Questa revisione sintetizza i risultati degli studi qualitativi sull'uso dell'IA nella formazione medica.

Metodi e strumenti. La revisione si basa su una revisione sistematica di studi qualitativi, sintetizzati con il metodo meta-etnografico da Nobles.

Risultati. Diciotto articoli selezionati hanno dato origine a cinque temi: convenienza dell'IA, somiglianza con la realtà dell'IA, comprensione critica degli utenti dell'IA, natura dialogica e sociale dell'IA, multidisciplinarità come ambiente di progettazione.

Discussione. L'uso dell'IA nell'istruzione è ancora in fase preliminare. Questa rassegna sostiene alcuni percorsi di ricerca.

Parole chiave:
intelligenza artificiale, educazione medica, simulazione, robotica, diffusione dell'innovazione

Take Home message

- The use of AI is quickly spreading in many fields, without a full understanding and awareness of its pitfall and opportunities
- There are conflicting experiences and attitudes both in teachers and learners about the use of AI in education
- A multiprofessional and multidisciplinary environment is needed to develop new theoretical models and implementations of AI in medical education

- L'uso dell'IA si sta rapidamente diffondendo in molti campi senza una piena comprensione e consapevolezza delle sue insidie e opportunità
- Esistono esperienze e atteggiamenti contrastanti sia negli insegnanti che negli studenti riguardo all'uso dell'IA nella formazione
- È necessario un ambiente multiprofessionale e multidisciplinare per sviluppare nuovi modelli teorici e implementazioni dell'IA nella formazione medica

INTRODUCTION

Medical education is facing several challenges in the current scenario. On one hand, an epistemological transformation impacts its practices, processes, habits, and outcomes. Despite the integration of Artificial Intelligence (AI) and digital solutions in care is still at its infancy (also due to a lack of regulatory maturity; Kudina, 2021), there is the need for understanding and assessing their real benefits, beyond enthusiasms and fears. On the other hand, the widespread of formal, informal, and non-formal education initiatives (see Tudor Car et al., 2022) is mitigating —not yet fully satisfying— several unmet needs for this field. Even when scientifically valid, most of these scattered opportunities are not yet fully recognized or certified according to the European (or other) standards.

Competencies must be proven because their practices directly affect the patients. While questioning about the accessibility and the equity of these educational chances for students and professionals' Lifelong Lifewide Learning (LLL), this great heterogeneity is representative of a "Liquid Education" (Vojtěch, 2022) scenario, with a jeopardized —in quality and quantity— andragogical offer. Political, academic, and industrial strategies contribute to shaping healthcare, the labour market, and the way we traditionally conceived career design; rapid changes and a high-performance culture impact on individual life projects in terms of choice and commitment to a career, while developing new "hybrid" professions that adapt themselves to the mutability and flexibility required by innovation, because the scientific and technological progress lies at the basis of their practices.

The lack of clarity can be also inked to the concept of AI itself, which is considered and umbrella term. The expression "Artificial Intelligence" was first used by John McCarthy in 1956, and the first artificial neural network was created in the following decade. A previously unthinkable phenomenon began to grow to today's commercially available AI assistants, such as Alexa or Siri (Nobles et al., 2020), and the Large Language Models (LLM) such as ChatGPT or Gemini.

The term AI is now part of the current discourse on the future development of the economy and work environment (Mamedov et al., 2018), social life and education (Williamson, 2024). The term AI brings together many different computational methods and applications. AI is gaining momentum in medical education as well, both as a topic of learning (Jafri et al., 2024) and as a tool to enhance the possibilities of teaching and the interactions with the students (Tozsin et al., 2024).

A special attention has been devoted to the use of the LLMs, as demonstrated by a recent systematic review focusing the use of ChatGPT (Xu et al., 2024). In a large survey, students showed to be willing to include AI as a topic in medical curriculum (Jackson et al., 2024). Some quantitative studies explored the effectiveness of AI based tools and methods for learning in medical education (Truong, 2016) (An & Wang, 2024)(Fazlollahi et al., 2022), mainly based on small scale experiments or self-administered surveys of satisfaction.

Less qualitative studies inquired the perception of these new learning environments from students and teachers. These studies were mainly addressed to inquire awareness and the need to include AI in medical curricula (Moldt et al., 2024) or qualitatively comparing the performance of an AI system with an human expert (Shamim et al., 2024). Most often, the conclusions of these studies have been reported with little consideration of the learning mechanisms at play and without an explicitly stated theoretical framework in the interpretation of the results.

The goal of this meta-ethnographic review is to offer a preliminary understanding of the benefits and drawbacks, as seen by educators and students, of utilizing AI-based tools and techniques in medical education. The intention is to provide a basic theoretical framework to guide future, better-quality studies in this area. From both the teacher and student point of view, the research questions were the following:

- What are the perceived positive aspects of the educational use of AI techniques in teaching and learning?
- What are the perceived negative aspects?

Which elements of the context influence the positive or negative perception?

MATERIALS AND METHODS

1. Meta-ethnography

Meta-ethnography was originally introduced by Noblit (Noblit, 1988) as a method to synthesize ethnographic studies, a type of qualitative research frequently used by social scientists, in which the researcher lives for some months in a community (an ethnos) to record behaviours and systems of thought. The method has been standardized and successfully adapted to synthesize other types of qualitative studies. In order to produce a trustworthy synthesis of data, diverse in nature and meaning, the core concept in meta-ethnography is the "translation" of the metaphors and concepts of the different studies.

Noblit used the term translation to mean the comparison between a concept in a study and a possibly related concept in another, as if it was the process of transforming a term in a language into a term in another language. The synthesis of the translations can be "reciprocal" (the concepts apparently related are similar in their focus) or "refutational" (the two studies refute each other, because the same concept has a different value and meaning).

This analysis is stepped and iterative, moving from the first comparison of two articles, their reciprocal or refutational synthesis, and moving to the comparison with further articles, until saturation is reached. Saturation means that no new concepts are present in any further examined article. Therefore, meta-ethnography does not mandatorily require a systematic bibliographic search and can be performed on a more limited number of articles than a traditional systematic review.

The steps of this meta-ethnographic synthesis were the following:

1. Identifying the synthesis's focus: the synthesis concentrated on the learning outcomes, teaching and learning activities, and curriculum organization that affected the way teachers and students perceived their

experiences using AI approaches. Both undergraduate and graduate education levels were considered.

2. Search strategy: the search strategy and inclusion/exclusion criteria were executed using a combination of search strings such as "artificial intelligence," "machine learning OR deep learning," "curriculum design," "teaching methods," "medical education/undergraduate," "medical education/graduate," and "qualitative research" in PubMed, Scopus, and the ERIC database. All qualitative research, including mixed-method studies, in which any method of AI was used in preclinical or clinical education for medical students (undergraduate and graduate) satisfied the inclusion criteria. We also selected commentaries and review articles, if they critically discussed their results from the teacher's perspective and not solely reported a list of fields of use of AI in education. Excluded publications included non-medical students; also, studies that solely reported the application of AI techniques (how I do it) or qualitative studies that did not provide quotes of the students' reaction.

   Finally, articles dealing with how best to train and prepare undergraduate and graduate students to the use of AI for clinical practice were also excluded. We did not assess the quality of the selected articles, but only considered the fact that they brought data relevant for our research questions. The bibliographic dataset was managed with the Zotero© software.

3. Data extraction: after carefully reading each article twice, the sentences that reflected an idea pertinent to the study goal were highlighted as raw data.

4. Concept correlation: a list of phrases or metaphors was generated for each article. To find common patterns, the final lists produced by the two researchers were compared. The definition of concept is "A meaningful idea that develops from comparing individual instances". Concepts, at their core, have to explain the data—not just describe it. In a meta-ethnography, "first order constructs" are the raw data; "second order

constructs" are the notions that emerge from comparing the lists of first order concepts.

5. Translation of studies: all publications were analysed for second order constructs, which were arranged chronologically. This arrangement made it easier to identify the emergence of a new idea and provided the chance to watch potential developments in the depiction of a thought. Reciprocal or refutational translations were used to indicate the similarities and differences between the studies.
6. Synthesis of translations: a restricted set of third-order interpretations, also referred to as themes, was constructed as a means of expressing the total as more than the sum of its parts, beginning with the second-order constructs and the assessment of reciprocal or refutational translation.
7. Synthesis of lines of argument: Lastly, the expression of the third order interpretation synthesis, called the line of arguments, was expressed in graphic form as a chart.

This article is compliant with the eNTreQ statement for the reporting of synthesis of qualitative research (Tong et al., 2012).

2. Data clustering

Table 1 summarizes some of the terms and their definitions, according to the Medical Subject Heading (MeSH, 2024) from the NIH Library of Medicine, which is defined as follows:

The Medical Subject Headings (MeSH) thesaurus is a controlled and hierarchically organized vocabulary produced by the National Library of Medicine. It is used for indexing, cataloging, and searching of biomedical and health-related information. MeSH includes the subject headings appearing in MEDLINE/PubMed, the NLM Catalog, and other NLM databases.

Consequently, every research article uploaded on PubMed and Medline large databases is catalogued according to keywords and nouns in the titles and abstracts that mirror the MeSH browser definitions. The choice of indicating the approved definitions by the NIH is to avoid any possible interpretation by the Authors and from the broader literature, due to the nature of the studied phenomenon (AI).

The list of definitions also helped us in the analysis of literature to further investigate if the researchers have common understanding on such topics or if they bring together different topics under the same name. In some cases, terms can be related to different subjects (e.g., Neural Networks can be related to Neurosciences as well as to Computer Sciences). Therefore, data descriptors have been integrated in Table 1.

To demystify the ambivalence, we reported the Unique ID as a code for term identification. Since annual changes may occur on the definitions, we added the date of establishment of the topic, the date of entry in the database, and the date of the last revision. Entry terms are keywords that bring to the same term of the first column when searched in the MeSH Browser: they might be considered as synonymous, despite substantial differences can occur.

Table 1. List of definitions on common terms surrounding Artificial Intelligence, retrieved from the MeSH browser (2024).

MESH HEADING / TERM	**DEFINITION**	**ENTRY TERMS***
Algorithms Unique ID: D000465 Date established: 1987 Data entry: 1986 Data revision: 2017	A procedure consisting of a sequence of algebraic formulas and/or logical steps to calculate or determine a given task.	N/A
Automation Unique ID: D001331 Date established: 1966 Data entry: 1999 Data revision: 2004	Controlled operation of an apparatus, process, or system by mechanical or electronic devices that take the place of human organs of observation, effort, and decision.	N/A
Artificial Intelligence (AI) Unique ID: D001185 Date established: 1986 Data entry: 1995 Data revision: 2017	Theory and development of <u>COMPUTER SYSTEMS</u> which perform tasks that normally require human intelligence. Such tasks may include speech recognition, <u>LEARNING; VISUAL PERCEPTION; MATHEMATICAL COMPUTING</u>; reasoning, <u>PROBLEM SOLVING, DECISION-MAKING</u>, and translation of language.	AI (Artificial Intelligence); Computational Intelligence; Computer Reasoning; Computer Vision Systems; Knowledge Acquisition (Computer); Knowledge Representation (Computer); Machine Intelligence
Big Data Unique ID: D000077558 Date established: 2019 Data entry: 2018 Data revision: 2018	Extremely large amounts of data which require rapid and often complex computational analyses to reveal patterns, trends, and associations, relating to various facets of human and non-human entities.	N/A
Decision Support Systems, Clinical (CDSS) Unique ID: D020000 Date established: 1998 Data entry: 1997	Computer-based information systems used to integrate clinical and patient information and provide support for decision-making in patient care.	Clinical Decision Support; Clinical Decision Support System; Clinical Decision Support Systems; Decision Support, Clinical

Data revision: 2020		
Computers Unique ID: D003201 Date established: 1966 Data entry: 1999 Data revision: 2018	Programmable electronic devices designed to accept data, perform prescribed mathematical and logical operations at high speed, and display the results of these operations.	Calculators, Programmable; Computer Hardware; Computers, Digital; Hardware, Computer
Computer simulation Unique ID: D003198 Date established: 1987 Data entry: 1986 Data revision: 2021	Computer-based representation of physical systems and phenomena such as chemical processes.	Computational Modeling; Computational Modelling; Computer Models; Computerized Models; In silico Modeling; In silico Models; In silico Simulation; Models, Computer
Computer Tutor (CT)	*No results for computer tutor in Main Heading Terms*	N/A
Computer Vision (CV) Redirected to: **Artificial Intelligence**	*No results for computer vision in Main Heading Terms*	N/A
Conversational agent	*No results for conversational agent in Main Heading Terms*	N/A
Neural Networks, Computer (NN) Unique ID: D016571 Date established: 1992 Data entry: 1991 Data revision: 2019 ***Artificial (ANN)*** **Convolutional (CNN)**	A computer architecture, implementable in either hardware or software, modeled after biological neural networks. Like the biological system in which the processing capability is a result of the interconnection strengths between arrays of nonlinear processing nodes, computerized neural networks, often called perceptrons or multilayer connectionist models, consist of neuron-like units. A homogeneous group of units makes up a layer. These networks are good at pattern recognition. They are adaptive, performing tasks by example, and thus are better for decision-making than are linear learning machines or cluster analysis. They do not require explicit programming. *No results for artificial neural network in Main Heading Terms* *No results for convolutional neural network in Main Heading Terms*	Computational Neural Networks; Connectionist Models; Models, Neural Network; Neural Network Models; Neural Networks (Computer); Perceptrons
Deep Learning (DL) Unique ID: D000077321 Date established: 2019 Data entry: 2018 Data revision: 2018	Supervised or unsupervised machine learning methods that use multiple layers of data representations generated by nonlinear transformations, instead of individual task-specific ALGORITHMS, to build and train neural network models.	Hierarchical Learning
Digital Health (DH) Unique ID: D000097103 Date established: 2024 Data entry: 2023 Data revision: 2023	Use of digital technologies in medicine and other health professions to manage illnesses and health risks and to promote wellness; includes the use of WEARABLE DEVICES; HEALTH INFORMATION TECHNOLOGY; ELECTRONIC HEALTH RECORDS; TELEMEDICINE; and PERSONALIZED MEDICINE	Digital Health Technology
Digital Medicine (DM)	*No results for digital medicine in Main Heading Terms*	N/A
Digital Therapeutics (DTx)	*No results for digital therapeutics in Main Heading Terms*	N/A
Generative Artificial Intelligence (Gen AI)	*No results for generative artificial intelligence in Main Heading Terms*	N/A
Generative Pre-trained Transformer (GPT)	*No results for generative pre-trained transformer in Main Heading Terms*	N/A
Hologram In: **Quantitative Phase Imaging** Unique ID: D000097913 Date established: 2024 Data entry: 2023 Data revision: 2023	*No results for hologram in Main Heading Terms* Contrast imaging generated from the measurements of phase shift changes of light waves as they pass through transparent objects of varying thickness and refractivity.	Digital Holographic Microscopy; Digital Holography; Holographic Microscopy; Microscopy, Quantitative Phase Contrast; Quantitative Phase Contrast Microscopy; Quantitative Phase Microscopy
Intelligent system	*No results for intelligent system in Main Heading Terms*	N/A
Internet of Things Unique ID: D000080487 Date established: 2020	*Networking capability which facilitates information flow to and from objects and devices using the INTERNET.*	N/A

Data entry: 2019 Data revision: 2019		
Large Language Model (LLM)	*No results for large language model in Main Heading Terms*	N/A
Machine Learning (ML) Unique ID: D000069550 Date established: 2016 Data entry: 2015 Data revision: 2019	A type of <u>ARTIFICIAL INTELLIGENCE</u> that enable <u>COMPUTERS</u> to independently initiate and execute <u>LEARNING</u> when exposed to new data.	Transfer Learning
Natural Language Processing (NLP) Unique ID: D009323 Date established: 1991 Data entry: 1986 Data revision: 2018	Computer processing of a language with rules that reflect and describe current usage rather than prescribed usage.	N/A
Picture Archiving and Communication System Redirected to: **Radiology Information Systems** Unique ID: D011873 Date established: 1991 Data entry: 1986 Data revision: 2020	Information systems, usually computer-assisted, designed to store, manipulate, and retrieve information for planning, organizing, directing, and controlling administrative activities associated with the provision and utilization of radiology services and facilities.	Archiving, Radiologic Picture; Information System, Radiologic; Information System, Radiology; Information Systems, Radiologic; Information Systems, Radiology; PACS (Radiology); Picture Archiving And Communication System; Picture Archiving and Communication Systems; Picture Archiving, Radiologic; Radiologic Information System; Radiologic Information Systems; Radiologic Picture Archiving; Radiology Information System; System, Radiologic Information; System, Radiology Information; Systems, Radiologic Information; Systems, Radiology Information; X-Ray Information Systems; Xray Information Systems
Patient simulation Unique ID: D016544 Date established: 1992 Data entry: 1991 Data revision: 2014	The use of persons coached to feign symptoms or conditions of real diseases in a life-like manner in order to teach or evaluate medical personnel.	N/A
Robotics Unique ID: D012371 Date established: 1987 Data entry: 1986 Data revision: 2023	The application of electronic, computerized control systems to mechanical devices designed to perform human functions. Formerly restricted to industry, but nowadays applied to artificial organs controlled by bionic (bioelectronic) devices, like automated insulin pumps and other prostheses.	Companion Robots; Humanoid Robots; Remote Operations (Robotics); Social Robots; Socially Assistive Robots; Soft Robotics; Telerobotics
Robotic Surgery (RS) In: **Robotic Surgical Procedures** Unique ID: D065287 Date established: 2015 Data entry: 2014 Data revision: 2020	*No results for robotic surgery in Main Heading Terms* Surgical procedures performed remotely using a computer that controls surgical instruments attached to mechanical arms designed to perform the tasks of the surgeon.	Robot Surgery; Robot-Assisted Surgery; Robot-Enhanced Procedures; Robot-Enhanced Surgery; Robotic-Assisted Surgery; Surgical Procedures, Robotic
Simulation training Unique ID: D000066908 Date established: 2016 Data entry: 2015 Data revision: 2014	A highly customized interactive medium or program that allows individuals to learn and practice real world activities in an accurate, realistic, safe and secure environment.	Interactive Learning
Social Robotics (SR) Redirected to: **Robotics**	*No results for social robotics in Main Heading Terms*	N/A
Support Vector Machine (SVM) Unique ID: D060388 Date established: 2021 Data entry: 2011 Data revision: 2017	<u>SUPERVISED MACHINE LEARNING</u> algorithm which learns to assign labels to objects from a set of training examples. Examples are learning to recognize fraudulent credit card activity by examining hundreds or thousands of fraudulent and non-fraudulent credit card activity or learning to make disease diagnosis or prognosis based on automatic classification of microarray gene expression profiles drawn from hundreds or thousands of samples.	Support Vector Network

* Other entry terms related to the exact definition on the MeSH Browser.

RESULTS

From a starting set of 7,818 articles, after removal of duplicates, reading of the title and exclusion of the articles not pertaining to the study, 96 articles were sought for retrieval. After reading the abstract and browsing the full text looking for qualitative data, 18 articles were retained for the analysis. Figure 1 shows the flow of selection and Table 2 lists the selected articles.

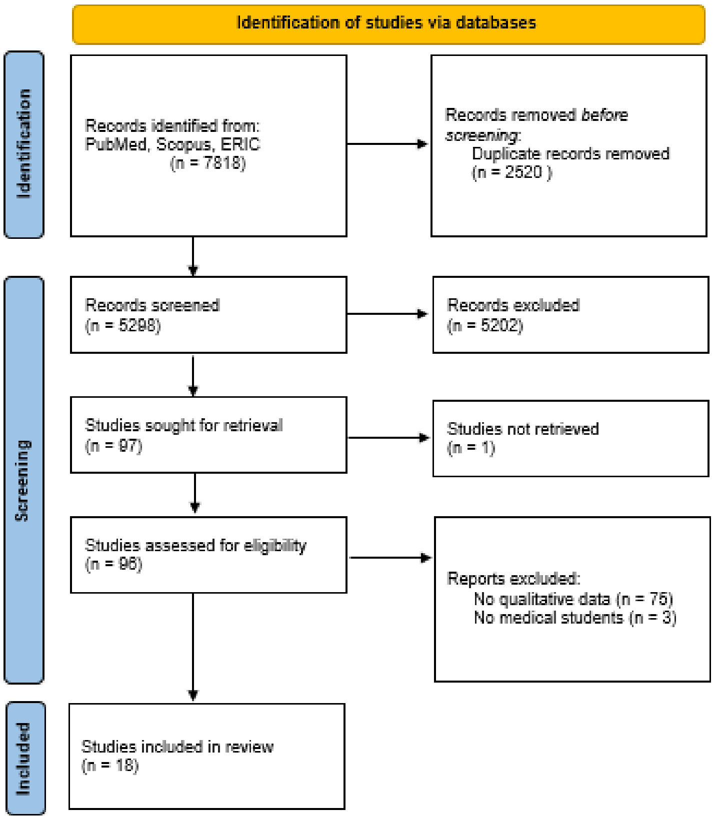

Figure 1. Research strategy for data selection.

Five articles dated from 2003 to 2019, all the rest from 2020 to 2024. Nine studies were about undergraduate students; in one study, the sample consisted of residents and teachers; one was about simulated patients; and seven were reviews or commentaries with significant discussion, making explicit the authors' opinion.

REFERENCE	TITLE	TYPE OF STUDY	TYPE OF AI	SAMPLE	TOPIC	LEARNING OUTCOME
(Asghar et al., 2022)	The potential scope of a humanoid robot in anatomy education: A review of a unique proposal.	Commented Review	Artificial Intelligence; Machine Learning; Social Robotics; Intelligent system; Automation; Computer;	N/A	Anatomy	Knowledge of Anatomy
(Bakshi et al., 2021)	The era of artificial intelligence and virtual reality: Transforming surgical education in ophthalmology.	Commented Review	Artificial Intelligence; Virtual Reality; Hologram; Convolutional Neural Network; Computer Vision; Machine Learning; Deep Learning;	N/A	Educational uses and industry role	Generic
(Blacketer et al., 2021)	Medical student knowledge and critical appraisal of machine learning: A multicentre international cross-sectional study	Mixed-Methods Survey	Machine Learning;	245 last year students from Australia, USA, New Zealand	Opinions on AI	Generic
(Chan & Zary, 2019)	Applications and Challenges of Implementing Artificial Intelligence in Medical Education: Integrative Review.	Commented Review	Artificial Intelligence; Artificial Neural Network; Big Data; Chatbot; Deep Learning; Machine Learning; Neural Network; Support Vector Machine; Virtual Machine;	N/A	Educational uses	Generic
(Duong et al., 2019)	Artificial intelligence for precision education in radiology.	Commented Review	Artificial Intelligence; Deep Learning; Decision Support System, Clinical; Picture Archiving and Communication System;	N/A	Educational uses	Radiology Skill
(Ejaz et al., 2022)	Artificial intelligence and medical education: A global mixed-methods study of medical students' perspectives.	Mixed-Methods	Artificial Intelligence; Deep Learning; Machine Learning;	128 students from 48 Countries	Opinions on educational uses	Generic
(Gilson et al., 2023)	How Does ChatGPT Perform on the United States Medical Licensing Examination (USMLE)? The Implications of Large Language Models for Medical Education and Knowledge Assessment.	Commented Quantitative Study	Artificial Intelligence; Chatbot; Conversational agent; Generative Pre-trained Transformer (GPT); Natural Language Processing (NLP); Large Language Models (LLM)	N/A	Questions from the UMLS exam	Generic
(Goldenberg, 2024)	Surgical Artificial Intelligence in Urology: Educational Applications.	Commented Review	Artificial Intelligence; Surgical Robot;	N/A	Educational uses	Surgical Skill
Hayasaka et al., 2018	Expectations for the Next Generation of Simulated Patients Born from Thoughtful Anticipation of Artificial Intelligence-Equipped Robot.	Interview	Artificial Intelligence; Patient simulation; Robot;	5 simulated patients	History taking	Clinical Competence
(Holderried et al., 2024)	A Generative Pretrained Transformer (GPT)–Powered Chatbot as a Simulated Patient to Practice History Taking: Prospective, Mixed Methods Study.	Mixed-Methods	Chatbot; Generative Pre-trained Transformer (GPT); Large Language Models (LLM)	26 students from 2-5 year	History taking	Clinical Competence

(Hudon et al., 2024)	Using ChatGPT in Psychiatry to Design Script Concordance Tests in Undergraduate Medical Education: Mixed Methods Study.	Mixed-Methods Survey	Artificial Intelligence; Generative Artificial Intelligence; Generative Pre-trained Transformer (GPT);	102 teachers and residents	Creation of script concordance test	Clinical Competence
(Jebreen et al., 2024)	Perceptions of undergraduate medical students on artificial intelligence in medicine: Mixed-methods survey study from Palestine.	Mixed-Methods Survey	Artificial Intelligence; Deep Learning; Machine Learning; Neural Network;	349 students from Palestine	Perceptions on educational uses	Clinical Reasoning
(Li et al., 2024)	Large language models and medical education: A paradigm shift in educator roles.	Narrative Review	Artificial Intelligence; Deep Learning; Large Language Models (LLM); Machine Learning; Computer tutor; Simulation training;	N/A	Teacher's role	Generic
(Liu et al., 2022)	Perceptions of US Medical Students on Artificial Intelligence in Medicine: Mixed Methods Survey Study.	Mixed-Methods Survey	Artificial Intelligence;	390 students from USA	Perceptions on educational uses	Generic
(Michael et al., 2003)	Learning from a Computer Tutor with Natural Language Capabilities. Interactive Learning Environments	Mixed-Methods	Computer tutor;	50 students from 1st year	Physiology of cardiac baroceptors reflex	Qualitative Reasoning
Persad et al., 2016	A novel approach to virtual patient simulation using natural language processing	Interview	Virtual Patient;	5 students from 3rd year	VP case of paediatric bloody diarrhoea	Clinical Reasoning
Pucchio et al., 2022	Exploration of exposure to artificial intelligence in undergraduate medical education: A Canadian cross-sectional mixed-methods study.	Mixed-Methods Survey	Artificial Intelligence;	486 students from Canada	Opinions on educational uses	Generic
Tsopra et al., 2023	Putting undergraduate medical students in AI-CDSS designers' shoes: An innovative teaching method to develop digital health critical thinking.	Case Study Mixed-Methods	Artificial Intelligence; Digital Health;	20 students from 4th year	Designing an AICDSS for solving a medical issue	Digital Health Skills and Critical Thinking

Table 2. List of the articles selected for the analysis, sorted for first author name.

Five themes emerged from this meta-ethnography, with both reciprocal and refutational position:

1) Convenience.

2) Resemblance.

3) Critical understanding.

4) Dialogic and social nature.

5) Multidisciplinarity.

They shed light on the different but interconnected facets of the use of AI in medical education. The results report the perceptions of students and teachers, based on an analysis of qualitative studies and narrative reviews, and they answer the first two research questions listed in the Introduction. Overall, they were divided into 11 sub-themes. Themes and sub-themes are illustrated with referenced quotes from the selected articles.

1. Convenience

Under this theme, all the concepts related to the effect of using AI techniques were gathered. Three sub-themes emerged, related to time-effectiveness, enjoyment and perceived educational effect, and a specific focus on the development of clinical reasoning. All the sub-themes had both reciprocal and refutational instances.

1.1 Time-effectiveness

Time was a very debated topics in the selected articles, considering AI either a way to decrease the workload of teachers and supervisors, as well as a further burden for students:

> *AI may be time-saving and beneficial for flipped learning, which emphasizes low-level supervision, such as comparing scans over time* (Deng, 2023)
> *The supervisor can utilize AI platforms as a first pass didactic tool and save time. AI can help trainees learn simpler concepts and indicate which topics trainees may need attending physicians to teach. Integrated platforms allow supervisors to monitor progress of trainees* (Duong, 2020).

Nevertheless, some student expressed their concerns on AI as something new to be added, stating that:

> *I feel as though the pre-clerkship curriculum is already pretty packed with a lot of very relevant things… There are things I'm going to need to know as a clerk that I don't feel like we've adequately covered* (Pucchio, 2024).

From my point of view, studying AI courses may hinder the comprehension and understanding of other courses related to medicine. For example, I need 3 hours a day to study the AI course, and this reduces the time required to study medical courses such as anatomy and physiology. This means that I will waste time studying topics that have no practical application in my country (Jebreen et al., 2024)
A case study discusses what skills medical undergraduate students should acquire to understand Digital Health (DH), enhancing their critical thinking by simulating they were AI-CDSS designers (Tsopra et al., 2023), thus recalling the role-playing methodology for knowledge transfer.

> *Some students explained that they better understood the issues related to AI and CDSS in medicine afterward. Others added that they felt more critical regarding technologies but also more confident in AI-CDSS use. However, some students complained about difficulty with overbooked agendas, especially as the AI-CDSS program was considered of lower priority than other clinical courses* (Tsopra et al., 2023).

This article is a step forward for the update of medical curricula, since the current scenario is headed to Digital Medicine (DM) and Digital Therapeutics (DTx).

> *Finally, 53 % of students would like to continue to increase their knowledge in digital health; 47 % would be interested in doing an internship in the medical informatics department; and 60 % would like to be involved in research projects on AI-CDSS* (ibidem).

For instance, Table 1 shows that a classification of DH/DM/DTx terms in the MeSH Browser is still missing, despite scientific literature in this field is growing, meaning that further knowledge is needed.

1.2 Perception of engagement and learning

In the examined articles, many instances of reciprocal translation were present on AI as a tool to produce personalised learning, engagement, and amusement, expressed both by teachers in their comments and by students in reported quotes.

AI's personalised learning paths cater to the unique needs of each student, enhancing engagement and significantly boosting collaborative skills and motivation through PBL-based models (Li, 2024).

I enjoyed using this program – it was fun (Michael, 2007).

On the other hand, previous knowledge and background can make the difference in Australian medical students' engagement with ML and AI topics, despite they are interested in learning more about it:

[…] medical students from a diverse array of backgrounds may struggle to interpret and critically appraise elements of medical ML research. Students endorse statements regarding the relevance of ML to their future practice and indicate interest in learning further about the topic (Blaketer, 2021).

Subject learning priorities and opportunities can change if students live in highly conflictual Countries:

[…] technological developments in the medical field must be kept updated because we live in Gaza and the economic situation in general is deteriorating, so it is difficult for a medical student to participate in any scientific event as long as there is no funding. If I have the opportunity to obtain funding from a donor, I will definitely take intensive training related to applying AI tools in my major and engaging with my peers from other countries (Jebreen et al., 2024).

A step forward on generative AI is achieved with a Computer Tutor (CT) trained to perform NLP-driven output via pre-defined sentences on the physiology of barioceptive cardiac reflexes to carry out qualitative reasoning by medical students at the first year. Students found the CT can be useful to learn some topics; still, the software was not able to provide constructive negative feedback nor to explain deeper physiological mechanisms:

> *The students' written comments were often quite enthusiastic about the program and its value as a learning resource. Such comments obviously parallel their numerical ratings. Their major criticism was leveled at the quality of the explanations that CIRCSIM-Tutor delivers when critiquing wrong answers. The program cannot yet describe the physiology underlying the causal relationships being tutored* (Michael et al., 2003).

1.3 Development of clinical reasoning

The usefulness of AI-based tools and methods has been one of the most controversial topics. From one side, AI was appreciated for its ability to be a guide in the development of clinical reasoning, providing students with hints and evidence-based support, like in the following quote:

> *AI teaches students how to use a diagnostic problem-solving method to address problems relating to diseases. Also, it offers an intelligent simulation tool or surgical aid* (Jebreen, 2024).

> *Clinical reasoning skills were exercised to a far greater degree than in other available VP simulators* (Persad, 2016).

Refutational instances were variously related to a possible failure of the algorithm, especially in training graduate students:

When guiding diagnosis, AI may not necessarily improve human interpretation. In the application of mammography, some forms of computer-assisted diagnosis can increase false-positive and false-negative errors (Duong, 2022).

Failure can be related also to ethical issues, derived from gender and social biases contained int the text that trained the neural network.

One significant concern surrounding the use of AI in surgical education is the role of bias in algorithm creation. Flawed data sets and inequitable access to care can perpetuate biases in predictive ML algorithms, leading to skewed predictions based on race or gender (Goldenberg, 2024).

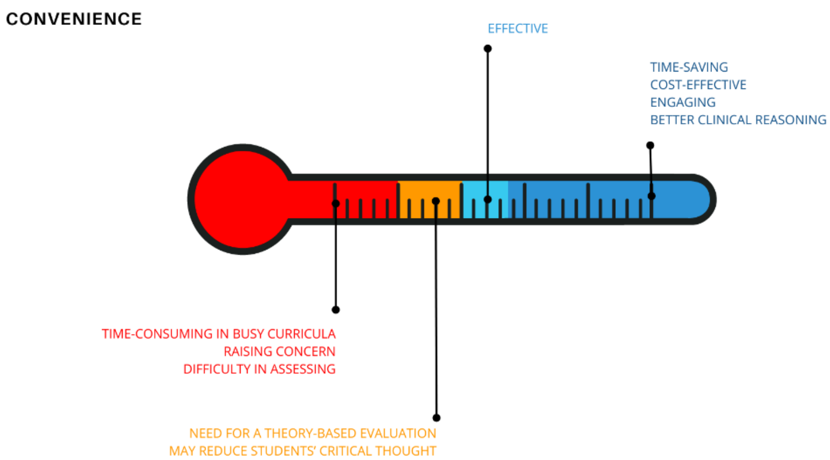

Figure 2. Line of arguments on the "Convenience" theme.

2. Resemblance

This theme gathered all the concepts related to the extent to which an interaction with an AI system resembled the interaction with concrete reality. The sub-themes were about the resemblance of the produced clinical cases for case-based learning, the production of texts and the empowerment of robotics. Such last sub-theme had only refutational instances.

2.1 Quality of generated clinical cases

Many articles reported reflections and comments on the production on clinical cases by AI systems (especially LLMs, chatbots, and GPTs) to support case-based learning. The production of cases for learning is a time-consuming activity, and if the process could be automated this would help teachers and small group supervisors. AI systems were particularly appreciated when integrated in a computer-based learning environment to play Virtual Patients (VP) simulations.

Holderried and colleagues (2024) found that chatbot could simulate a patient to help medical students train their clinical competencies, especially history taking. However, the most common complaint with AI-generated clinical cases is that they are too simplistic and stereotypical to engage students. Even in this sub-theme we found both reciprocal and refutational translations, as in the following exemplary quotes.

> *All participants described the programme and case design as 'excellent' and a 'significant improvement' over VP* (Persal, 2016).

> *It would be interesting to add more challenging prompts as they [clinical cases] tend to be very simplistic and poorly represent complex clinical cases as they are very stereotypical to what is found in the DSM-5* (Hudon, 2024).

2.2 Quality of generated texts

This sub-theme gathers the instances dealing with the quality of the texts generated by the AI systems in answering to the prompts of the users. All instances were positive and reciprocal, valuing the answer rich and an opening to further questioning and study, as in the following example:

> *We found value in plugging questions into ChatGPT and engaging with follow-up dialogue, because it could unearth context relevant to the question and effectively trigger recall for specific lectures that taught the material relevant to the problem. This suggests that the context that ChatGPT provides in an initial answer could*

open the door for further questioning that naturally digs into the foundational knowledge required to justify the given underlying medical reasoning (Gilson, 2023).

Gilson et al. (2023) commented the implications of using LLM in medical education, by comparing the performance of 3 chatbots in succeeding the United States Medical Licensing Examination [USMLE]. On the other hand, the ability of AI systems to understand and produce rich text was also appreciated to overcome one limitation of the current VP systems:

VP simulators also limit students to the selection of options from lists, which we believe limits further development and application of clinical reasoning skills […] All felt that the free-text input provided a more personalized learning experience (Persad, 2016).

2.3 Effectiveness in empowering robots

Only two of the selected articles mentioned the use of AI to empower robots as a teacher or standardized patient, and both expressed concerns for the present performance of these systems, that lacks the emotional component of real human communication:

[…] robot teachers are more rapid and repetitive, whereas human teachers are more socially behaviored, empathetic and naturally creative (Ashgar, 2022)
[…] concerns as to how much a robot SP [Standardized Patient] can deal with human diversity. In addition, SPs had "an expectation of emotional intelligence (EI) (Hayasaka, 2020).

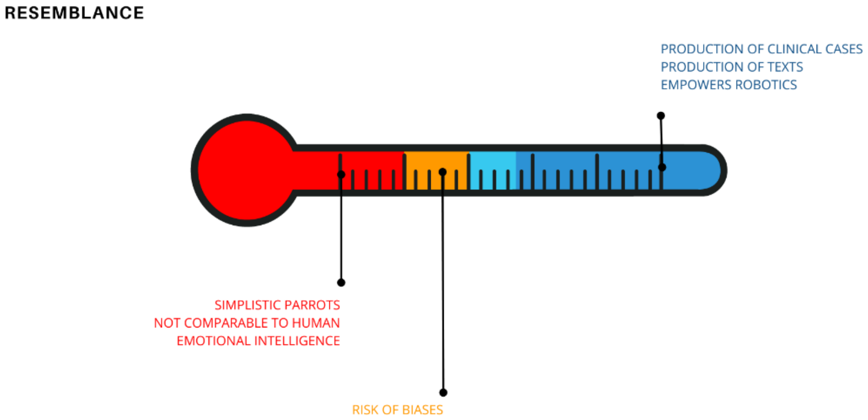

Figure 3. Line of arguments on the "Resemblance" theme.

3. Critical understanding

This theme reflected the challenges of understanding in depth the implications of the use of AI in medical education. The sub-themes were about the best way to integrate the AI-based methods into the curriculum, the need of a new vision of the future, the threats related to the use of AI.

3.1 Integration into the curriculum

Although we excluded the articles only dealing with the composition of a curriculum on AI in a medical degree course, some instances of the selected articles pointed at the issue of integrating this topic in the curriculum.

> *The majority of students who participated in the study stressed that AI-related teaching needs to be incorporated into the core medical curriculum* (Ejaz, 2022).

A review diving into predictive modeling and NN also stated that:

> *Without a structured curriculum, it is difficult to select a knowledge base for the AI system, which may, in part, explain the lower prevalence of AI use in CME as compared to that in medical undergraduates* (Chan & Zharv, 2019).

This instance, together with the already reported quote about the fear of an addition to an already "pretty packed" pre-clerkship curriculum, expresses the ambivalence of the students' feeling about something new, still unexplored. Liu and colleagues proposed an AI curriculum only in postgraduate education:

> *AI is interesting and will be incorporated into medical practices more and more through our lifetimes. But I don't think it's necessary for pre-clinical medical students to understand how to utilize AI. I feel that it's more important to learn about AI in residency versus medical school* (Liu, 2022).

Some researchers report a futuristic vision of the potential of AI and Virtual Reality (VR) in ophthalmology, including holograms and virtual simulations, believing that some standardization is needed:

> *In the process, the challenges of accurately rendering the orbit, interaction of instruments with tissue surfaces, and the surgical field must be addressed to yield meaningful educational experiences for learners. As the utility of VR within healthcare continues to grow, the development of standardised guidelines for clinical validation (as with AI) will also be essential* (Bakshi et al, 2021).

3.2 Deep understanding of the implications of AI

Beyond the specific topic of curricular integration, many reciprocal instances discussed the need of a deep understanding of the implications of the use of AI in teaching and learning. The following quote lists several warnings:

> *When utilizing LLMs in medical education, it is vital to be aware of information accuracy, potential biases, the risks of over-reliance, academic misconduct, and copyright infringement* (Li, 2024).

AI is perceived as a still fluid domain, in rapid evolution. This is why the results of research in this domain should be considered:

The critical appraisal of medical ML research is an area of ongoing development [...] in particular with respect to model performance analysis and stage of model development (Blacketer, 2023).

A critical understanding of AI is mandatory also to support a change in the teacher's role:

> *Shift from traditional teaching roles to becoming "information management experts". To guide students in the wise use of LLMs, prevent overreliance, eliminate biases, and cultivate critical thinking, teachers should act as "learning navigators". Additionally, to foster students' sense of responsibility and research integrity, educators must also take on the role of "academic integrity guardians"* (Li, 2024).

3.3 Other threats of the use of AI

Many threats were listed in the selected articles, not only those connected to student's misconduct or bias in AI responses, as already reported in other sub-themes above. AI was perceived as threatening the occupational rate of physicians:

> *Maybe there'll be less jobs for physicians in certain fields that AI is more applicable to, like radiology or pathology* (Pucchio, 2022).

Researchers reported a vague sense of threat, due to a poor understanding of what AI is and can do:

> *I don't have a great understanding of what AI is capable of or what it even is. It definitely could have implications in different fields of healthcare, and I think we all need to be prepared in the future of medicine* (Pucchio et al, 2022).

One article focused on medical imaging (PACS) as the predictive model can enhance CDSS, finding that:

When guiding diagnosis, AI may not necessarily improve human interpretation. In the application of mammography, some forms of computer-assisted diagnosis can increase false-positive and false-negative errors (Duong, 2022).

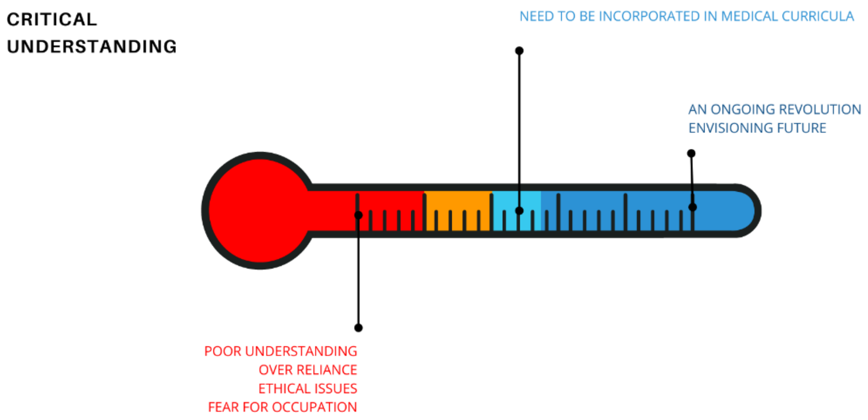

Figure 4. Line of arguments on the "Criticalunderstanding" theme.

4. Dialogic and social nature

The fourth theme tackled the highly sensible topic of student-system interaction. Two sub-themes emerged, the first one had reciprocal instances, related to the effective educational value, while the second one was refutational, about the poor quality of emotional interaction.

4.1 Educational value of the interaction

Most of the selected articles considered LLM systems and explored their ability to understand free-text questions and to generate meaningful text in response. It is therefore not surprising that the main educational effect is related to the LLM's ability to engage students in a dialogical interaction, even to the point of considering its role as a supervisor.

An aspect of small group education that is often beneficial is the ability of students to test ideas off of each other and receive feedback. With its dialogic

> *interface, ChatGPT is able to provide many of these same benefits for students when they are studying independently [...] One potential use case to highlight for the use of ChatGPT is as an adjunct or surrogate for small (peer) group education* (Gilson, 2023).

> *A common theme in many of the responses emphasising the importance of the information being provided in a comprehensible fashion delivered at the level of clinicians and focussed on the relevance to clinicians* (Blacketer, 2021).

Asghar et al. (2022) refer to "humanoid robots" (robots that resemble human bodies, but not necessarily having some sort of AI) as useful tools anatomy education. In this review, human robots are also Social Robots (SR), since they are conceived as potentially substitutes of anatomy teachers. In addition to the limited ability to simulate empathy (see sub-theme 4.2 below), there have been some cases of failure in dialogue:

> *[...] a negligible number of students stating that the chatbot did not recognize their inputs* (Holderried, 2024).

Despite that, it is known that AI-driven technologies might support teachers in conveying some forms of contents if teachers are considered as merely "knowledge transmitters", thus flattening the pedagogical domains of learning and growing as human beings. This conception of medical pedagogy conveys a deep "mechanistic" representation of education:

> *The role of educators is undergoing a profound metamorphosis, and shortly, they are anticipated to transition from conventional knowledge transmitters to multi-dimensional roles. LLMs present opportunities to medical education by infusing traditional teacher roles with new connotations. As knowledge conveyors, educators are no longer mere unilateral transmitters of information but utilise LLMs*

as potent auxiliary tools to bolster teaching efficiency and enrich and engage the educational content, thereby magnifying their teaching impact (Li et al., 2024).

4.2 Emotional intelligence and empathy

A major current limitation is the lack of an emotional component in LLM's response in a dialogue, both in showing understanding of the students' emotional situation and in providing meaningful feedback to an emotional expression:

Although LLMs can be trained to express empathy similarly to humans, and sometimes even better, they lack the ability to provide genuine emotional experiences [...] There were a few instances when GPT provided implausible responses, however, and our content analysis revealed a tendency toward socially desirable answers (Holderried, 2024).

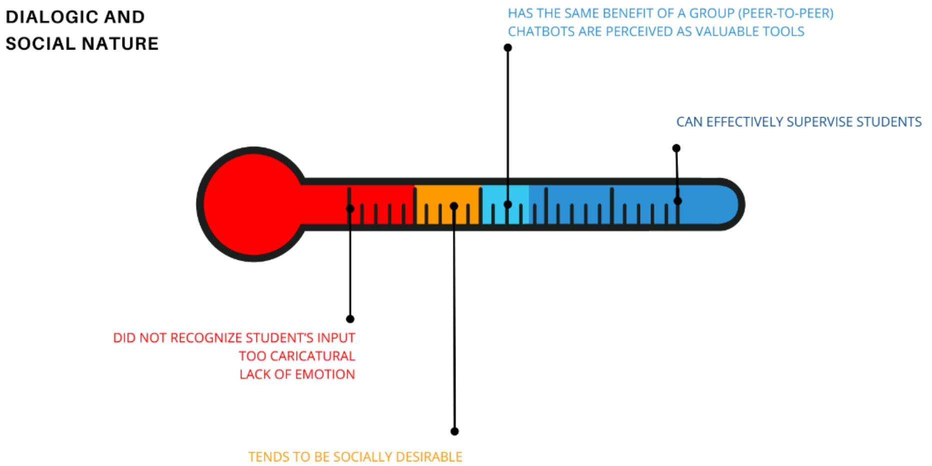

Figure 5. Line of arguments on the "Dialogic and social nature" theme.

5. Multidisciplinarity

The last theme brings together all the reciprocal instances on how the use of AI in medical training can foster interprofessional collaboration. Likewise, the adoption of AI in training requires interprofessional collaboration to be fully effective:

An essential aspect of developing an AI system is the need for a multidisciplinary team, including educational experts, data scientists for management of the large pool of data, physicians for ensuring the clinical relevance, and accuracy of the AI system (Chan, 2020).

Many of these avenues toward the advancement of surgical education are obstructed only by our ability to collaborate across institutional and disciplinary boundaries (Goldenberg, 2024).

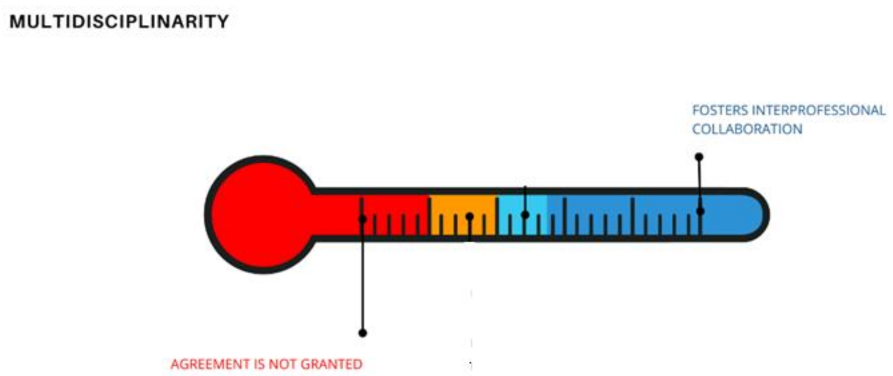

Figure 6. Line of arguments on the "Multidisciplinarity" theme.

DISCUSSION

Except for the "Multidisciplinarity" theme, all the other four themes grouped articles with both reciprocal and refutational translations, demonstrating how controversial this topic still is. Disorientation may be due to poor understanding and awareness on these topics, and this is a mirror of what the society is facing (including policy makers and other stakeholders). Therefore, it is important to go beyond the enthusiasms and fears about these integrations without neglecting the necessity to start constructing and validate novel approaches to education, both at an individual and at a societal level.

Our meta-ethnographic synthesis explored the perception of teachers and students on the use of AI in medical education, highlighting the need for digital literacy on the main definition of technologies (at least), to create a common ground for understanding. To this aim, we reported in Table 1 the principal definitions but some of them are still missing in the MeSH Browser, meaning that some topics should be consolidated more in the research field. Moreover, we suggest that customized educational trainings on AI topics and applications could help gain more understanding of their true potential, going beyond the hype of enthusiasms and fears.

To conceptualize the current situation of AI adoption in medical education, we used Rogers' model of innovation diffusion (Rogers, 1983). Innovation decision is the process by which an individual (or other decision-making unit) goes from the first knowledge of an innovation to the formation of an attitude towards the innovation, the decision to adopt or reject it, the implementation of the new idea, and the confirmation of that decision. Knowledge occurs when a possible decision maker is exposed to the existence of the innovation and understands how it works. Persuasion occurs when an individual (or other decision-making unit) forms a favourable or unfavourable attitude towards the innovation. According to Rogers' model, we believe that the introduction of AI in medical education is still in the first phase of "knowledge," moving slowly only in selected places towards the "persuasion" phase.

Our third research question inquired the role of context in promoting a positive or negative perception and attitude towards the use of AI in medical training. In our view, an answer to this question comes from the fifth theme, which unanimously showed the perceived need for a multidisciplinary environment as a context for developing the application of AI in medical education.

In their review, Issroff and Scanlon (2002) discussed the role of theoretical frameworks in the use of technology in education; they explored the role of theories in the current practice of educational technology, concluding that the use of technologies such as AI requires "a multilevel approach to understanding complex learning situations" (Issroff & Scanlon, 2002 p. 10) and the contamination of traditional educational theories with technological and social theories, such as Activity Theory or Distributed Cognition. Therefore, a second possible answer to the third research question is that the use of AI and other technologies in medical education requires a new type of educational design with a circular relationship between objectives, methods, tools, and results because tools and methods of AI can no longer be regarded as passive elements to be used by educators and learners.

CONCLUSION

The European Commission (2022) supports plans of education and training, including vocational education and adult learning institutions, to provide digital competencies while reducing the existing gaps, as well as to improve the quality and equity in education and training.

Current challenges can be synthetized as follows:

1) Understanding the many interpretations of data with AI, its real potential and threat. Integrating ethics as a key perspective within data-driven practices is an essential part of data literacy.
2) Acknowledging the real maturity of the AI-driven and digital solutions for medical purposes, considering that regulatory flaws still occur, and strong evidence-based best practices are needed before their usability in the clinical field, since the final user will always be the patient.
3) Updating the medical education agenda to enhance AI and digital competences. This effort might be extended to undergraduate and graduate students, to professionals during Continuing Medical Education (CME), and to faculty of biomedical sciences as well.
4) Assessing and validating the user experience and acceptance of these technologies, especially about their access, usability, and ease to be understood.
5) Stressing the ethical concerns about the delegation of decision making and the interpretation of results, especially about choices that concern health and life.
6) Recognizing that phenomenological differences in time, spaces, relationships, languages, meanings, and therefore experiences do occur within several “realities” (physical, digital, virtual, augmented, phygital).

Since these challenges are affecting medical education, to our experience, it is important to investigate the sustainability of a literacy convergence in order to teach best practices and the necessary specialistic and transversal skills in the domain of AI, while valorizing hidden skills of students that might be exposed to AI without knowing (e.g., Google Ads, semi-automated cars, chatbots, videogames, etc.) to bridge some existing gaps that open disparities in educational opportunities and access among medical students.

In this "Liquid Education" era, how long the literacy gap will be sustainable? What will be the future of the educative "offer"? How do we, as educative professionals, should adapt our agenda as well? Still, can we leave medical education become a "market" in which tutors and teachers can be "substituted" or can we preserve the pedagogical value and meaning of teaching and learning?

REFERENCES

An, P., & Wang, Z. (2024). Application value of an artificial intelligence-based diagnosis and recognition system in gastroscopy training for graduate students in gastroenterology: A preliminary study. *Wiener Medizinische Wochenschrift*, *174*(9), 173–180. https://doi.org/10.1007/s10354-023-01020-w

Asghar, A., Patra, A., & Ravi, K. S. (2022). The potential scope of a humanoid robot in anatomy education: A review of a unique proposal. *Surgical and Radiologic Anatomy: SRA*, *44*(10), 1309–1317. https://doi.org/10.1007/s00276-022-03020-8

Bakshi, S. K., Lin, S. R., Ting, D. S. W., Chiang, M. F., & Chodosh, J. (2021). The era of artificial intelligence and virtual reality: Transforming surgical education in ophthalmology. *The British Journal of Ophthalmology*, *105*(10), 1325–1328. https://doi.org/10.1136/bjophthalmol-2020-316845

Blacketer, C., Parnis, R., B Franke, K., Wagner, M., Wang, D., Tan, Y., Oakden-Rayner, L., Gallagher, S., Perry, S. W., Licinio, J., Symonds, I., Thomas, J., Duggan, P., & Bacchi, S. (2021). Medical student knowledge and critical appraisal of machine learning: A multicentre international cross-sectional study. *Internal Medicine Journal*, *51*(9), 1539–1542. https://doi.org/10.1111/imj.15479

Chan, K. S., & Zary, N. (2019). Applications and Challenges of Implementing Artificial Intelligence in Medical Education: Integrative Review. *JMIR Medical Education*, *5*(1), e13930. https://doi.org/10.2196/13930

Duong, M. T., Rauschecker, A. M., Rudie, J. D., Chen, P.-H., Cook, T. S., Bryan, R. N., & Mohan, S. (2019). Artificial intelligence for precision education in radiology. *British Journal of Radiology*, *92*(1103), 20190389. https://doi.org/10.1259/bjr.20190389

Ejaz, H., McGrath, H., Wong, B. L., Guise, A., Vercauteren, T., & Shapey, J. (2022). Artificial intelligence and medical education: A global mixed-methods study of medical students' perspectives. *Digital Health*, *8*, 20552076221089099. https://doi.org/10.1177/20552076221089099

European Commission (2022). Digital Education. https://education.ec.europa.eu/focus-topics/digital-education/action-plan/action-5 Last consultation: 04/12/2024

Fazlollahi, A. M., Bakhaidar, M., Alsayegh, A., Yilmaz, R., Winkler-Schwartz, A., Mirchi, N., Langleben, I., Ledwos, N., Sabbagh, A. J., Bajunaid, K., Harley, J. M., & Del Maestro, R. F. (2022). Effect of Artificial Intelligence Tutoring vs Expert Instruction on Learning Simulated Surgical Skills Among Medical Students. *JAMA Network Open*, *5*(2), e2149008. tps://doi.org/10.1001/jamanetworkopen.2021.49008

Gilson, A., Safranek, C. W., Huang, T., Socrates, V., Chi, L., Taylor, R. A., & Chartash, D. (2023). How Does ChatGPT Perform on the United States Medical Licensing Examination (USMLE)? The Implications of Large Language Models for Medical Education and Knowledge Assessment. *JMIR Medical Education*, *9*, e45312. https://doi.org/10.2196/45312

Goldenberg, M. G. (2024). Surgical Artificial Intelligence in Urology: Educational Applications. *Urologic Clinics of North America*, *51*(1), 105–115. https://doi.org/10.1016/j.ucl.2023.06.003

Holderried, F., Stegemann–Philipps, C., Herschbach, L., Moldt, J.-A., Nevins, A., Griewatz, J., Holderried, M., Herrmann-Werner, A., Festl-Wietek, T., & Mahling, M. (2024). A Generative Pretrained Transformer (GPT)–Powered Chatbot as a Simulated Patient to Practice History Taking: Prospective, Mixed Methods Study. *JMIR Medical Education*, *10*, e53961. https://doi.org/10.2196/53961

Hudon, A., Kiepura, B., Pelletier, M., & Phan, V. (2024). Using ChatGPT in Psychiatry to Design Script Concordance Tests in Undergraduate Medical Education: Mixed Methods Study. *JMIR Medical Education*, *10*(1), e54067. https://doi.org/10.2196/54067

Issroff, K., & Scanlon, E. (2002). Educational Technology: The influence of theory. *Journal of Interactive Media in Education*, *6*. https://oro.open.ac.uk/44/

Jackson, P., Ponath Sukumaran, G., Babu, C., Tony, M. C., Jack, D. S., Reshma, V. R., Davis, D., Kurian, N., & John, A. (2024). Artificial intelligence in medical education—Perception among medical students. *BMC Medical Education*, *24*(1), 804. https://doi.org/10.1186/s12909-024-05760-0

Jafri, L., Farooqui, A. J., Grant, J., Omer, U., Gale, R., Ahmed, S., Khan, A. H., Siddiqui, I., Ghani, F., & Majid, H. (2024). Insights from semi-structured interviews on integrating artificial intelligence in clinical chemistry laboratory practices. *BMC Medical Education*, *24*(1), 170. https://doi.org/10.1186/s12909-024-05078-x

Jebreen, K., Radwan, E., Kammoun-Rebai, W., Alattar, E., Radwan, A., Safi, W., Radwan, W., & Alajez, M. (2024). Perceptions of undergraduate medical students on artificial intelligence in medicine: Mixed-methods survey study from Palestine. *BMC Medical Education*, *24*(1), 507. https://doi.org/10.1186/s12909-024-05465-4

Li, Z., Li, F., Fu, Q., Wang, X., Liu, H., Zhao, Y., & Ren, W. (2024). Large language models and medical education: A paradigm shift in educator roles. *Smart Learning Environments*, *11*(1), 26. https://doi.org/10.1186/s40561-024-00313-w

Liu, D. S., Sawyer, J., Luna, A., Aoun, J., Wang, J., Boachie, Lord, Halabi, S., & Joe, B. (2022). Perceptions of US Medical Students on Artificial Intelligence in Medicine: Mixed Methods Survey Study. *JMIR Medical Education*, *8*(4), e38325. https://doi.org/10.2196/38325

Mamedov, O., Tumanyan, Y., Ishchenko-Padukova, O., & Movchan, I. (2018). Sustainable economic development and post-economy of artificial intelligence. *Entrepreneurship and Sustainability Issues*, *6*(2), 1028–1040. https://doi.org/10.9770/jesi.2018.6.2(37)

Michael, J., Rovick, A., Glass, M., Zhou, Y., & Evens, M. (2003). Learning from a Computer Tutor with Natural Language Capabilities. *Interactive Learning Environments*, *11*(3), 233–262. https://doi.org/10.1076/ilee.11.3.233.16543

Moldt, J.-A., Festl-Wietek, T., Fuhl, W., Zabel, S., Claassen, M., Wagner, S., Nieselt, K., & Herrmann-Werner, A. (2024). Assessing AI Awareness and Identifying Essential Competencies: Insights From Key Stakeholders in Integrating AI Into Medical Education. *JMIR Medical Education*, *10*, e58355. https://doi.org/10.2196/58355

Nobles, A. L., Leas, E. C., Caputi, T. L., Zhu, S.-H., Strathdee, S. A., & Ayers, J. W. (2020). Responses to addiction help-seeking from Alexa, Siri, Google Assistant, Cortana, and

Bixby intelligent virtual assistants. *NPJ Digital Medicine*, *3*, 11. https://doi.org/10.1038/s41746-019-0215-9

Persad, A., Stroulia, E., & Forgie, S. (2016). A novel approach to virtual patient simulation using natural language processing. *Medical Education*, *50*(11), 1162–1163. https://doi.org/10.1111/medu.13197

Pucchio, A., Rathagirishnan, R., Caton, N., Gariscsak, P. J., Del Papa, J., Nabhen, J. J., Vo, V., Lee, W., & Moraes, F. Y. (2022). Exploration of exposure to artificial intelligence in undergraduate medical education: A Canadian cross-sectional mixed-methods study. *BMC Medical Education*, *22*(1), 815. https://doi.org/10.1186/s12909-022-03896-5

Tong, A., Flemming, K., McInnes, E., Oliver, S., & Craig, J. (2012). Enhancing transparency in reporting the synthesis of qualitative research: ENTREQ. *BMC Medical Research Methodology*, *12*(1), 181. https://doi.org/10.1186/1471-2288-12-181

Truong, H. M. (2016). Integrating learning styles and adaptive e-learning system: Current developments, problems and opportunities. *Computers in Human Behavior*, *55*, 1185–1193. https://doi.org/10.1016/j.chb.2015.02.014

Tsopra, R., Peiffer-Smadja, N., Charlier, C., Campeotto, F., Lemogne, C., Ruszniewski, P., Vivien, B., & Burgun, A. (2023). Putting undergraduate medical students in AI-CDSS designers' shoes: An innovative teaching method to develop digital health critical thinking. *International Journal of Medical Informatics*, *171*, 104980. https://doi.org/10.1016/j.ijmedinf.2022.104980

Williamson, B. (2024). The Social life of AI in Education. *International Journal of Artificial Intelligence in Education*, *34*(1), 97–104. https://doi.org/10.1007/s40593-023-00342-5

Xu, X., Chen, Y., & Miao, J. (2024). Opportunities, challenges, and future directions of large language models, including ChatGPT in medical education: A systematic scoping review. *Journal of Educational Evaluation for Health Professions*, *21*, 6. https://doi.org/10.3352/jeehp.2024.21.6