

Utilização de tecnologias de inteligência artificial no desenvolvimento de competências comunicacionais clínicas em estudantes pré-graduados da área da saúde: protocolo de *scoping review*

The use of artificial intelligence technologies in the development of clinical communication skills among undergraduate health professions students: a scoping review protocol

Mariana Lima^{1*} , Lígia Lima² , Celeste Bastos² , Teresa Martins² 

¹Unidade Local de Saúde de Santo António, Porto, Portugal

²RISE-Health, Escola Superior de Enfermagem, Universidade do Porto, Porto, Portugal

*Autor correspondente/Corresponding author: mariana.vale.lima@gmail.com

Recebido/Received: 23-03-2026; Revisto/Revised: 02-06-2026; Aceite/Accepted: 16-06-2026

Resumo

Introdução: A comunicação clínica é uma competência central na formação em saúde, associada à qualidade dos cuidados e à segurança do doente. Em paralelo, a rápida incorporação de tecnologias de inteligência artificial (IA) em contextos educativos tem produzido resultados inconsistentes quanto a modelos pedagógicos, instrumentos de avaliação e desfechos reportados no treino de competências comunicacionais clínicas. **Objetivo:** Mapear a utilização de tecnologias educativas baseadas em IA no desenvolvimento de competências comunicacionais clínicas em estudantes pré-graduados da área da saúde. **Material e Métodos:** Será conduzida uma *scoping review* segundo a metodologia do *Joanna Briggs Institute*, com questão estruturada pelo modelo PCC e reporte PRISMA-ScR. Serão pesquisadas as bases MEDLINE, CINAHL, Scopus, Web of Science, ERIC e PsycINFO, complementadas por literatura cinzenta. A seleção dos estudos será realizada em duas fases (título/resumo e texto integral) por dois revisores independentes, com resolução de divergências por consenso ou por terceiro revisor. **Resultados:** Os dados serão sintetizados de forma descritiva e analítica, recorrendo a tabelas e matrizes de evidência que relacionem tecnologia, componentes pedagógicas e avaliação, sendo os desfechos organizados segundo a Pirâmide de *Miller* e, quando aplicável, segundo o nível 1 do modelo de *Kirkpatrick*. **Discussão:** A revisão permitirá caracterizar abordagens educativas, identificar lacunas e clarificar a natureza da evidência disponível. **Conclusão:** Espera-se contribuir para a sistematização do conhecimento e apoiar decisões curriculares e futuras investigações neste domínio.

Palavras-chave: Inteligência Artificial; Comunicação; Competência Clínica; Educação em Saúde; Estudantes de Saúde; Revisão de Escopo.

Abstract

Introduction: Clinical communication is a core competency in health professions education, closely associated with quality of care and patient safety. In parallel, the rapid integration of artificial intelligence (AI) technologies in educational contexts has produced inconsistent results regarding pedagogical models, assessment instruments, and reported outcomes in training clinical communication skills. **Objective:** To map the use of AI-based educational technologies in the development of clinical communication skills among undergraduate health professions students. **Methods:** A scoping review will be conducted following the Joanna Briggs Institute methodology, with the research question structured according to the PCC framework and reported in line with PRISMA-ScR guidelines. The databases MEDLINE, CINAHL, Scopus, Web of Science, ERIC, and PsycINFO will be searched, complemented by grey literature. Study selection will be performed in two stages (title/abstract and full text) by two independent reviewers, with disagreements resolved by consensus or by a third reviewer. **Results:** Data will be synthesized descriptively and analytically using tables and evidence matrices that relate technologies, pedagogical components, and assessment methods. Outcomes will be organized according to Miller's Pyramid and, when applicable, according to level 1 of the Kirkpatrick model. **Discussion:** This review will enable the characterization of educational approaches, identification of gaps, and clarification of the nature of the available evidence. **Conclusion:** The review is expected to contribute to the systematization of knowledge and to support *syllabus* decision-making and future research in this field.

Keywords: Artificial Intelligence; Communication; Clinical Competence; Health Education; Students, Health Occupations; Scoping Review.

1. INTRODUÇÃO

A comunicação clínica constitui um eixo estruturante da prática dos profissionais de saúde, sendo amplamente reconhecida como determinante da qualidade assistencial, da segurança do doente e dos resultados em saúde (Aktan & Khorshid, 2021; Leonard et al., 2004; Stiefel et al., 2024). Mais do que uma competência instrumental, a comunicação configura-se como um processo relacional complexo que sustenta a construção de significados partilhados, a negociação terapêutica e a tomada de decisão centrada na pessoa.

Enquanto competência profissional, assume natureza multidimensional, integrando componentes verbais, não-verbais e paralinguísticas, bem como dimensões afetivas, cognitivas e éticas (Arnold & Boggs, 2003; Kim et al., 2024). A sua expressão em contexto clínico traduz-se na capacidade de estabelecer alianças terapêuticas, interpretar necessidades explícitas e implícitas e adaptar a interação às características singulares do doente, da família e da situação clínica (Arnold & Boggs, 2003; Kim et al., 2024; Kourkouta & Papathanasiou, 2014; McCabe & Timmins, 2013).

A literatura sugere de forma consistente que falhas comunicacionais contribuem para eventos adversos, erros evitáveis, fragmentação dos cuidados e deterioração da experiência do doente, particularmente em contextos de elevada complexidade clínica e exigência emocional (Chang et al., 2025; Leonard et al., 2004; O'Daniel & Rosenstein, 2008; Stiefel et al., 2024). Assim, a competência comunicacional constitui uma dimensão estruturante da prática profissional, sustentando a segurança e a efetividade clínica.

A formação pré-graduada na área da saúde assume, por conseguinte, um papel decisivo na consolidação de padrões comunicacionais que moldarão o desempenho profissional futuro. Tradicionalmente, o ensino destas competências tem recorrido a metodologias com componente experiencial, incluindo role-play estruturado, simulação com doentes padronizados, *objective structured clinical examination* (OSCE) e *debriefing* reflexivo, entre outros (Chant et al., 2002; Goh et al., 2023). A investigação tem demonstrado que a simulação baseada em cenários clínicos contribui de forma significativa para a aquisição de competências comunicacionais clinicamente relevantes, operacionalizadas através de indicadores comportamentais mensuráveis e passíveis de avaliação no desempenho clínico (Cook et al., 2011; McGaghie et al., 2010). Sínteses sistemáticas em educação em saúde sugerem ainda que intervenções estruturadas de treino comunicacional, particularmente quando integradas com prática deliberada e feedback individualizado, tendem a associar-se a melhorias nas competências avaliadas, embora a magnitude dos efeitos varie em função do desenho pedagógico e dos instrumentos utilizados (Cook et al., 2011; Gilligan et al., 2016). Contudo, estas abordagens apresentam limitações reconhecidas, nomeadamente exigência logística elevada, custos significativos, dependência de avaliadores especializados e variabilidade interobservador na avaliação do desempenho (Dewan et al., 2024; Hodges, 2003; McGaghie et al., 2010). Adicionalmente, a necessidade de escalabilidade e personalização da aprendizagem, em contextos de crescente massificação do ensino superior, tem evidenciado

1. INTRODUCTION

Clinical communication is a fundamental component of health professionals' practice and is widely recognized as a determinant of quality of care, patient safety, and health outcomes (Aktan & Khorshid, 2021; Leonard et al., 2004; Stiefel et al., 2024). More than an instrumental competency, clinical communication constitutes a complex relational process that underpins the construction of shared meanings, therapeutic negotiation, and person-centred decision-making.

As a professional competency, it is multidimensional in nature, integrating verbal, non-verbal, and paralinguistic components, as well as affective, cognitive, and ethical dimensions (Arnold & Boggs, 2003; Kim et al., 2024). In clinical settings, it is reflected in the ability to establish therapeutic alliances, interpret explicit and implicit needs, and adapt interactions to the unique characteristics of the patient, the family, and the clinical situation (Arnold & Boggs, 2003; Kim et al., 2024; Kourkouta & Papathanasiou, 2014; McCabe & Timmins, 2013).

The existing literature consistently demonstrates that ineffective clinical communication is a significant contributor to adverse events, preventable errors, fragmented care, and poorer patient experiences, particularly in healthcare settings marked by high clinical complexity and emotional intensity. (Chang et al., 2025; Leonard et al., 2004; O'Daniel & Rosenstein, 2008; Stiefel et al., 2024). Communication competence therefore constitutes a core dimension of professional practice, underpinning patient safety and clinical effectiveness.

As a result, undergraduate health professions education serves as a critical foundation for the development of clinical communication behaviours that will guide future professional practice. Traditionally, the teaching of these skills has relied on methodologies with an experiential component, including structured role-play, simulation with standardized patients, Objective Structured Clinical Examinations (OSCEs), and reflective debriefing, among others (Chant et al., 2002; Goh et al., 2023).

Research has demonstrated that simulation based on clinical scenarios contributes significantly to the acquisition of clinically relevant communication skills, operationalized through measurable behavioural indicators that can be assessed in clinical performance (Cook et al., 2011; McGaghie et al., 2010). Systematic reviews in health professions education further suggest that structured communication training interventions, particularly when combined with deliberate practice and individualized feedback, tend to be associated with improvements in the assessed competencies, although the magnitude of the effects varies according to the pedagogical design and the instruments used (Cook et al., 2011; Gilligan et al., 2016). Nevertheless, these approaches have recognized limitations, namely high logistical demands, significant costs, dependence on specialized assessors, and inter-rater variability in performance assessment (Dewan et al., 2024; Hodges, 2003; McGaghie et al., 2010). In addition, the need for scalability and personalization of learning, in increasingly large higher education cohorts, has highlighted constraints on deliberate repetition and individualized feedback.

constrangimentos na repetição deliberada e no feedback individualizado.

Neste contexto de transformação pedagógica, as tecnologias de inteligência artificial (IA) emergem como ferramentas potencialmente transformadoras no ensino das profissões da saúde. A integração de algoritmos de aprendizagem automática e processamento de linguagem natural permite a criação de ambientes interativos capazes de simular diálogos clínicos, adaptar cenários ao desempenho do estudante e fornecer feedback automatizado em tempo real (Chan & Zary, 2019; Masters, 2019; Zarei et al., 2024).

Mais recentemente, a incorporação de modelos de linguagem de grande dimensão e agentes conversacionais em contextos educativos tem suscitado interesse significativo, quer pelo seu potencial na simulação de raciocínio clínico e interação comunicacional, quer pela possibilidade de treino autónomo e prática deliberada ilimitada (Kung et al., 2023; Sallam, 2023).

Apesar do interesse crescente, a evidência permanece fragmentada quanto ao tipo de tecnologia utilizada, aos fundamentos pedagógicos, aos instrumentos de avaliação e aos resultados formativos reportados. Esta rápida incorporação de IA em contextos educativos tem, por vezes, precedido a consolidação de enquadramentos pedagógicos explicitamente fundamentados e de medidas de avaliação consistentes, contribuindo para dispersão terminológica e heterogeneidade metodológica no domínio da comunicação clínica.

Antes de avançar para análises comparativas de eficácia, torna-se necessário mapear sistematicamente o estado atual do conhecimento, identificando que tecnologias educativas baseadas em IA estão a ser utilizadas, como são operacionalizadas no desenho pedagógico, com que instrumentos são avaliadas e que tipos de evidência suportam os resultados reportados. Neste enquadramento, o presente protocolo propõe a realização de uma *scoping review* destinada a mapear a utilização de tecnologias de inteligência artificial aplicadas ao desenvolvimento de competências comunicacionais clínicas em estudantes pré-graduados da área da saúde.

2. MATERIAL E MÉTODOS

A presente revisão será conduzida como uma *scoping review*, seguindo a metodologia do Joanna Briggs Institute (JBI) (Peters et al., 2020), complementada pelos princípios fundadores e recomendações de aperfeiçoamento metodológico para *scoping reviews* (Arksey & O'Malley, 2005; Levac, Colquhoun, & O'Brien, 2010). Este desenho é particularmente indicado quando se pretende mapear um domínio emergente caracterizado por diversidade conceptual, multiplicidade de intervenções e variabilidade de medidas de avaliação, permitindo clarificar definições operacionais, cartografar abordagens pedagógicas e identificar lacunas de investigação (Vilelas, 2017).

O protocolo foi concebido para assegurar transparência e rastreabilidade em todas as etapas (definição da questão, elegibilidade, pesquisa, seleção, extração e síntese), e o reporte final seguirá as recomendações da PRISMA-ScR (Tricco et al., 2018). O protocolo foi registado prospetivamente na Open

Within this context of pedagogical transformation, artificial intelligence (AI) technologies are emerging as potentially transformative tools in health professions education. The integration of machine learning algorithms and natural language processing enables the creation of interactive environments capable of simulating clinical dialogues, adapting scenarios to student performance, and providing automated real-time feedback (Chan & Zary, 2019; Masters, 2019; Zarei et al., 2024).

More recently, the incorporation of large language models and conversational agents into educational contexts has attracted considerable interest, both for their potential in simulating clinical reasoning and communicative interaction, and for the possibility of autonomous training and unlimited deliberate practice (Kung et al., 2023; Sallam, 2023).

Despite this growing interest, the evidence remains fragmented regarding the type of technology used, the underlying pedagogical foundations, the assessment instruments, and the reported educational outcomes. This rapid incorporation of AI into educational contexts has, at times, preceded the consolidation of explicitly grounded pedagogical frameworks and consistent assessment measures, contributing to terminological dispersion and methodological heterogeneity in the domain of clinical communication.

Before advancing to comparative analyses of effectiveness, it is necessary to systematically map the current state of knowledge by identifying which AI-based educational technologies are being used, how they are operationalized within the pedagogical design, with which instruments they are assessed, and what types of evidence support the reported outcomes. Within this framework, the present protocol proposes a *scoping review* aimed at mapping the use of artificial intelligence technologies applied to the development of clinical communication skills among undergraduate health professions students.

2. MATERIALS AND METHODS

The present review will be conducted as a *scoping review*, following the Joanna Briggs Institute (JBI) methodology (Peters et al., 2020), complemented by the founding principles and methodological refinement recommendations for *scoping reviews* (Arksey & O'Malley, 2005; Levac, Colquhoun, & O'Brien, 2010). This design is particularly suited to mapping an emerging domain characterized by conceptual diversity, a multiplicity of interventions, and variability in assessment measures, allowing operational definitions to be clarified, pedagogical approaches to be mapped, and research gaps to be identified (Vilelas, 2017).

The protocol was designed to ensure transparency and traceability across all stages (question definition, eligibility, search, selection, extraction, and synthesis), and the final report will follow the PRISMA-ScR recommendations (Tricco et al., 2018). The protocol was prospectively registered on the Open Science Framework (OSF) (<https://osf.io/bns69>), ensuring methodological transparency and traceability.

The research question was structured according to the Population–Concept–Context (PCC) framework. The population of interest comprises undergraduate health

Science Framework (OSF) (<https://osf.io/bns69>), assegurando transparência e rastreabilidade metodológica.

A questão de investigação foi estruturada segundo o modelo Population–Concept–Context (PCC). A população de interesse integra estudantes pré-graduados da área da saúde. O conceito centra-se na utilização de tecnologias de inteligência artificial aplicadas ao desenvolvimento de competências comunicacionais clínicas. O contexto corresponde a ambientes formais de aprendizagem em comunicação no âmbito da formação pré-graduada, incluindo contextos presenciais, digitais ou híbridos, quando integrados em programas curriculares. A revisão orienta-se pela seguinte questão: que tecnologias de inteligência artificial têm sido utilizadas no desenvolvimento de competências de comunicação clínica em estudantes pré-graduados da área da saúde, e de que modo são operacionalizadas no que respeita ao enquadramento pedagógico, aos instrumentos de avaliação e à evidência empírica dos resultados formativos?

2.1. CRITÉRIOS DE ELEGIBILIDADE

Serão incluídas fontes de evidência que satisfaçam cumulativamente os critérios relativos à população, conceito e contexto (PCC), bem como os critérios de tipo de evidência, idioma e período temporal.

Relativamente à população, serão elegíveis estudos que incluam estudantes pré-graduados da área da saúde, entendidos como estudantes em programas conducentes à qualificação profissional inicial (por exemplo, licenciatura ou mestrado integrado) em enfermagem, medicina, medicina dentária/odontologia, farmácia e profissões de diagnóstico e terapêutica, entre outras formações equivalentes na área da saúde. Serão excluídos estudos conduzidos unicamente com profissionais já graduados em formação contínua e com internos/residentes, quando explicitamente enquadrados como formação pós-graduada. Em casos limite, nos quais o estatuto formativo dos participantes não seja claramente descrito, será considerada elegível apenas a população explicitamente enquadrada pelos autores como formação pré-graduada ou equivalente. Estudos com amostras mistas, incluindo pré-graduados e pós-graduados/profissionais, serão incluídos apenas quando os dados relativos a estudantes pré-graduados sejam extraíveis de forma autónoma; caso tal não seja possível, o estudo será excluído, por não permitir atribuição inequívoca dos resultados à população-alvo.

Quanto ao conceito, serão incluídos estudos que descrevam intervenções educativas em que a inteligência artificial desempenhe um papel ativo no processo de aprendizagem de competências comunicacionais clínicas. Para efeitos desta revisão, considera-se “IA aplicada diretamente à aprendizagem” qualquer tecnologia baseada em inteligência artificial que desempenhe uma função pedagógica direta no desenvolvimento de competências comunicacionais clínicas dos estudantes. Serão incluídos sistemas que recorram explicitamente a técnicas de IA, tais como aprendizagem automática, processamento de linguagem natural, modelos de linguagem de grande dimensão (LLM), agentes conversacionais, tutores inteligentes ou sistemas adaptativos, para interagir com o estudante, analisar a sua produção

professions students. The concept focuses on the use of artificial intelligence technologies applied to the development of clinical communication skills. The context corresponds to formal communication-learning environments within undergraduate education, including face-to-face, digital, or hybrid settings, when integrated into curricular programmes. The review is guided by the following question: which artificial intelligence technologies have been used in the development of clinical communication skills among undergraduate health professions students, and how are they operationalized with respect to the pedagogical framework, the assessment instruments, and the empirical evidence of educational outcomes?

2.1 ELIGIBILITY CRITERIA

Sources of evidence will be included if they cumulatively meet the criteria relating to population, concept, and context (PCC), as well as the criteria concerning type of evidence, language, and time period.

With regard to the population, eligible studies will be those including undergraduate health professions students, understood as students enrolled in programmes leading to initial professional qualification (for example, a bachelor's degree or an integrated master's degree) in nursing, medicine, dentistry, pharmacy, and diagnostic and therapeutic professions, among other equivalent health-related programmes. Studies conducted solely with already-graduated professionals in continuing education and with interns/residents, when explicitly framed as postgraduate training, will be excluded. In borderline cases in which the educational status of the participants is not clearly described, only the population explicitly framed by the authors as undergraduate education or equivalent will be considered eligible. Studies with mixed samples, including undergraduate and postgraduate students/professionals, will be included only when data relating to undergraduate students can be extracted independently; where this is not possible, the study will be excluded, as it would not allow unequivocal attribution of the results to the target population.

Regarding the concept, studies describing educational interventions in which artificial intelligence plays an active role in the process of learning clinical communication skills will be included. For the purposes of this review, “AI applied directly to learning” is understood as any artificial intelligence-based technology that performs a direct pedagogical function in the development of students' clinical communication skills. Systems that explicitly draw on AI techniques will be included, such as machine learning, natural language processing, large language models (LLMs), conversational agents, intelligent tutors, or adaptive systems, used to interact with the student, analyse their written or oral output, simulate clinical interlocutors, generate responses or scenarios, provide pedagogical guidance or automated feedback, or produce adaptive instructional materials, including personalized clinical cases, provided that these are used directly in learning activities.

Merely digital technologies without an AI component, such as non-adaptive e-learning, will be excluded, as will

textual ou oral, simular interlocutores clínicos, gerar respostas ou cenários, fornecer orientação pedagógica ou feedback automatizado, ou produzir materiais pedagógicos adaptativos, incluindo casos clínicos personalizados, desde que estes sejam utilizados diretamente em atividades de aprendizagem.

Serão excluídas tecnologias meramente digitais sem componente de IA, como e-learning não adaptativo, bem como aplicações de IA sem função pedagógica direta no treino comunicacional. Incluem-se neste caso ferramentas exclusivamente administrativas, de triagem, de apoio à decisão clínica sem componente formativa centrada na comunicação, ou ferramentas utilizadas apenas para preparação genérica de materiais, sem aplicação direta documentada em atividades de aprendizagem de competências comunicacionais clínicas.

A “competência comunicacional clínica” será operacionalizada como comunicação profissional com o doente e a família, incluindo, por exemplo, entrevista clínica, comunicação terapêutica, empatia, educação para a saúde, comunicação de más notícias e gestão de emoções ou conflito em contexto clínico. Componentes de comunicação interprofissional serão incluídas apenas quando explicitamente enquadradas na prestação de cuidados (por exemplo, *handover*) e avaliadas com indicadores comunicacionais.

No âmbito desta revisão, a empatia será considerada como uma dimensão da comunicação clínica em que o estudante procura compreender, reconhecer e responder à experiência emocional, cognitiva ou vivencial do doente e/ou da família, demonstrando essa compreensão através da linguagem verbal e não-verbal. Esta dimensão poderá ser operacionalizada através da escuta ativa, da validação emocional, do reconhecimento das preocupações expressas pelo doente e da formulação de respostas que transmitam compreensão, respeito e atenção à experiência da pessoa em situação de cuidado. Serão excluídos estudos que abordem a empatia exclusivamente como traço disposicional ou construto psicológico, sem ligação explícita a comportamentos de comunicação clínica.

Relativamente ao contexto, serão elegíveis estudos realizados em ambientes educativos formais (presenciais, digitais ou híbridos) integrados em programas curriculares pré-graduados, incluindo unidades curriculares, programas estruturados e cenários de simulação. Serão excluídos estudos realizados fora de estruturas educativas formais ou intervenções no local de trabalho sem desenho formativo dirigido a estudantes.

No que respeita ao tipo de evidência, serão incluídos estudos empíricos primários quantitativos, qualitativos ou de métodos mistos e avaliações de programas educativos, desde que disponibilizem texto integral e descrevam a intervenção com detalhe suficiente para extração de dados (características da IA, componentes pedagógicas e forma de avaliação). Serão excluídos editoriais, comentários, cartas e textos opinativos sem dados, bem como resumos de congresso sem texto integral. Serão considerados estudos publicados em português, inglês ou espanhol. Não será aplicado limite temporal à pesquisa, de modo a assegurar a identificação abrangente de evidência relevante sobre a utilização de tecnologias de inteligência artificial na formação em comunicação clínica.

AI applications without a direct pedagogical function in clinical communication training. This includes exclusively administrative or triage tools, clinical decision-support tools without a training component focused on communication, and tools used solely for the generic preparation of materials, without documented direct application in clinical communication learning activities.

“Clinical communication competence” will be conceptualised as the ability to engage in professional communication with patients and their families. This includes, but is not limited to, conducting clinical interviews, demonstrating therapeutic communication and empathy, providing health education, delivering bad news, and effectively managing emotions and conflict in clinical contexts. Interprofessional communication components will be included only when explicitly framed within the provision of care (for example, *handover*) and assessed using communication indicators.

Within this review, empathy will be regarded as a dimension of clinical communication in which the student seeks to understand, recognize, and respond to the emotional, cognitive, or experiential reality of the patient and/or family, demonstrating this understanding through verbal and non-verbal language. This dimension may be operationalized through active listening, emotional validation, recognition of the concerns expressed by the patient, and the formulation of responses that convey understanding, respect, and attention to the experience of the person receiving care. Studies that address empathy exclusively as a dispositional trait or psychological construct, without explicit links to clinical communication behaviours, will be excluded.

Regarding context, eligible studies will be those carried out in formal educational settings (face-to-face, digital, or hybrid) integrated into undergraduate curricular programmes, including curricular units, structured programmes, and simulation scenarios. Studies conducted outside formal educational structures, or workplace interventions without an instructional design directed at students, will be excluded.

About the type of evidence, primary empirical studies—quantitative, qualitative, or mixed-methods—and evaluations of educational programmes will be included, provided that they offer full text and describe the intervention in sufficient detail for data extraction (AI characteristics, pedagogical components, and form of assessment). Editorials, commentaries, letters, and opinion pieces without data, as well as conference abstracts without full text, will be excluded. Studies published in Portuguese, English, or Spanish will be considered. No time limit will be applied to the search, to ensure the comprehensive identification of relevant evidence on the use of artificial intelligence technologies in clinical communication training.

2.2 SEARCH STRATEGY AND IDENTIFICATION OF STUDIES

The search strategy will be conducted in a systematic and reproducible manner, following three stages: (i) an initial exploratory search to refine terms; (ii) a comprehensive search of the selected databases, with syntactic adaptation for each database; and (iii) systematic screening of the references and

2.2. ESTRATÉGIA DE PESQUISA E IDENTIFICAÇÃO DOS ESTUDOS

A estratégia de pesquisa será conduzida de forma sistemática e reprodutível, seguindo três etapas: (i) pesquisa exploratória inicial para refinamento de termos; (ii) pesquisa abrangente nas bases selecionadas com adaptação sintática por base; e (iii) rastreamento sistemático de referências e citações dos estudos incluídos. A estratégia será estruturada em quatro blocos operacionais: estudantes pré-graduados da área da saúde, competências comunicacionais clínicas, inteligência artificial aplicada diretamente à aprendizagem e contexto educacional/formativo. Os três primeiros blocos correspondem aos elementos centrais da questão PCC, enquanto o quarto bloco será utilizado como componente operacional de delimitação do contexto educativo. A inclusão deste quarto bloco será testada durante a fase exploratória da estratégia, comparando-se o número e a relevância dos registros recuperados com e sem a sua aplicação. Caso se verifique perda de estudos potencialmente elegíveis, a estratégia será ajustada de modo a privilegiar a sensibilidade da pesquisa.

A pesquisa exploratória inicial foi realizada em 18 de março de 2026, na MEDLINE via PubMed, utilizando a estratégia preliminar apresentada neste protocolo, tendo recuperado 382 registros. Esta etapa permitiu testar a viabilidade da estratégia de pesquisa, estimar a ordem de grandeza dos registros potencialmente recuperados e identificar eventuais necessidades de ajustamento dos termos utilizados.

Serão pesquisadas as bases de dados MEDLINE (via PubMed), CINAHL, Scopus, Web of Science Core Collection, ERIC e PsycINFO. Serão incluídos estudos até à data da pesquisa, em português, inglês ou espanhol. Não serão aplicados filtros por desenho de estudo na pesquisa eletrônica, sendo os critérios de elegibilidade aplicados durante a fase de triagem. A estratégia de pesquisa será atualizada antes da conclusão da revisão, sendo reportadas as datas de execução e o número de novos registros identificados no fluxograma PRISMA-ScR.

Antes da execução final, a estratégia eletrônica será revista com o apoio de uma profissional da área da Ciência da Informação e Gestão de Bibliotecas, com experiência em pesquisa bibliográfica e gestão da produção científica em contexto acadêmico. Esta revisão será orientada pelos princípios da checklist *Peer Review of Electronic Search Strategies (PRESS): 2015 Guideline Statement*, incluindo a adequação da tradução da questão de investigação em conceitos pesquisáveis, o uso de vocabulário controlado e termos livres, a combinação de operadores booleanos, a adaptação da sintaxe a cada base de dados e a coerência entre os blocos da estratégia. As alterações resultantes deste processo serão registadas e reportadas no manuscrito final da revisão (McGowan et al., 2016).

Com o objetivo de reforçar a transparência na construção da estratégia de pesquisa, apresenta-se na Tabela 1 a estrutura conceptual e operacional utilizada para identificar descritores controlados e termos livres correspondentes aos principais componentes do modelo Population–Concept–Context.

citations of the included studies. The strategy will be structured into four operational blocks: undergraduate health professions students, clinical communication skills, artificial intelligence applied directly to learning, and educational/training context. The first three blocks correspond to the central elements of the PCC question, while the fourth block will be used as an operational component for delimiting the educational context. The inclusion of this fourth block will be tested during the exploratory phase of the strategy, comparing the number and relevance of the records retrieved with and without its application. If potentially eligible studies are missed, the strategy will be adjusted to favour the sensitivity of the search.

The initial exploratory search was carried out on 18 March 2026 in MEDLINE via PubMed, using the preliminary strategy presented in this protocol, and retrieved 382 records. This stage made it possible to test the feasibility of the search strategy, estimate the order of magnitude of the records potentially retrievable, and identify any need to adjust the terms used.

The MEDLINE (via PubMed), CINAHL, Scopus, Web of Science Core Collection, ERIC, and PsycINFO databases will be searched. Studies up to the date of the search, in Portuguese, English, or Spanish, will be included. No study-design filters will be applied in the electronic search; eligibility criteria will be applied during the screening phase. The search strategy will be updated before the conclusion of the review, and the execution dates and number of newly identified records will be reported in the PRISMA-ScR flow diagram.

Before final execution, the electronic strategy will be reviewed with the support of a professional in the field of Information Science and Library Management, with experience in bibliographic searching and the management of scientific output in academic settings. This review will be guided by the principles of the Peer Review of Electronic Search Strategies (PRESS): 2015 Guideline Statement checklist, including the adequacy of translating the research question into searchable concepts, the use of controlled vocabulary and free-text terms, the combination of Boolean operators, the adaptation of syntax to each database, and the coherence between the blocks of the strategy. The changes resulting from this process will be recorded and reported in the final manuscript of the review (McGowan et al., 2016).

To strengthen transparency in the construction of the search strategy, Table 1 presents the conceptual and operational structure used to identify controlled descriptors and free-text terms corresponding to the main components of the Population–Concept–Context framework.

Tabela/Table 1: Estrutura conceitual e operacional da estratégia de pesquisa/Conceptual and operational structure of the search strategy.

Componente PCC/ PCC Component	Descritores controlados (MeSH/CINAHL)/ Controlled descriptors (MeSH/CINAHL)	Termos livres/ Free-text terms
População/ Population	"Students, Health Occupations"; "Students, Nursing"; "Students, Medical"	student*; undergraduate*; "health professions student*"; "nursing student*"; "medical student*"; "dental student*"; "pharmacy student*"
Conceito – Competências comunicacionais clí- nicas/ Concept – Clinical communication skills	"Communication"; "Interpersonal Communication"	"communication skill*"; "clinical communication"; "therapeutic communication"; "patient-centered communication"; empathy; "breaking bad news"; "shared decision*"
Conceito – Inteligência artificial/ Concept – Artificial intelligence	"Artificial Intelligence"; "Machine Learning"; "Natural Language Processing"	chatbot*; "conversational agent*"; "large language model*"; ChatGPT; "virtual patient*"; "intelligent tutor*"; "automated feedback"
Contexto – Educacional/formativo/ Context – Educational/training	"Education"; "Simulation"; "Objective Structured Clinical Examination"; "Standardized Patients"	education; training; teach*; learn*; curriculum; simulation; OSCE; "objective structured clinical exa- mination"; "standardized patient*"

Adicionalmente, para assegurar a reprodutibilidade da pesquisa, apresenta-se abaixo a estratégia completa prevista para MEDLINE, via PubMed, estruturada segundo os quatro blocos operacionais descritos.

Estratégia PubMed (MEDLINE)

((("Students, Health Occupations"[MeSH] OR student*[tiab] OR undergraduate*[tiab] OR "health professions student*[tiab]) AND ("Communication"[MeSH] OR "InterpersonalCommunication"[MeSH]OR"communication skill*[tiab] OR "clinical communication"[tiab] OR "therapeutic communication"[tiab] OR empathy[tiab] OR "breaking bad news"[tiab] OR "shared decision*[tiab]) AND ("Artificial Intelligence"[MeSH] OR "Machine Learning"[MeSH] OR "Natural Language Processing"[MeSH] OR chatbot*[tiab] OR "conversational agent*[tiab] OR "large language model*[tiab] OR ChatGPT[tiab] OR "virtual patient*[tiab] OR "intelligent tutor*[tiab] OR "automated feedback"[tiab]) AND (education[tiab] OR training[tiab] OR curriculum[tiab] OR simulation[tiab])).

A estratégia será adaptada às restantes bases, utilizando CINAHL Headings, Thesaurus ERIC e Thesaurus PsycINFO, quando aplicável, mantendo os blocos operacionais e ajustando campos e operadores, como truncaturas e proximidade, conforme a plataforma. Para cada base serão registados e reportados no manuscrito: data de execução, limites aplicados e número de registos recuperados.

Para reduzir o risco de omissão de evidência relevante em domínios tecnológicos emergentes, será pesquisada literatura cinzenta no Google Scholar, RCAAP, OpenAlex e, quando disponível através de acesso institucional, ProQuest Dissertations & Theses Global. Adicionalmente, dada a rápida evolução do campo da inteligência artificial aplicada à educação, serão pesquisados repositórios de preprints relevantes, nomeadamente medRxiv e EdArXiv. No Google Scholar, será realizada triagem das primeiras 200 entradas ordenadas por relevância; a expressão de pesquisa exata, a data/hora de execução e o número de registos triados serão registados e reportados no manuscrito. Nas restantes fontes de literatura cinzenta, serão registadas a expressão de pesquisa utilizada, a data de execução e o número de registos recuperados.

Os registos serão importados diretamente para o Rayyan, onde será realizada a remoção de duplicados e a triagem dos

Additionally, to ensure the reproducibility of the search, the complete strategy planned for MEDLINE via PubMed is presented below, structured according to the four operational blocks described.

PubMed search strategy (MEDLINE)

((("Students, Health Occupations"[MeSH] OR student*[tiab] OR undergraduate*[tiab] OR "health professions student*[tiab]) AND ("Communication"[MeSH] OR "InterpersonalCommunication"[MeSH]OR"communication skill*[tiab] OR "clinical communication"[tiab] OR "therapeutic communication"[tiab] OR empathy[tiab] OR "breaking bad news"[tiab] OR "shared decision*[tiab]) AND ("Artificial Intelligence"[MeSH] OR "Machine Learning"[MeSH] OR "Natural Language Processing"[MeSH] OR chatbot*[tiab] OR "conversational agent*[tiab] OR "large language model*[tiab] OR ChatGPT[tiab] OR "virtual patient*[tiab] OR "intelligent tutor*[tiab] OR "automated feedback"[tiab]) AND (education[tiab] OR training[tiab] OR curriculum[tiab] OR simulation[tiab])).

The strategy will be adapted to the remaining databases, using CINAHL Headings, the ERIC Thesaurus, and the PsycINFO Thesaurus where applicable, maintaining the operational blocks and adjusting fields and operators—such as truncation and proximity—according to the platform. For each database, the execution date, applied limits, and number of records retrieved will be recorded and reported in the manuscript.

To reduce the risk of omitting relevant evidence in emerging technological domains, grey literature will be searched in Google Scholar, RCAAP, OpenAlex, and—when available through institutional access—ProQuest Dissertations & Theses Global. Additionally, given the rapid evolution of the field of artificial intelligence applied to education, relevant preprint repositories will be searched, namely medRxiv and EdArXiv. In Google Scholar, the first 200 entries sorted by relevance will be screened; the exact search expression, the date/time of execution, and the number of records screened will be recorded and reported in the manuscript. For the remaining grey-literature sources, the search expression used, the execution date, and the number of records retrieved will be recorded.

Records will be imported directly into Rayyan, where deduplication and study screening will be carried out. Selection

estudos. A seleção decorrerá em duas fases (título/resumo e texto integral) por dois revisores independentes, precedida de calibração em 50 registos. A concordância será quantificada através da percentagem de acordo e do coeficiente kappa de Cohen, considerando-se adequada uma concordância correspondente a $\text{kappa} \geq 0,70$. Caso este limiar não seja atingido, as discordâncias serão discutidas entre os revisores, os critérios de elegibilidade serão clarificados e, se necessário, o manual de decisão será revisto. Nessa situação, será realizada nova ronda de calibração com um conjunto adicional de registos antes do início da triagem independente. Divergências subsequentes serão resolvidas por consenso e, quando necessário, por terceiro revisor. As razões de exclusão em texto integral serão registadas. Quando forem identificados múltiplos relatórios relativos ao mesmo estudo, estes serão agregados e tratados como uma única unidade de análise. Será privilegiada a publicação revista por pares, quando disponível; caso existam várias versões com estatuto semelhante, será selecionada a versão que apresente informação metodológica e resultados mais completos; em caso de equivalência quanto à completude da informação, será considerada a versão mais recente. O processo de identificação, triagem, elegibilidade e inclusão dos estudos será apresentado através de um fluxograma PRISMA-ScR.

2.3. EXTRAÇÃO DOS DADOS

A extração será realizada através de um instrumento de extração de dados desenvolvido pela equipa e apresentado na Tabela 2. O instrumento foi estruturado por domínios, campos de extração e tipo de dado a recolher, incluindo, sempre que aplicável, opções ou categorias de codificação, de modo a favorecer a consistência entre revisores e a reprodutibilidade do processo de extração.

will take place in two stages (title/abstract and full text) by two independent reviewers, preceded by calibration on 50 records. Agreement will be quantified through the percentage of agreement and Cohen’s kappa coefficient, with an agreement corresponding to $\text{kappa} \geq 0.70$ considered adequate. Should this threshold not be reached, the disagreements will be discussed between the reviewers, the eligibility criteria will be clarified, and, if necessary, the decision manual will be revised. In that situation, a new calibration round will be carried out with an additional set of records before independent screening begins. Subsequent disagreements will be resolved by consensus and, where necessary, by a third reviewer. The reasons for full-text exclusion will be recorded. When multiple reports relating to the same study are identified, they will be aggregated and treated as a single unit of analysis. Peer-reviewed publication will be favoured where available; where several versions with a similar status exist, the version presenting the most complete methodological information and results will be selected; in the case of equivalence regarding the completeness of the information, the most recent version will be considered. The process of identification, screening, eligibility, and inclusion of studies will be presented through a PRISMA-ScR flow diagram.

2.3 DATA EXTRACTION

Extraction will be carried out using a data-extraction instrument developed by the team and presented in Table 2. The instrument was structured by domains, extraction fields, and the type of data to be collected, including, where applicable, coding options or categories, so as to promote consistency between reviewers and the reproducibility of the extraction process.

Tabela/Table 2: Instrumento preliminar de extração de dados/Preliminary data-extraction instrument.

Título/Title	Texto livre/Free text
Ano/Year	Numérico/Numeric
Autor/Author	Texto livre/Free text
País/Country	Texto livre/Free text
Revista/Fonte/Journal/Source	Texto livre/Free text
Tipo de estudo/Type of study	Quantitativo; qualitativo; métodos mistos; avaliação de programa; quase-experimental; experimental; observacional; outro
Curso/Programme	Enfermagem; Medicina; Medicina Dentária/Odontologia; Farmácia; Psicologia; Fisioterapia; outra profissão da saúde; amostra mista/
Ano curricular/Academic year	Numérico; texto livre; não reportado/
Número de participantes/Number of participants	Numérico; não reportado/
Contexto/Setting	Sala de aula; simulação; online; híbrido; outro; não reportado/
Tipo de tecnologia/Type of technology	Chatbot/agente conversacional; paciente virtual; modelo de linguagem de grande dimensão (LLM); tutor inteligente; sistema adaptativo; processamento de linguagem natural; aprendizagem automática; outro; não reportado/
Modalidade/Modality	Texto; voz; vídeo; multimodal; não reportado/
Função da IA/AI function	Interação com estudante; simulação de interlocutor clínico; geração de cenários/casos; análise de produção textual/oral; feedback automatizado; orientação pedagógica; outro/
Grau de adaptabilidade/Degree of adaptability	Não adaptativo; parcialmente adaptativo; adaptativo; não reportado/
Objetivo de aprendizagem/Learning objective	Texto livre
Tipo de intervenção/Type of intervention	Simulação; treino autónomo; atividade curricular; prática deliberada; feedback automatizado; outro/

Integração curricular/ Curricular integration	Curricular; extracurricular; não reportado/
Duração das sessões/ Session duration	Texto livre; não reportado/
Supervisão docente/ Teacher supervision	Sim; não; não reportado/
Tipo de competência comunicacional analisada/Type of clinical communication competency analysed	Entrevista clínica; comunicação terapêutica; empatia; educação para a saúde; comunicação de más notícias; gestão de emoções/conflito; comunicação interprofissional/handover; outro; não reportado/
Instrumento utilizado/Instrument used	Texto livre; não reportado/
Tipo de outcome/resultado/ Type of outcome/result	Conhecimento; desempenho observado; autorrelato; satisfação/aceitabilidade; outro/
Miller	Knows; Knows how; Shows how; Does; não classificável/
Kirkpatrick – nível 1/Reação/ level 1/Reaction	Satisfação; aceitabilidade; percepção de utilidade; envolvimento; experiência de aprendizagem; não aplicável; não reportado/
Barreiras identificadas/ facilitadores/ Barriers/facilitators identified	Texto livre; não reportado/
Considerações éticas e de segurança/ Ethical and safety considerations	Privacidade/proteção de dados; vies algorítmico; outputs incorretos ou enviesados; integridade académica; segurança pedagógica; outro; não reportado/

Em termos de conteúdo, o instrumento contempla domínios relativos à identificação do estudo, características da amostra, caracterização técnica da tecnologia de IA, caracterização pedagógica da intervenção, competências comunicacionais clínicas abordadas, instrumentos e tipo de avaliação, classificação dos outcomes segundo a Pirâmide de Miller e, quando aplicável, segundo o nível 1 do modelo de Kirkpatrick, aspetos de implementação e considerações éticas e de segurança.

O instrumento será testado em cinco estudos incluídos, com refinamento iterativo antes da extração definitiva. Esta etapa permitirá avaliar a clareza dos campos, a adequação das opções de codificação e a consistência entre revisores. As discrepâncias ou dificuldades de codificação identificadas durante este teste serão discutidas entre os revisores e resolvidas por consenso, podendo originar a clarificação de definições operacionais ou o ajustamento das opções de codificação. A extração definitiva será conduzida por dois revisores independentes, com resolução de discrepâncias por consenso e, quando necessário, por terceiro revisor.

Para maximizar comparabilidade e rigor em intervenções baseadas em IA, serão extraídos, no mínimo, os seguintes elementos: (i) características do estudo (autor, ano, país, desenho); (ii) características da amostra (curso, ano curricular, dimensão, contexto); (iii) caracterização técnica da IA (tipo de sistema: LLM/agente conversacional/tutor inteligente/ processamento de linguagem natural (NLP) *scoring*; modalidade de entrada/saída: texto/voz/vídeo; grau de adaptabilidade; autonomia do estudante); (iv) caracterização pedagógica (objetivos de aprendizagem, modelo/estratégia pedagógica explícita, dose/duração, integração curricular, supervisão docente, natureza do feedback: imediato ou diferido; formativo ou somativo; rubricado ou livre); (v) avaliação (instrumentos utilizados, validade/evidência de validação quando disponível, nível de *outcome*: desempenho observável ou autorrelato); (vi) resultados formativos e aceitabilidade; (vii) aspetos de implementação (barreiras/facilitadores); e (viii) considerações éticas e de segurança relevantes (por exemplo, privacidade e proteção de dados, integridade académica, mitigação

In terms of content, the instrument encompasses domains relating to study identification, sample characteristics, technical characterization of the AI technology, pedagogical characterization of the intervention, the clinical communication skills addressed, the instruments and type of assessment, the classification of outcomes according to Miller’s Pyramid and, when applicable, according to level 1 of the Kirkpatrick model, implementation aspects, and ethical and safety considerations.

The instrument will be tested on five included studies, with iterative refinement before definitive extraction. This stage will make it possible to assess the clarity of the fields, the adequacy of the coding options, and the consistency between reviewers. Any discrepancies or coding difficulties identified during this test will be discussed between the reviewers and resolved by consensus, potentially leading to the clarification of operational definitions or the adjustment of coding options. Final data extraction will be conducted by two independent reviewers, with discrepancies resolved by consensus and, where necessary, by a third reviewer.

To maximize comparability and rigour in AI-based interventions, at least the following elements will be extracted: (i) study characteristics (author, year, country, design); (ii) sample characteristics (programme, academic year, size, setting); (iii) technical characterization of the AI (type of system: LLM/conversational agent/intelligent tutor/natural language processing (NLP) scoring; input/output modality: text/voice/video; degree of adaptability; student autonomy); (iv) pedagogical characterization (learning objectives, explicit pedagogical model/strategy, dose/duration, curricular integration, teacher supervision, nature of the feedback: immediate or deferred; formative or summative; rubric-based or open); (v) assessment (instruments used, validity/ validation evidence where available, outcome level: observable performance or self-report); (vi) educational outcomes and acceptability; (vii) implementation aspects (barriers/ facilitators); and (viii) relevant ethical and safety considerations (for example, privacy and data protection, academic integrity, mitigation of incorrect or biased outputs, and potential adverse effects associated with immersive environments—namely

de outputs incorretos ou enviesados, e potenciais efeitos adversos associados a ambientes imersivos — nomeadamente *cybersickness* — bem como fadiga cognitiva e fenómenos de desrealização, quando aplicável).

2.4. SÍNTESE DOS DADOS

Os resultados serão sintetizados de forma descritiva e analítica, privilegiando o mapeamento do campo. A síntese integrará: (i) caracterização quantitativa dos estudos incluídos (frequências por tipo de tecnologia, área de formação, contexto educativo e tipo de avaliação); (ii) categorização temática das intervenções educativas (componentes pedagógicas, mecanismos de feedback e formas de integração curricular); e (iii) construção de matrizes de evidência relacionando tecnologia de inteligência artificial, dimensão da competência comunicacional abordada, instrumentos de avaliação utilizados e tipo de evidência reportada.

Adicionalmente, será desenvolvido um mapa conceptual da evidência, com o objetivo de visualizar de forma integrada a relação entre os diferentes tipos de tecnologias de inteligência artificial identificados, as dimensões da competência comunicacional abordadas, as estratégias pedagógicas utilizadas e os níveis de avaliação educacional reportados nos estudos incluídos.

Para reforçar o rigor interpretativo, os resultados serão organizados principalmente segundo a Pirâmide de Miller (Knows, Knows how, Shows how, Does), distinguindo a evidência centrada em conhecimento da evidência baseada em desempenho comunicacional observável (Miller, 1990). Assim, avaliações baseadas em OSCE e grelhas observacionais em simulação serão classificadas como “Shows how”; avaliação em prática clínica real, quando existente, como “Does”; e testes de conhecimento como “Knows” ou “Knows how”.

A utilização complementar do modelo de Kirkpatrick será limitada ao nível 1, correspondente à reação dos participantes, de modo a captar dimensões como satisfação, aceitabilidade, perceção de utilidade, envolvimento e experiência de aprendizagem. Esta opção justifica-se pelo facto de os níveis 2 e 3 de Kirkpatrick, relativos à aprendizagem e ao comportamento, apresentarem sobreposição operacional com os níveis da Pirâmide de Miller, nomeadamente “Knows/Knows how” e “Shows how/Does”. Assim, o nível 1 de Kirkpatrick será utilizado apenas como dimensão complementar, por permitir mapear indicadores centrados na experiência e aceitação da tecnologia pelos estudantes, aspetos que não têm correspondência direta na Pirâmide de Miller.

Quando um estudo reportar múltiplos desfechos, cada desfecho será classificado separadamente; quando a informação não permitir classificação, o desfecho será registado como “não classificável”, com justificação. A atribuição das categorias segundo a Pirâmide de Miller e, quando aplicável, segundo o nível 1 do modelo de Kirkpatrick, será realizada por dois revisores independentes, após calibração do procedimento, com resolução de discrepâncias por consenso e, se necessário, por terceiro revisor.

cybersickness—as well as cognitive fatigue and derealization phenomena, where applicable).

2.4 DATA SYNTHESIS

The results will be synthesized descriptively and analytically, prioritizing the mapping of the field. The synthesis will integrate: (i) a quantitative characterization of the included studies (frequencies by type of technology, area of training, educational context, and type of assessment); (ii) a thematic categorization of the educational interventions (pedagogical components, feedback mechanisms, and forms of curricular integration); and (iii) the construction of evidence matrices relating the artificial intelligence technology, the dimension of communication competence addressed, the assessment instruments used, and the type of evidence reported.

Additionally, a conceptual evidence map will be developed, with the aim of providing an integrated visualization of the relationship between the different types of artificial intelligence technologies identified, the dimensions of communication competence addressed, the pedagogical strategies used, and the levels of educational assessment reported in the included studies.

To strengthen interpretive rigour, the results will be organized primarily according to Miller’s Pyramid (Knows, Knows how, Shows how, Does), distinguishing knowledge-focused evidence from evidence based on observable communicative performance (Miller, 1990). Thus, assessments based on OSCEs and observational grids in simulation will be classified as “Shows how”; assessment in real clinical practice, where present, as “Does”; and knowledge tests as “Knows” or “Knows how”.

The complementary use of the Kirkpatrick model will be limited to level 1, corresponding to participants’ reaction, in order to capture dimensions such as satisfaction, acceptability, perceived usefulness, engagement, and the learning experience. This choice is justified by the fact that Kirkpatrick’s levels 2 and 3, relating to learning and behaviour, present operational overlap with the levels of Miller’s Pyramid, namely “Knows/Knows how” and “Shows how/Does”. Accordingly, Kirkpatrick’s level 1 will be used only as a complementary dimension, as it allows the mapping of indicators centred on students’ experience and acceptance of the technology—aspects that have no direct correspondence in Miller’s Pyramid.

When a study reports multiple outcomes, each outcome will be classified separately; when the information does not allow classification, the outcome will be recorded as “not classifiable”, with justification. The assignment of categories according to Miller’s Pyramid and, when applicable, according to level 1 of the Kirkpatrick model, will be carried out by two independent reviewers, following calibration of the procedure, with discrepancies resolved by consensus and, if necessary, by a third reviewer.

3. RESULTS

The results will be presented in characterization and synthesis tables, complemented by a structured narrative. Evidence matrices will be constructed relating the type of AI technology, the dimension of communication competence,

3. RESULTADOS

Os resultados serão apresentados em tabelas de caracterização e síntese, complementadas por narrativa estruturada. Serão construídas matrizes de evidência relacionando tipo de tecnologia de IA, dimensão da competência comunicacional, enquadramento pedagógico, instrumentos de avaliação e tipo de evidência. A interpretação privilegiará a identificação de convergências, lacunas e implicações para a formação pré-graduada, com recomendações para investigação futura e para a eventual construção de revisões sistemáticas de eficácia centradas em subconjuntos mais homogêneos.

4. DISCUSSÃO

Atendendo à natureza protocolar da presente revisão, esta seção incide no contributo esperado do estudo e nos desafios metodológicos inerentes ao domínio em análise. A utilização de tecnologias de inteligência artificial no desenvolvimento de competências comunicacionais clínicas em estudantes pré-graduados da área da saúde constitui um campo emergente, caracterizado por diversidade terminológica, heterogeneidade de intervenções educativas e variabilidade dos instrumentos de avaliação, circunstâncias que sustentam a opção por uma *scoping review*.

A estratégia metodológica delineada pretende favorecer uma leitura estruturada da evidência, através da articulação entre tipo de tecnologia de inteligência artificial, dimensão da competência comunicacional abordada, enquadramento pedagógico, instrumentos de avaliação e natureza dos desfechos reportados. A organização analítica dos resultados com recurso a matrizes de evidência e a sua interpretação à luz da Pirâmide de Miller, complementada, quando aplicável, pelo nível de “Reação” do modelo de Kirkpatrick, poderá contribuir para uma caracterização mais consistente dos níveis de aprendizagem avaliados e da natureza da evidência disponível.

Neste sentido, espera-se que a revisão permita não apenas mapear as tecnologias e abordagens pedagógicas predominantes, mas também evidenciar áreas de dispersão conceptual e metodológica, lacunas na avaliação dos resultados formativos e limitações na descrição das intervenções. Tal mapeamento poderá clarificar em que medida a evidência existente se centra em conhecimento, desempenho observável, aceitabilidade ou outros indicadores educacionais.

Entre as limitações antecipáveis, destaca-se a provável heterogeneidade dos estudos incluídos, quer ao nível dos desenhos metodológicos, quer quanto às características técnicas das tecnologias utilizadas, ao grau de adaptabilidade dos sistemas e aos contextos de implementação. Acresce que a descrição insuficiente de componentes pedagógicas e das propriedades dos instrumentos de avaliação, quando presente nos estudos primários, poderá condicionar a comparabilidade entre intervenções. Adicionalmente, tratando-se de um domínio em rápida evolução, é previsível a emergência contínua de nova evidência ao longo do processo de revisão.

Não obstante estas limitações, considera-se que o protocolo proposto reúne condições para produzir uma síntese abrangente, transparente e metodologicamente fundamentada do estado da arte, com potencial para apoiar decisões curriculares, orientar futuras investigações e sustentar

the pedagogical framework, the assessment instruments, and the type of evidence. The interpretation will prioritize the identification of convergences, gaps, and implications for undergraduate education, with recommendations for future research and for the possible construction of systematic effectiveness reviews focused on more homogeneous subsets.

4. DISCUSSION

Given that this is a review protocol, this section focuses on the expected contribution of the study and on the methodological challenges inherent to the domain under analysis. The use of artificial intelligence technologies in the development of clinical communication skills among undergraduate health professions students constitutes an emerging field, characterized by terminological diversity, heterogeneity of educational interventions, and variability of assessment instruments—circumstances that support the choice of a scoping review.

The methodological strategy outlined aims to favour a structured reading of the evidence, through the articulation of the type of artificial intelligence technology, the dimension of communication competence addressed, the pedagogical framework, the assessment instruments, and the nature of the reported outcomes. The analytical organization of the results using evidence matrices, and their interpretation in light of Miller’s Pyramid—complemented, where applicable, by the “Reaction” level of the Kirkpatrick model—may contribute to a more consistent characterization of the levels of learning assessed and of the nature of the available evidence.

In this sense, the review is expected to allow not only the mapping of the predominant technologies and pedagogical approaches, but also the identification of areas of conceptual and methodological dispersion, gaps in the assessment of educational outcomes, and limitations in the description of interventions. Such mapping may clarify the extent to which the existing evidence focuses on knowledge, observable performance, acceptability, or other educational indicators.

Among the anticipated limitations, the probable heterogeneity of the included studies stands out, both in terms of methodological designs and with respect to the technical characteristics of the technologies used, the degree of adaptability of the systems, and the implementation contexts. Furthermore, the insufficient description of pedagogical components and of the properties of the assessment instruments, when present in the primary studies, may constrain the comparability between interventions. Additionally, as this is a rapidly evolving domain, the continuous emergence of new evidence throughout the review process is foreseeable.

Notwithstanding these limitations, the proposed protocol is considered to meet the conditions for producing a comprehensive, transparent, and methodologically grounded synthesis of the state of the art, with the potential to support curricular decisions, guide future research, and underpin the development of subsequent systematic reviews in more homogeneous subdomains.

5. CONCLUSIONS

This protocol describes a scoping review aimed at

o desenvolvimento de revisões sistemáticas subsequentes em subdomínios mais homogêneos.

5. CONCLUSÕES

Este protocolo descreve uma *scoping review* destinada a mapear, de forma sistemática e transparente, a utilização de tecnologias de inteligência artificial aplicadas diretamente ao processo de aprendizagem no desenvolvimento de competências comunicacionais clínicas em estudantes pré-graduados da área da saúde, bem como a descrever os resultados formativos reportados e as limitações associadas ao seu uso. A revisão irá clarificar as tecnologias e abordagens pedagógicas existentes, os instrumentos de avaliação utilizados e a natureza da evidência disponível sobre os desfechos formativos, identificando lacunas que orientem prioridades de investigação e decisões curriculares.

CONFLITO DE INTERESSES

Os autores declaram não existir conflitos de interesse.

AGRADECIMENTOS

As autoras agradecem a Laura Mafalda Lopes, Técnica Superior na Escola Superior de Enfermagem da Universidade do Porto, pelo apoio prestado na revisão da estratégia de pesquisa, nomeadamente na adequação dos termos, vocabulário controlado e adaptação da estratégia às bases de dados.

CONTRIBUIÇÕES AUTORAIS

Conceptualização: Mariana Lima, Teresa Martins, Lígia Lima, Celeste Bastos; Metodologia: Mariana Lima, Teresa Martins, Lígia Lima, Celeste Bastos; Análise formal: Mariana Lima, Teresa Martins, Lígia Lima, Celeste Bastos; Investigação: Mariana Lima, Teresa Martins, Lígia Lima, Celeste Bastos; Redação - preparação do draft original: Mariana Lima, Teresa Martins, Lígia Lima, Celeste Bastos; Redação - revisão e edição: Mariana Lima, Teresa Martins, Lígia Lima, Celeste Bastos. Todos os autores leram e concordaram com a versão publicada do manuscrito.

REFERÊNCIAS BIBLIOGRÁFICAS/REFERENCES

- AA Aktan GG, Khorshid L. Communication skills and perceived stress of nursing students during first clinical experience. *Journal of Anatolia Nursing and Health Sciences* 24:202-211, 2021.
- Arnold E, Boggs KU. Interpersonal relationships: professional communication skills for nurses. Saunders, St. Louis, 2003.
- Chan KS, Zary N. Applications and challenges of implementing artificial intelligence in medical education: integrative review. *JMIR Medical Education* 5:e13930, 2019.
- Chang CM, Li JJ, Li IF, Lee YH. Divergent perspectives: a cross-sectional study unveiling disparities in cancer patients' and oncology nurses' perceptions on communication and empathy. *European Journal of Oncology Nursing* 76:102877, 2025.
- Chant S, Tim U, Randle J, Russell G, Webb C. Communication skills training in healthcare: a review of the literature. *Nurse Education Today* 22:189-202, 2002.
- Cook DA, Hatala R, Brydges R, Zendejas B, Szostek JH, Wang AT, Erwin PJ, Hamstra SJ. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. *JAMA* 306:978-988, 2011.
- Dewan P, Khalil S, Gupta P. Objective structured clinical examination for teaching and assessment: evidence-based critique. *Clinical Epidemiology and Global Health* 25:101477, 2024.
- Gilligan C, James EL, Snow P, Outram S, Ward BM, Powell M, Lonsdale C, Cushing AM, Silverman J, Regan T, Harvey P, Lynagh MC. Interventions for improving medical students' interpersonal communication in medical consultations. *Cochrane Database of Systematic Reviews* CD012418, 2016.
- Goh PRQ, Ng GYJ, Shorey S, Lim S. Impacts of standardized patients as a teaching tool to develop communication skills in nursing education: a mixed-studies systematic review. *Clinical Simulation in Nursing* 84:101464, 2023.
- Hodges B. OSCE! Variations on a theme by Harden. *Medical Education* 37:1134-1140, 2003.
- Kim EJ, Koo YR, Nam IC. Patients and healthcare providers' perspectives on patient experience factors and a model of patient-centered care communication: a systematic review. *Healthcare* 12:1109, 2024.

mapping, in a systematic and transparent manner, the use of artificial intelligence technologies applied directly to the learning process in the development of clinical communication skills among undergraduate health professions students, as well as at describing the reported educational outcomes and the limitations associated with their use. The review will clarify the existing technologies and pedagogical approaches, the assessment instruments used, and the nature of the available evidence on educational outcomes, identifying gaps that may guide research priorities and curricular decisions.

CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest.

ACKNOWLEDGEMENTS

The authors thank Laura Mafalda Lopes, Technical Officer at the Nursing School of Porto, University of Porto, for the support provided in reviewing the search strategy, particularly regarding the adequacy of terms, controlled vocabulary, and adaptation of the strategy to the databases.

AUTHOR CONTRIBUTIONS

Conceptualization: Mariana Lima, Teresa Martins, Lígia Lima, Celeste Bastos; Methodology: Mariana Lima, Teresa Martins, Lígia Lima, Celeste Bastos; Formal analysis: Mariana Lima, Teresa Martins, Lígia Lima, Celeste Bastos; Investigation: Mariana Lima, Teresa Martins, Lígia Lima, Celeste Bastos; Writing – original draft preparation: Mariana Lima, Teresa Martins, Lígia Lima, Celeste Bastos; Writing – review and editing: Mariana Lima, Teresa Martins, Lígia Lima, Celeste Bastos. All authors have read and agreed to the published version of the manuscript.

- Kirkpatrick D, Kirkpatrick J. Evaluating training programs: the four levels. Berrett-Koehler Publishers, San Francisco, 2006.
- Kourkouta L, Papathanasiou I. Communication in nursing practice. *Materia Socio Medica* 26:65-67, 2014.
- Kung TH, Cheatham M, Medenilla A, Sillos C, De Leon L, Elepaño C, Madriaga M, Aggabao R, Diaz-Candido G, Maningo J, Tseng V. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. *PLOS Digital Health* 2:e0000198, 2023.
- Leonard M, Graham S, Bonacum D. The human factor: the critical importance of effective teamwork and communication in providing safe care. *Quality and Safety in Health Care* 13:i85-i90, 2004.
- Masters K. Artificial intelligence in medical education. *Medical Teacher* 41:976-980, 2019.
- McCabe C, Timmins F. Communication skills for nursing practice. Bloomsbury Publishing, London, 2013.
- McGaghie WC, Issenberg SB, Petrusa ER, Scalese RJ. A critical review of simulation-based medical education research: 2003-2009. *Medical Education* 44:50-63, 2010.
- McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS peer review of electronic search strategies: 2015 guideline statement. *Journal of Clinical Epidemiology* 75:40-46, 2016.
- Miller GE. The assessment of clinical skills/competence/performance. *Academic Medicine* 65:S63-S67, 1990.
- O'Daniel M, Rosenstein AH. Professional communication and team collaboration. In: Patient safety and quality: an evidence-based handbook for nurses, Hughes RG (ed). Agency for Healthcare Research and Quality, Rockville, Chapter 33:2008.
- Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. *Healthcare* 11:887, 2023.
- Stiefel F, Bourquin C, Salmon P, Ahtari Jeanneret L, Dauchy S, Ernstmann N, Grassi L, Libert Y, Vitinius F, Santini D, Ripamonti CI. Communication and support of patients and caregivers in chronic cancer care: ESMO clinical practice guideline. *ESMO Open* 9:103496, 2024.
- Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, Moher D, Peters MDJ, Horsley T, Weeks L, Hempel S, Akl E, Chang C, McGowan J, Stewart L, Hartling L, Aldcroft A, Wilson MG, Garritty C, Lewin S, Godfrey CM, Macdonald MT, Langlois EV, Soares-Weiser K, Moriarty J, Clifford T, Tunçalp Ö, Straus SE. PRISMA extension for scoping reviews: checklist and explanation. *Annals of Internal Medicine* 169:467-473, 2018.
- Vilelas J. Investigação: o processo de construção do conhecimento. Edições Sílabo, Lisboa, 2017.
- Zarei M, Eftekhari Mamaghani H, Abbasi A, Hosseini MS. Application of artificial intelligence in medical education: a review of benefits, challenges, and solutions. *Medicina Clínica Práctica* 7:100422, 2024.