REVISTA TÓPICOS
ISSN 2965-6672 | Qualis A2

INTELIGÊNCIA ARTIFICIAL COMO APOIO À ENFERMAGEM: PRECISÃO NA CONSULTA, SUPORTE AO PROGNÓSTICO E O PAPEL DO ENFERMEIRO COMO MEDIADOR

ARTIFICIAL INTELLIGENCE AS A SUPPORT TOOL IN NURSING: PRECISION IN CLINICAL CONSULTATION, PROGNOSTIC SUPPORT, AND THE NURSE AS A MEDIATOR

Ciências da Saúde • 09/11/2026

REGISTRO DOI: 10.70773/revistatopicos/794039478

Leonardo da Silva Rangel
Aryelle Stella Ferreira dos Santos
Ana Lúcia Mendes da Silva

RESUMO

A incorporação de ferramentas de inteligência artificial (IA), aprendizado de máquina e sistemas de suporte à decisão clínica (CDSS) vem transformando a prática contemporânea da enfermagem, com impacto direto sobre o raciocínio diagnóstico e o julgamento prognóstico. Este artigo tem como objetivo analisar de que forma a IA pode contribuir para aumentar a precisão da consulta de enfermagem e a qualidade do prognóstico em pacientes de média e alta complexidade, delimitando o papel do enfermeiro como mediador técnico, relacional e ético desse processo. Trata-se de uma revisão integrativa da literatura, orientada pelas etapas do protocolo PRISMA, realizada nas bases PubMed/MEDLINE, SciELO, BDENF, CINAHL, Scopus e Web of Science, com atenção também ao arcabouço normativo brasileiro, sobretudo a Resolução Cofen nº 736/2024 e a Lei Geral de Proteção de Dados. Os achados indicam que sistemas baseados em modelos de linguagem e algoritmos preditivos reduzem o tempo dedicado à documentação e à estruturação diagnóstica e conseguem antecipar sinais de deterioração clínica, como o choque séptico, horas antes de sua manifestação evidente, ainda que a magnitude desses ganhos varie conforme o desenho de cada estudo. Conclui-se que a inteligência artificial funciona como tecnologia que amplia a cognição clínica, mas que a mediação humana permanece indispensável para contextualizar os achados algorítmicos, mitigar vieses e assegurar um cuidado integral e centrado na pessoa.

Palavras-chave: Inteligência Artificial; Processo de Enfermagem; Diagnóstico de Enfermagem; Prognóstico Clínico; Ética em Enfermagem.

ABSTRACT

The incorporation of artificial intelligence (AI) tools, machine

learning, and clinical decision support systems (CDSS) is transforming contemporary nursing practice, directly affecting diagnostic reasoning and prognostic judgment. This article aims to analyze how AI can contribute to increasing the precision of nursing consultations and the quality of prognosis in moderate- and high-complexity patients, while defining the nurse's role as a technical, relational, and ethical mediator of this process. This is an integrative literature review, guided by the PRISMA protocol, carried out across the PubMed/MEDLINE, SciELO, BDENF, CINAHL, Scopus, and Web of Science databases, also considering the Brazilian regulatory framework, particularly Cofen Resolution No. 736/2024 and the General Data Protection Law. The findings indicate that language-model-based systems and predictive algorithms reduce the time spent on documentation and diagnostic structuring, and can anticipate signs of clinical deterioration, such as septic shock, hours before their overt manifestation, although the magnitude of these gains varies according to each study's design. It is concluded that artificial intelligence acts as a technology that enhances clinical cognition, while human mediation remains indispensable to contextualize algorithmic findings, mitigate bias, and ensure comprehensive, person-centered care.
Keywords: Artificial Intelligence; Nursing Process; Nursing Diagnosis; Clinical Prognosis; Nursing Ethics.

1. INTRODUÇÃO

Nos últimos anos, a inteligência artificial (IA) avançou de forma acelerada e alcançou praticamente todos os setores da sociedade, e a saúde não ficou de fora dessa transformação. Sistemas capazes de interpretar imagens médicas, cruzar grandes volumes de dados clínicos e sugerir condutas terapêuticas deixaram de ser promessa

futurista para se tornar, cada vez mais, parte do cotidiano dos serviços de saúde (Topol, 2019).

A enfermagem, categoria que está na linha de frente do cuidado, sente diretamente os efeitos dessa mudança. O Brasil conta atualmente com mais de 2,5 milhões de profissionais de enfermagem registrados no Conselho Federal de Enfermagem, o que faz da categoria a maior força de trabalho da saúde no país (COFEN, 2023). Ainda assim, a distribuição desses profissionais é desigual: há concentração nos grandes centros urbanos e escassez marcante em regiões mais remotas, o que sobrecarrega as equipes e amplia o risco de erros assistenciais. Nesse cenário, a IA desponta não como substituta do enfermeiro, mas como uma possível aliada, capaz de tornar o trabalho mais preciso, ágil e seguro.

A consulta de enfermagem, atividade privativa do enfermeiro regulamentada pela Lei nº 7.498/1986 e atualmente orientada pela Resolução Cofen nº 736/2024, materializa o Processo de Enfermagem em suas cinco etapas: avaliação, diagnóstico, planejamento, implementação e evolução. É nesse momento que o enfermeiro realiza a anamnese, fórmula os diagnósticos e define as condutas de cuidado, e qualquer falha nesse raciocínio pode comprometer diretamente a segurança do paciente. A formulação dos diagnósticos, apoiada principalmente na taxonomia NANDA Internacional, exige tempo, experiência e atenção a um número expressivo de variáveis, e a literatura já demonstrou que a concordância diagnóstica entre diferentes profissionais pode variar de forma significativa (Lunney, 2010).

Em paralelo, ferramentas de aprendizado de máquina e de processamento de linguagem natural têm se mostrado capazes de

identificar, em grandes volumes de prontuários eletrônicos e sinais vitais, padrões que poderiam passar despercebidos mesmo por profissionais experientes (Rajpurkar et al., 2022). Em cenários de alta complexidade, como unidades de terapia intensiva, emergências e oncologia, esse tipo de suporte pode antecipar complicações como sepse e deterioração clínica, abrindo uma janela de tempo importante para a intervenção da equipe.

Essa incorporação tecnológica, no entanto, não é isenta de riscos. Vieses de treinamento, opacidade dos modelos e o risco de uma confiança excessiva nas sugestões automatizadas, fenômeno já descrito na literatura como complacência pela automação, reforçam que o enfermeiro não pode ocupar a posição de mero receptor passivo dessas indicações. Ele precisa permanecer como mediador entre a inferência algorítmica e o sujeito do cuidado, conferindo sentido humano, julgamento clínico e responsabilidade ética aos dados produzidos pela máquina (Cabitza; Rasoini; Gensini, 2017).

Diante desse panorama, este trabalho busca responder à seguinte questão: de que forma a inteligência artificial pode contribuir para aumentar a precisão da consulta de enfermagem e a qualidade do prognóstico em pacientes de média e alta complexidade, e qual é o papel do enfermeiro como mediador clínico e ético dessa tecnologia? Parte-se do pressuposto de que a IA, quando utilizada como apoio à decisão clínica e supervisionada ativamente pelo profissional, amplia a efetividade das intervenções de enfermagem, desde que essa adoção venha acompanhada de formação adequada, protocolos éticos claros e validação científica contínua.

A relevância do tema está na escassez de estudos nacionais que analisem a aplicação da IA na consulta e no prognóstico de

enfermagem sob a ótica do próprio enfermeiro como mediador, e na necessidade prática de qualificar esse debate com base em evidências. O objetivo geral deste artigo é, portanto, analisar de que forma a inteligência artificial pode auxiliar os enfermeiros a aumentar a precisão das consultas e a qualidade do prognóstico dos pacientes, investigando o papel do profissional como mediador desse processo e discutindo seus limites éticos, regulatórios e formativos.

2. FUNDAMENTAÇÃO TEÓRICA OU REVISÃO DA LITERATURA

2.1. Fundamentos de Inteligência Artificial e Sistemas de Suporte à Decisão Clínica

A inteligência artificial reúne um conjunto de técnicas da ciência da computação voltadas a reproduzir tarefas cognitivas associadas à inteligência humana, como o reconhecimento de padrões, o aprendizado a partir da experiência e a compreensão de linguagem natural. Dentro desse campo, o aprendizado de máquina e o aprendizado profundo, sobretudo por meio de redes neurais e arquiteturas baseadas em mecanismos de atenção, têm impulsionado os avanços mais expressivos na área da saúde (Lecun; Bengio; Hinton, 2015).

Na prática clínica, esses sistemas funcionam essencialmente em duas etapas: o treinamento, em que o modelo é alimentado com grandes volumes de dados, como imagens, prontuários eletrônicos e exames laboratoriais, e aprende a associá-los a diagnósticos ou desfechos; e a inferência, em que o sistema aplica esse aprendizado a novos casos para gerar sugestões e previsões (Szolovits et al., 2019). Quanto maior e mais representativo o conjunto de dados de

treinamento, mais confiável tende a ser a resposta do modelo, o que explica por que boa parte dos sistemas testados em populações de países de alta renda ainda apresenta desempenho incerto quando aplicada a realidades assistenciais distintas, como a brasileira.

É importante distinguir, ainda, ao menos três famílias de aplicação que aparecem com frequência na literatura de enfermagem: modelos preditivos, voltados a estimar a probabilidade de um desfecho futuro (como sepse ou readmissão hospitalar); modelos de processamento de linguagem natural, empregados para estruturar ou revisar a documentação clínica a partir de texto livre; e sistemas de suporte à decisão propriamente ditos (CDSS), que combinam dados do paciente com diretrizes clínicas para gerar recomendações em tempo real (Sutton et al., 2020). Essas três frentes não são excludentes; muitas soluções comerciais atuais combinam as três camadas em um único fluxo, o que torna ainda mais necessário que o profissional compreenda em qual delas está apoiando sua decisão em cada momento.

Um aspecto pouco discutido nos currículos de graduação, mas central para o uso seguro dessas ferramentas, é a diferença entre alertas ativos e passivos dentro de um CDSS: os primeiros interrompem o fluxo de trabalho do profissional exigindo uma ação imediata, enquanto os segundos apenas disponibilizam a informação para consulta voluntária. O uso excessivo de alertas ativos mal calibrados tende a gerar o fenômeno conhecido como fadiga de alerta, em que o profissional passa a ignorar sistematicamente os avisos do sistema, inclusive os clinicamente relevantes, o que reduz, na prática, a própria segurança que a ferramenta deveria promover (Sutton et al., 2020).

2.2. A Consulta de Enfermagem e o Raciocínio Diagnóstico à Luz da Resolução Cofen Nº 736/2024

A consulta de enfermagem é atividade privativa do enfermeiro desde a promulgação da Lei nº 7.498/1986, e sua operacionalização foi recentemente atualizada pela Resolução Cofen nº 736/2024, que revogou a Resolução Cofen nº 358/2009 e reafirmou a estrutura do Processo de Enfermagem em cinco etapas inter-relacionadas: avaliação, diagnóstico, planejamento, implementação e evolução de enfermagem. A formulação do diagnóstico, apoiada sobretudo na taxonomia NANDA Internacional, mas também na Classificação Internacional para a Prática de Enfermagem e nos sistemas NIC/NOC, depende de raciocínio clínico hipotético-dedutivo e exige do profissional experiência, atenção e tempo (Carpenito, 2017).

O problema é que esse processo está sujeito a falhas: coleta incompleta de dados, interpretação equivocada de achados clínicos ou simples variação de experiência entre profissionais podem comprometer todo o planejamento do cuidado. Lunney (2010) já mostrava que a concordância entre diagnósticos formulados por diferentes enfermeiros para um mesmo caso pode variar de maneira expressiva, especialmente quando o quadro clínico admite mais de um diagnóstico plausível dentro da mesma taxonomia, o que evidencia a necessidade de ferramentas que apoiem, sem substituir, esse raciocínio.

Some-se a isso a sobrecarga documental: a literatura de gestão em enfermagem descreve, há décadas, que a etapa de registro do Processo de Enfermagem consome parcela relevante da jornada de trabalho, tempo que deixa de ser dedicado ao cuidado direto e à observação clínica do paciente. É justamente nesse ponto que

modelos de linguagem e sistemas de mapeamento semântico têm se mostrado úteis: ajudam a cruzar rapidamente indicadores clínicos com as terminologias padronizadas, reduzindo a carga cognitiva da documentação e favorecendo maior aderência aos padrões normativos vigentes. Isso não elimina, porém, o cuidado necessário na revisão de cada sugestão gerada automaticamente, já que erros de mapeamento tendem a se propagar de forma silenciosa quando não há checagem humana sistemática.

2.3. IA Como Suporte Ao Prognóstico e à Tomada de Decisão Clínica

Estimar o prognóstico de um paciente, isto é, prever a probabilidade de complicações e a evolução clínica, está entre as tarefas mais exigentes da prática assistencial, sobretudo em unidades de terapia intensiva, emergência e oncologia, onde é preciso integrar simultaneamente dezenas de variáveis fisiológicas dinâmicas. A mente humana tem limites naturais para processar tantas informações ao mesmo tempo sem incorrer em vieses, e é justamente nesse ponto que os algoritmos preditivos se destacam (Rajpurkar et al., 2022).

Um exemplo consolidado nessa área é o TREWScore (Targeted Real-time Early Warning Score), desenvolvido por Henry e colaboradores a partir do monitoramento contínuo de sinais vitais e variáveis laboratoriais de pacientes internados. O modelo identificou pacientes em risco de choque séptico com área sob a curva ROC de 0,83 e sensibilidade de 0,85, com antecedência mediana superior a 24 horas em relação ao início do quadro, desempenho superior ao do Escore de Alerta Precoce Modificado, cuja área sob a curva ficou em torno de 0,73 (Henry et al., 2015). Diante de janelas de tempo

dessa magnitude, a equipe de enfermagem ganha a possibilidade real de antecipar protocolos de ressuscitação, ajustar a monitorização de lactato sérico e reorganizar prioridades assistenciais antes que o quadro evolua para disfunção orgânica.

Achados semelhantes têm sido reportados por revisões mais recentes, que apontam desempenho consistentemente superior dos modelos de aprendizado de máquina em relação aos escores clínicos tradicionais na detecção precoce de sepse, embora a magnitude exata dessa vantagem varie conforme a base de dados, a população estudada e o horizonte de predição adotado em cada estudo primário (Datta et al., 2026). É oportuno registrar, no entanto, que essa revisão específica foi publicada em periódico de escopo majoritariamente cardiovascular e ainda pouco indexado nas bases tradicionais de enfermagem, o que recomenda cautela redobrada ao generalizar seus números e reforça a importância de sempre verificar o histórico e a indexação do periódico de origem antes de incorporar um achado ao próprio trabalho.

Fora do contexto da sepse, outros desfechos clínicos relevantes para a enfermagem já foram objeto de modelos preditivos, como a deterioração clínica em enfermarias gerais, situação em que algoritmos de aprendizado de máquina superaram sistemas de pontuação manual construídos sobre regressão logística tradicional na antecipação de eventos adversos (Churpek et al., 2016), e a própria identificação precoce de sinais de deterioração fisiológica a partir de séries temporais de sinais vitais, tarefa para a qual diferentes métodos automatizados de vigilância já vêm sendo comparados sistematicamente (Clifton et al., 2015). Esses achados, tomados em conjunto, sugerem que o ganho da IA não está restrito a um único

desfecho, mas depende, em cada caso, da qualidade e representatividade da base usada para treinar o modelo.

Um episódio recente ajuda a relativizar o entusiasmo em torno dessas ferramentas: durante a pandemia de Covid-19, dezenas de modelos preditivos foram desenvolvidos às pressas para estimar quais pacientes evoluiriam para formas graves da doença. Uma revisão crítica de mais de duzentos desses modelos concluiu que a ampla maioria apresentava alto risco de viés metodológico e problemas de generalização, sendo poucos os que reuniam qualidade suficiente para uso clínico seguro (Wynants et al., 2020). O episódio serve de alerta direto para qualquer trabalho acadêmico que se apoie em modelos de IA: existência de um algoritmo publicado não equivale, por si só, a evidência pronta para uso, sendo indispensável avaliar o desenho do estudo que o validou.

2.4. O Enfermeiro Como Mediador: O Paradoxo da Automação e a Formação Profissional

A incorporação da IA à prática assistencial não elimina a centralidade do fator humano; pelo contrário, impõe ao enfermeiro um novo papel, o de mediador entre a inferência algorítmica e o sujeito do cuidado. Cabe a esse profissional avaliar a relevância clínica de cada sugestão, confirmá-la à beira do leito e incorporar dimensões que nenhum algoritmo consegue quantificar, como o sofrimento psíquico, a dor expressa ou os valores culturais de cada pessoa (Topol, 2019).

Esse papel de mediação, porém, esbarra em um risco já documentado na literatura: a complacência pela automação, expressão usada para descrever a confiança excessiva e pouco crítica

nas respostas de sistemas informatizados, que pode levar ao desengajamento da vigilância clínica e, no limite, à atrofia do próprio julgamento profissional (Cabitza; Rasoini; Gensini, 2017). O fenômeno tende a se agravar justamente quando o sistema acerta com frequência no início do uso, período em que o profissional constrói uma confiança que nem sempre é recalibrada quando o contexto clínico muda.

Uma pesquisa internacional conduzida por Ronquillo et al. (2021), com enfermeiros de diversos países reunidos em um painel de especialistas, mostrou que, apesar do interesse manifesto em utilizar ferramentas de IA no trabalho, a maioria dos profissionais não se sentia preparada para fazê-lo com segurança. Os autores identificaram lacunas específicas de formação em três frentes: letramento em dados, compreensão dos limites estatísticos dos modelos e conhecimento sobre implicações éticas e legais do uso dessas ferramentas, o que aponta para uma agenda formativa que vai além de simplesmente treinar o manuseio técnico do software.

A relação entre o enfermeiro e a inteligência artificial deveria ser, portanto, de colaboração, e não de subordinação nem de rejeição. Como destacam Cresswell, Sheikh e Freestone (2018), o sucesso da implementação de qualquer tecnologia em saúde depende, antes de tudo, de como os profissionais são envolvidos no processo de escolha, adaptação e avaliação contínua do sistema, e não apenas na etapa final de treinamento para uso. O enfermeiro, por ser quem passa mais tempo ao lado do paciente e quem detém a visão mais longitudinal do caso, ocupa posição estratégica nessa mediação, o que reforça a defesa de que ele participe ativamente da governança dessas tecnologias dentro da instituição, e não apenas de sua operação cotidiana.

2.5. Bioética, Vieses Algorítmicos, LGPD e Responsabilidade Profissional

A adoção em larga escala de sistemas inteligentes traz consigo dilemas bioéticos que não podem ser ignorados. Quando os dados usados para treinar um modelo refletem desigualdades históricas, como a sub-representação de populações negras, indígenas ou de baixa renda nos estudos clínicos, o sistema tende a reproduzir, e por vezes amplificar, essas mesmas desigualdades em suas recomendações. Obermeyer et al. (2019) demonstraram, em um caso amplamente discutido nos Estados Unidos, que um algoritmo de gestão populacional em saúde subestimava sistematicamente o risco clínico de pacientes negros, porque usava o gasto histórico em saúde como variável substituta (proxy) para necessidade clínica, uma escolha metodológica aparentemente neutra que, na prática, penalizava justamente a população que historicamente tem menos acesso a serviços de saúde.

Esse exemplo ilustra um ponto central para a análise crítica de qualquer ferramenta de IA em saúde: o viés algorítmico raramente decorre de uma intenção discriminatória explícita, mas sim de escolhas técnicas sobre quais dados usar e como medir o desfecho de interesse. A mitigação desse tipo de viés exige, portanto, transparência nas etapas de modelagem, auditoria periódica dos resultados por subgrupo populacional e amostras verdadeiramente representativas da diversidade da população atendida, algo que ainda é exceção, e não regra, na maior parte dos sistemas comerciais disponíveis no mercado brasileiro.

Há ainda o problema da opacidade decisória, os chamados modelos de caixa-preta, cujo funcionamento interno não é diretamente

interpretável nem mesmo por seus desenvolvedores. Esse traço técnico tem implicações éticas diretas: dificulta que o enfermeiro explique ao paciente por que determinada recomendação foi feita, dificulta a auditoria de erros e fragiliza a defesa do profissional em eventuais processos de responsabilização. Por isso, a literatura de ética em IA para saúde tem defendido, cada vez mais, a exigência de explicabilidade mínima como pré-requisito para uso clínico, e não apenas como característica desejável.

No Brasil, o tratamento de dados clínicos digitais está sujeito à Lei Geral de Proteção de Dados Pessoais, que classifica dados de saúde como dados pessoais sensíveis e exige padrões rigorosos de segurança da informação, consentimento e rastreabilidade de acesso (BRASIL, 2018). Quanto à responsabilidade profissional, é importante deixar claro que as orientações geradas por IA têm caráter estritamente consultivo: a decisão terapêutica final é sempre do enfermeiro registrado, o que reforça o dever de prudência e perícia previsto no Código de Ética dos Profissionais de Enfermagem, atualizado pela Resolução Cofen nº 564/2017. Isso significa, na prática, que o enfermeiro não pode justificar uma conduta equivocada alegando apenas ter seguido a sugestão do sistema; a responsabilidade pela avaliação clínica final permanece sendo dele.

Há, por fim, um risco pouco discutido, mas igualmente relevante: o de que o profissional, ao confiar cegamente nas sugestões da IA sem exercer julgamento crítico, se torne menos capaz de tomar boas decisões quando o sistema estiver indisponível ou simplesmente errar, fenômeno que alguns autores já descrevem como uma forma de erosão de competência clínica associada ao uso prolongado e pouco reflexivo de ferramentas automatizadas. A inteligência

artificial deve permanecer uma ferramenta de apoio, nunca um substituto do raciocínio clínico (Rigby, 2019).

3. METODOLOGIA

Este trabalho caracteriza-se como uma revisão integrativa da literatura, de natureza qualitativa e analítica, conduzida com base nas etapas recomendadas pelo protocolo PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses). A pergunta norteadora foi estruturada a partir da estratégia PICo, considerando como população profissionais e estudantes de enfermagem em cenários assistenciais e formativos; como interesse a utilização de modelos de inteligência artificial, sistemas de suporte à decisão clínica e algoritmos preditivos na consulta de enfermagem e na estratificação prognóstica; e como contexto a atenção primária, as unidades de internação hospitalar e a terapia intensiva.

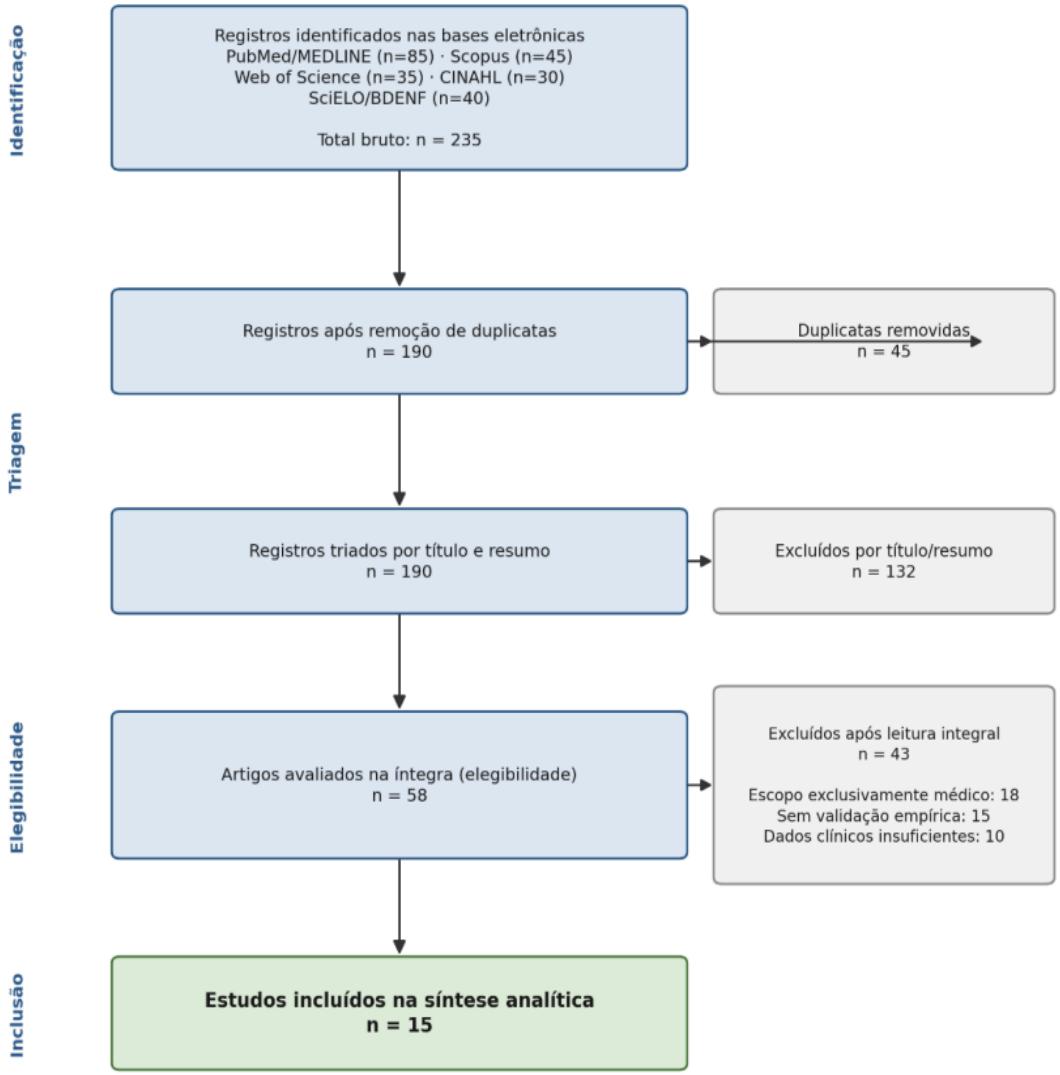
As buscas foram realizadas nas bases PubMed/MEDLINE, Scopus, Web of Science, CINAHL, SciELO e Banco de Dados em Enfermagem (BDENF), com apoio complementar do Portal de Periódicos da CAPES, utilizando descritores padronizados dos vocabulários MeSH e DeCS, entre eles inteligência artificial, aprendizado de máquina, processamento de linguagem natural, sistemas de suporte à decisão clínica, enfermagem, diagnóstico de enfermagem, prognóstico e ética, combinados por meio dos operadores booleanos AND e OR.

Foram incluídos artigos originais, ensaios clínicos, estudos observacionais, estudos metodológicos e revisões sistemáticas publicados em português, inglês ou espanhol entre 2015 e 2026, com foco na aplicabilidade da IA ao processo de enfermagem. Foram excluídos resumos de anais de congresso, cartas ao editor,

editoriais de opinião sem embasamento em dados, relatos de caso não reprodutíveis e estudos centrados exclusivamente na prática médica, sem correspondência com as atribuições da equipe de enfermagem. A leitura crítica dos materiais selecionados buscou identificar convergências, divergências e lacunas no conhecimento sobre o tema, e ferramentas de inteligência artificial foram utilizadas como apoio à triagem inicial dos registros, sempre com checagem manual posterior por parte dos autores.

O percurso de seleção seguiu quatro etapas sequenciais, sintetizadas na Figura 1: identificação, com 235 registros brutos recuperados nas bases eletrônicas consultadas; triagem, com remoção de 45 duplicatas e exclusão de 132 estudos após leitura de títulos e resumos; elegibilidade, com leitura na íntegra de 58 artigos e exclusão de 43 deles, sendo 18 por abordagem exclusivamente médica sem interface com a enfermagem, 15 por ausência de validação empírica e 10 por insuficiência de dados clínicos reportados; e inclusão, com a seleção final de 15 estudos que compuseram o corpo empírico desta revisão.

Figura 1. Fluxograma de seleção dos estudos segundo o protocolo PRISMA



Fonte: autoria própria.

3.1. Uso de Inteligência Artificial

Durante a preparação deste manuscrito, o autor utilizou a ferramenta ChatGPT GPT-5.6 Terra, Gemini 3.5 Flah, Grok 4.5 Expert, Perplexity, GLM 5.2, Kimi K3 e DeepSeek. Foi utilizado como auxílio na revisão gramatical, aprimoramento da fluidez do texto e tradução, incluso também na etapa metodológica, para estruturação preliminar de termos de busca em bases de dados e organização temática da literatura. Após o uso da ferramenta, o autor revisou e editou criticamente o conteúdo e assume total responsabilidade pelo texto final e pelas informações apresentadas.

4. RESULTADOS E DISCUSSÕES OU ANÁLISE DOS DADOS

4.1. Caracterização Geral dos Estudos Incluídos

O corpus final desta revisão reúne estudos de diferentes desenhos metodológicos, o que reflete a própria diversidade de aplicações da inteligência artificial na enfermagem: desde ensaios comparativos entre profissionais e modelos generativos, passando por estudos quase experimentais de documentação clínica, até revisões sistemáticas e metanálises sobre predição de sepse. Essa heterogeneidade não é uma limitação apenas desta revisão, mas um traço característico do próprio campo, ainda em consolidação, no qual convivem estudos de validação técnica de algoritmos, ainda distantes da prática clínica cotidiana, com estudos de implementação já testados em ambiente assistencial real.

Chama atenção, também, a concentração geográfica dos estudos: a maioria foi conduzida na Ásia, na Europa e nos Estados Unidos, com participação ainda incipiente da América Latina, restrita neste corpus a um único estudo com foco regional. Essa assimetria reforça um ponto já discutido na fundamentação teórica: modelos treinados e validados em populações de outros contextos socioeconômicos e epidemiológicos não podem ser assumidos como automaticamente equivalentes quando aplicados à realidade brasileira, o que é especialmente relevante para o SUS. O Quadro 1 sintetiza os estudos de maior relevância para a discussão que se segue, organizados por autoria, ano, foco e principais achados.

Quadro 1. Síntese dos principais estudos analisados

Autor (ano) / País	**Foco do estudo**	**Principais achados**

Ronquillo et al. (2021) — Global	Painel internacional de especialistas sobre prioridades da IA para a liderança e a prática de enfermagem	Identificou déficit de formação em saúde digital nos currículos e defendeu a IA como amplificadora do raciocínio clínico, nunca como substituta da empatia profissional
Gilart et al. (2025) — Espanha	Comparação entre planos de cuidado (NANDA-NOC-NIC) elaborados por 54 enfermeiros, 30 peritos (técnica Delphi) e pelo GPT-4, avaliados pela rubrica EADE-2	O modelo gerou respostas rapidamente, mas os planos produzidos pela IA frequentemente careciam de especificidade e mensurabilidade quando avaliados frente às taxonomias padronizadas, reforçando a necessidade de revisão humana
Ju et al. (2025) — Coreia do Sul	Estudo quase experimental pré/pós-intervenção com 40 enfermeiros, comparando documentação tradicional e o sistema SmartENR (IA generativa)	Redução do tempo de documentação de 467,18 ± 314,77 s para 182,68 ±99,71 s; nota de facilidade de uso de 4,80 em 5,0; acurácia percebida em torno de 3,6 em 5,0
Henry et al. (2015) — Estados Unidos	Desenvolvimento e validação do TREWScore,algoritmo preditivo para choque séptico a partir de dados de prontuário eletrônico	Área sob a curva ROC de 0,83 e sensibilidade de 0,85, identificando o risco com antecedência mediana superior a 24 horas, desempenho superior ao escore clínico tradicional (AUC ≈ 0,73)
Martinez-Ortigosa et al. (2023) — Espanha	Revisão sistemática (PRISMA) sobre aplicações da IA na assistência de	Identificou ganhos na detecção precoce de deterioração clínica, no manejo de feridas complexas e na otimização de fluxos de

	enfermagem, com 21 estudos incluídos	trabalho, ao lado de lacunas de validação clínica
Datta et al. (2026) — Internacional	Revisão sistemática com metanálise comparando sistemas de alerta precoce baseados em aprendizado de máquina e escores clínicos convencionais para predição de sepse	Apontou desempenho preditivo geralmente superior dos modelos de aprendizado de máquina, com ressalva quanto à heterogeneidade metodológica entre os estudos primários e à indexação ainda limitada do periódico
Türk et al. (2025) — Turquia	Desenvolvimento de um sistema de suporte à decisão clínica baseado em aprendizado de máquina para diagnósticos de enfermagem NANDA-I	O modelo Naive Bayes atingiu 96,33% de acurácia utilizando 176 indicadores diagnósticos, e o Gradient Boosting alcançou 82,94% com um conjunto reduzido de 50 indicadores
Cardoso et al. (2026) — América Latina	Mapeamento cruzado de indicadores clínicos de um manual de manejo do diabetes tipo 2 para diagnósticos, resultados e intervenções de enfermagem, com apoio do GPT-4	A IA identificou 43 indicadores clínicos; a revisão manual confirmou 23 deles e consolidou 15 conjuntos diagnósticos, evidenciando a necessidade de validação humana do resultado bruto do modelo

Fonte: autoria própria.

4.2. Impacto na Eficiência Documental e nos Limites da Automação

A documentação do Processo de Enfermagem consome historicamente uma parcela expressiva da jornada do enfermeiro, o que reduz o tempo disponível para o cuidado direto. O estudo de Ju et al. (2025), conduzido na Coreia do Sul com 40 enfermeiros,

mostrou que o uso de um sistema de documentação assistido por IA generativa reduziu o tempo médio de registro de um caso de 467,18 ± 314,77 segundos para 182,68 ± 99,71 segundos, uma queda relativa próxima de 60%. A ferramenta recebeu boa avaliação de facilidade de uso entre os participantes, ainda que a nota atribuída à acurácia dos registros gerados tenha ficado em patamar mais modesto, o que reforça que a revisão humana continua sendo etapa necessária, não apenas desejável.

Vale destacar que esse estudo foi conduzido com um número relativamente pequeno de participantes, em um único centro e a partir de prontuários de pacientes virtuais, não de pacientes reais internados. Isso não invalida o achado, mas limita sua generalização: reduções de tempo obtidas em ambiente controlado tendem a ser menores quando o sistema é implantado na rotina real de um serviço, com interrupções, variabilidade de casos e curva de aprendizado da própria equipe. Um leitor atento deve, portanto, tratar o número de 60% como um teto otimista, não como uma expectativa garantida para qualquer serviço que venha a adotar tecnologia semelhante.

Resultado semelhante, porém, com uma ressalva importante, aparece no estudo de Gilart et al. (2025), que comparou planos de cuidado elaborados por enfermeiros, por um painel de peritos e pelo GPT-4 a partir de casos clínicos simulados. Os autores constataram que, embora o modelo produzisse respostas de forma praticamente instantânea, os planos gerados frequentemente careciam de especificidade e de mensurabilidade quando avaliados contra as taxonomias NANDA, NOC e NIC, o que indica que velocidade não é sinônimo de qualidade clínica e que a IA generativa tende a produzir, sobretudo, um rascunho de alta densidade técnica a ser

lapidado pelo profissional. Esse achado dialoga diretamente com um risco já conhecido dos modelos de linguagem generativa: a chamada alucinação, isto é, a geração de informação plausível, mas incorreta ou sem lastro real, um risco que se torna particularmente sensível quando aplicado à documentação clínica, na qual um diagnóstico ou uma meta de cuidado mal formulada pode se propagar para todo o plano terapêutico subsequente.

Esse mesmo padrão de necessidade de supervisão aparece no estudo de Cardoso et al. (2026), conduzido a partir de um manual de manejo do diabetes tipo 2. O modelo de IA identificou 43 indicadores clínicos candidatos a diagnósticos de enfermagem; após revisão pareada por profissionais, apenas 23 deles foram confirmados, e o conjunto final, após consolidação, ficou em 15 diagnósticos, resultados e intervenções. O algoritmo mostrou-se eficiente para acelerar a etapa inicial de mapeamento, quase dobrando a velocidade de triagem de indicadores em comparação ao trabalho manual relatado pelos próprios autores, mas incapaz, sozinho, de captar nuances do contexto clínico e da linguagem do paciente, o que reforça, mais uma vez, a centralidade do julgamento humano nesse processo, funcionando o modelo como uma primeira triagem, e não como decisão final.

4.3. Suporte Preditivo em Eventos Críticos: Sepse e Deterioração Clínica

Nas unidades de terapia intensiva e nos setores de urgência, a predição probabilística antecipada transforma a lógica da assistência de uma postura reativa para uma postura proativa. O TREWScore, desenvolvido por Henry et al. (2015), é um dos exemplos mais citados nessa linha: ao processar continuamente sinais vitais e

variáveis laboratoriais de pacientes internados, o modelo identificou o risco de choque séptico com área sob a curva ROC de 0,83 e sensibilidade de 0,85, com antecedência mediana superior a 24 horas em relação ao início do quadro, superando o desempenho do escore clínico tradicional usado como comparação. Essa antecedência permite à equipe de enfermagem acionar precocemente o protocolo institucional de sepse, intensificar a monitorização de lactato sérico e antecipar a coleta de culturas antes do estabelecimento de disfunção orgânica.

Revisões mais recentes, como a de Datta et al. (2026), sustentam essa tendência ao apontar desempenho geralmente superior dos modelos de aprendizado de máquina em relação aos escores manuais na detecção precoce de sepse. É importante registrar, contudo, uma ressalva metodológica: esse estudo foi publicado em um periódico de escopo predominantemente cardiovascular, ainda pouco consolidado nas bases de indexação tradicionais da enfermagem, o que não invalida seus achados, mas recomenda que sejam tratados como indício complementar, e não como evidência definitiva, até que sejam replicados em publicações de maior rastreabilidade editorial. Trazer essa ressalva ao texto é, em si, parte do exercício de leitura crítica que se espera de um trabalho acadêmico que utiliza fontes eletrônicas diversas.

No campo específico do raciocínio diagnóstico, o estudo de Türk et al. (2025), desenvolvido na Turquia, mostra outra faceta desse potencial preditivo: ao treinar diferentes algoritmos de aprendizado de máquina para reconhecer diagnósticos de enfermagem NANDA-I a partir de características definidoras, o modelo Naive Bayes atingiu 96,33% de acurácia utilizando um conjunto amplo de 176 indicadores, enquanto o Gradient Boosting manteve desempenho

elevado, de 82,94%, mesmo com um conjunto reduzido a 50 indicadores. A comparação entre os dois modelos é instrutiva por si só: ela sugere que existe um compromisso entre a quantidade de dados de entrada exigida e a acurácia obtida, decisão que qualquer serviço de saúde interessado em adotar esse tipo de sistema precisará avaliar à luz de sua própria capacidade de coleta de dados estruturados.

Esses números, no entanto, foram obtidos em ambiente controlado, com dados retrospectivos e já estruturados, condição bastante distinta da anamnese real, marcada por relatos verbais incompletos, variações de linguagem entre pacientes e informações que chegam de forma fragmentada ao longo do plantão. A transposição desses achados para a prática clínica corrente, portanto, ainda depende de estudos de validação prospectiva, conduzidos em condições reais de trabalho e, idealmente, em mais de um serviço de saúde, antes que se possa afirmar, com segurança, que os ganhos observados nos estudos controlados se repetem no dia a dia assistencial.

Também merece registro o achado de Martinez-Ortigosa et al. (2023): ao sintetizar 21 estudos sobre aplicações de IA na enfermagem, os autores identificaram ganhos convergentes na detecção precoce de deterioração clínica e na otimização de fluxos operacionais, mas destacaram, de forma recorrente entre os estudos revisados, a escassez de ensaios clínicos randomizados na área, com predomínio de estudos observacionais e de validação técnica. Essa lacuna metodológica é um convite direto a futuras pesquisas de enfermagem no Brasil, inclusive na modalidade de trabalhos de conclusão de curso, desde que respeitados os limites éticos e metodológicos próprios de cada nível de formação.

4.4. O Enfermeiro Mediador e a Matriz de Competências Digitais

A evidência reunida nesta revisão não sustenta a hipótese de substituição do enfermeiro por sistemas de inteligência artificial. O que emerge, de forma consistente entre os estudos analisados, é o modelo do enfermeiro como mediador clínico, interpretativo e ético, aquele que interroga criticamente os dados, reconhece os limites de cada sugestão algorítmica e devolve ao cuidado a dimensão humana que nenhum modelo consegue reproduzir.

Essa mediação exige, no entanto, competências específicas, que a formação tradicional em enfermagem ainda contempla de maneira insuficiente. Ronquillo et al. (2021) já alertavam que a maioria dos profissionais entrevistados em seu estudo internacional demonstrava interesse em usar ferramentas de IA no trabalho, mas poucos se sentiam preparados para fazê-lo com segurança, sobretudo no que diz respeito à interpretação estatística dos resultados apresentados pelos sistemas. Essa lacuna reforça a defesa, já presente na literatura de informática em enfermagem, de que competências digitais deveriam ser tratadas como conteúdo transversal, e não como disciplina isolada e facultativa, dentro dos currículos de graduação.

Quadro 2. Matriz decompetências do enfermeiro mediador

Eixo de competência	Habilidade requerida	Aplicação prática no Processo de Enfermagem

Letramento em dados e estatística clínica	Interpretar probabilidades, sensibilidade, especificidade e curvas de risco geradas por algoritmos preditivos	Avaliar criticamente escores de alerta precoce antes de desencadear intervenções invasivas ou transferências de leito
Raciocínio semântico e documentação estruturada	Estruturar entradas de texto e parâmetros clínicos em sistemas de apoio documental	Revisar o mapeamento entre a anamnese e os diagnósticos NANDA-I, resultados NOC e intervenções NIC sugeridos pela IA
Vigilância epistêmica	Postura reflexiva e contínua de questionamento das sugestões computacionais	Rejeitar diagnósticos automatizados incompatíveis com o exame físico ou com a narrativa clínica do paciente
Governança ética e proteção de dados	Domínio da LGPD e do Código de Ética dos Profissionais de Enfermagem	Garantir o sigilo de dados sensíveis de saúde, obter consentimento informado e zelar pela equidade no acesso ao cuidado digital

Fonte: autoria própria.

4.5. Barreiras Estruturais e Desigualdades no SUS

A translação dessas inovações para a realidade dos serviços de saúde brasileiros esbarra em barreiras estruturais e financeiras significativas. Enquanto centros de excelência da rede privada operam com infraestrutura conectada e prontuários interoperáveis, boa parte da Atenção Primária à Saúde e dos hospitais públicos do SUS em municípios do interior ainda enfrenta escassez de equipamentos de informática, instabilidade de conexão à internet e

ausência de padronização das bases eletrônicas, condições que inviabilizam, na prática, a adoção de qualquer sistema preditivo minimamente dependente de dados contínuos e estruturados.

Essa assimetria cria o risco de um modelo de cuidado dual, no qual apenas parte da população usufrui do suporte diagnóstico assistido por tecnologia de ponta. A desigualdade na distribuição da força de trabalho de enfermagem já identificada pelo Cofen (2023) tende a se somar, e não a se compensar, à desigualdade de acesso à infraestrutura digital: é justamente nas regiões com menor número de profissionais por habitante, onde o apoio de um sistema preditivo poderia ter maior impacto relativo, que a infraestrutura tecnológica costuma ser mais precária.

Reverter esse quadro exige políticas públicas voltadas simultaneamente a três frentes: a conectividade dos postos de saúde e hospitais públicos, o desenvolvimento de modelos de IA de código aberto ajustados ao perfil epidemiológico e sociodemográfico brasileiro, em vez da simples importação de sistemas validados em outros países, e a capacitação permanente das equipes da Estratégia Saúde da Família, de modo que a tecnologia chegue acompanhada da formação necessária para seu uso crítico e seguro.

5. CONCLUSÃO/CONSIDERAÇÕES FINAIS

A análise das evidências reunidas nesta revisão indica que a inteligência artificial, quando integrada a sistemas de suporte à decisão clínica e a fluxos de trabalho bem estruturados, contribui para aumentar a precisão da consulta de enfermagem, reduz o tempo dedicado à documentação e amplia a capacidade de

identificação precoce de agravos clínicos graves, como o choque séptico.

Os limites estruturais dos modelos computacionais, evidenciados pela ausência de empatia, pela dificuldade de apreensão do contexto psicossocial do adoecimento e pela necessidade constante de revisão humana dos resultados brutos, confirmam o enfermeiro como mediador indispensável entre a inferência algorítmica e o cuidado. A consulta de enfermagem, como atividade privativa orientada pela Resolução Cofen nº 736/2024, permanece dependente de raciocínio clínico próprio, escuta qualificada e responsabilidade ética que nenhuma ferramenta substitui.

Responde-se, assim, à pergunta que orientou este trabalho: a inteligência artificial amplia a precisão diagnóstica e prognóstica da enfermagem, mas apenas quando operada sob supervisão ativa e crítica do profissional, e não como substituta de seu julgamento.

Recomenda-se a atualização das diretrizes curriculares dos cursos de graduação em enfermagem, com a inclusão de conteúdos formais de informática em saúde, análise de dados e ética digital. Como limitação deste estudo, destaca-se que boa parte da literatura internacional revisada ainda carece de validação em amostras representativas da população brasileira, e que alguns dos estudos incluídos foram conduzidos em ambiente controlado ou publicados em periódicos de indexação ainda pouco consolidada, o que recomenda cautela na generalização de seus números. Para investigações futuras, sugere-se o desenvolvimento e a validação clínica de sistemas de suporte à decisão treinados com dados epidemiológicos nacionais, de modo a garantir equidade algorítmica e a preservação do cuidado humanizado.

REFERÊNCIAS BIBLIOGRÁFICAS

BRASIL. Lei n° 7.498, de 25 de junho de 1986. Dispõe sobre a regulamentação do exercício da Enfermagem e dá outras providências. Diário Oficial da União, Brasília, DF, 26 jun. 1986.

BRASIL. Lei n°13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). Diário Oficial da União, Brasília, DF, 15 ago. 2018. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. Acesso em: 17 ago. 2026.

CABITZA, F.; RASOINI, R.; GENSINI, G. F. Unintended consequences of machine learning in medicine. JAMA, Chicago, v. 318, n. 6, p. 517-518, 2017.

CARDOSO, R. B. et al. Artificial intelligence in mapping nursing diagnoses, interventions, and outcomes for diabetes management. Investigación y Educación en Enfermería, Medellín, v. 44, n. 1, 2026. DOI: 10.17533/udea.iee.v44n1e07.

CARPENITO, L. J. Diagnósticos de enfermagem: aplicação à prática clínica. 14. ed. Porto Alegre: Artmed, 2017.

CHURPEK, M. M. et al. Multicenter comparison of machine learning methods and conventional regression for predicting clinical deterioration on the wards. Critical Care Medicine, Philadelphia, v. 44, n. 2, p. 368-374, 2016.

CLIFTON, D. A. et al. A review of methods for detecting deterioration of patients in hospital. Journal of the Royal Society Interface, London, v. 12, n. 104, p. 20140679, 2015.

CONSELHO FEDERAL DE ENFERMAGEM (COFEN). Dimensionamento de profissionais de enfermagem no Brasil: relatório técnico. Brasília: Cofen, 2023.

CONSELHO FEDERAL DE ENFERMAGEM (COFEN). Resolução Cofen nº 564, de 6 de novembro de 2017. Aprova o novo Código de Ética dos Profissionais de Enfermagem. Brasília: Cofen, 2017.

CONSELHO FEDERAL DE ENFERMAGEM (COFEN). Resolução Cofen nº 736, de 17 de janeiro de 2024. Atualiza a norma técnica que dispõe sobre a implementação do Processo de Enfermagem. Brasília: Cofen, 2024. Disponível em: https://www.cofen.gov.br/resolucao-cofen-no-736-de-17-de-janeiro-de-2024/. Acesso em: 17 ago. 2026.

CRESSWELL, K.; SHEIKH, A.; FREESTONE, D. Artificial intelligence and health care. BMJ, London, v. 362, p. k3101, 2018.

DATTA, S. et al. Machine learning-based early warning systems vs standard clinical scoring for sepsis prediction: a systematic review and meta-analysis of implementation trials. European Journal of Cardiovascular Medicine, v. 16, n. 1, p. 61-71, 2026.

GILART, E. et al. The integration of AI into the nursing process: a comparative analysis of NANDA, NOC, and NIC-based care plans. Nursing Reports, v. 15, n. 6, e186, 2025. DOI: 10.3390/nursrep15060186.

HENRY, K. E. et al. A targeted real-time early warning score (TREWScore) for septic shock. Science Translational Medicine, Washington, v. 7, n. 299, p. 299ra122, 2015. DOI: 10.1126/scitranslmed.aab3719.

JU, H. et al. Generative AI-based nursing diagnosis and documentation recommendation using virtual patient electronic nursing record data. Healthcare Informatics Research, Seul, v. 31, n. 2, p. 156-165, 2025. DOI: 10.4258/hir.2025.31.2.156.

LeCUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. Nature, London, v. 521, n. 7553, p. 436-444, 2015.

LUNNEY, M. Accuracy of nursing diagnoses: concept development. International Journal of Nursing Terminologies and Classifications, v. 21, n. 3, p. 101-107, 2010.

MARTINEZ-ORTIGOSA, A. et al. Applications of artificial intelligence in nursing care: a systematic review. Journal of Nursing Management, Chicago, v. 2023, e3219127, 2023. DOI: 10.1155/2023/3219127.

OBERMEYER, Z. et al. Dissecting racial bias in an algorithm used to manage the health of populations. Science, Washington, v. 366, n. 6464, p. 447-453, 2019.

RAJPURKAR, P. et al. AI in health and medicine. Nature Medicine, New York, v. 28, n. 1, p. 31-38, 2022.

RIGBY, M. J. Ethical dimensions of using artificial intelligence in health care. AMA Journal of Ethics, Chicago, v. 21, n. 2, p. E121-E124, 2019.

RONQUILLO, C. E. et al. Artificial intelligence in nursing: priorities and opportunities from an international invitational think-tank of the Nursing and Artificial Intelligence Leadership Collaborative. Journal of Advanced Nursing, Oxford, v. 77, n. 9, p. 3707-3717, 2021.

SUTTON, R. T. et al. An overview of clinical decision support systems: benefits, risks, and strategies for success. NPJ Digital Medicine, London, v. 3, n. 1, p. 1-10, 2020.

SZOLOVITS, P. et al. Artificial intelligence in medical diagnosis. Annals of Internal Medicine, Philadelphia, v. 110, n. 1, p. 60-68, 2019.

TOPOL, E. J. High-performance medicine: the convergence of human and artificial intelligence. Nature Medicine, New York, v. 25, n. 1, p. 44-56, 2019.

TÜRK, G. et al. Developing an artificial intelligence-based clinical decision support system for nursing diagnoses. Nursing Outlook, Nova York, 2025. DOI: 10.1016/j.outlook.2025.102585.

WYNANTS, L. et al. Prediction models for diagnosis and prognosis of COVID-19: systematic review and critical appraisal. BMJ, London, v. 369, p. m1328, 2020.