



BACHARELADO EM MEDICINA

**CLÉSIA MICHELLI ROCHA PIMENTEL
IZABELLY THAYS RAMOS SILVA
MATHEUS GOMES CUNHA MENEZES**

**A UTILIZAÇÃO DO PROCESSAMENTO DE LINGUAGEM NATURAL NO
DIAGNÓSTICO DE DELIRIUM EM PACIENTES IDOSOS ADMITIDOS NA
UNIDADE DE TERAPIA INTENSIVA: Uma Revisão Sistemática sem Metanálise**

Jaboatão dos Guararapes

2025

**CLÉSIA MICHELLI ROCHA PIMENTEL
IZABELLY THAYS RAMOS SILVA
MATHEUS GOMES CUNHA MENEZES**

**A UTILIZAÇÃO DO PROCESSAMENTO DE LINGUAGEM NATURAL NO
DIAGNÓSTICO DE DELIRIUM EM PACIENTES IDOSOS ADMITIDOS NA
UNIDADE DE TERAPIA INTENSIVA: Uma Revisão Sistemática sem Metanálise**

Trabalho de Conclusão de Curso apresentado
como requisito para o cumprimento da
disciplina de TCC II, para obtenção de título
em Bacharel em Medicina.

Orientador(a): Profa. Dra. Isaura Romero
Peixoto

Jaboatão dos Guararapes

2025

**A UTILIZAÇÃO DO PROCESSAMENTO DE LINGUAGEM
NATURAL NO DIAGNÓSTICO DE DELIRIUM EM PACIENTES
IDOSOS ADMITIDOS NA UNIDADE DE TERAPIA INTENSIVA: UMA
REVISÃO SISTEMÁTICA SEM METANÁLISE**

**THE USE OF NATURAL LANGUAGE PROCESSING IN THE
DIAGNOSIS OF DELIRIUM IN ELDERLY PATIENTS ADMITTED
TO THE INTENSIVE CARE UNIT: A SYSTEMATIC REVIEW
WITHOUT META-ANALYSIS**

Resumo

O *delirium* é uma condição cerebral aguda e de curso flutuante, potencialmente fatal em pessoas idosas, que está associado, sobretudo quando subdiagnosticado pelos profissionais de saúde, a inúmeras complicações. Para tanto, algoritmos de Processamento de Linguagem Natural podem ser úteis na conversão de dados não estruturados, disponíveis em prontuários eletrônicos, em formatos estruturados para pesquisa médica, vigilância em saúde e suporte à decisão clínica. Esta pesquisa propôs-se a analisar as repercussões e os desafios da utilização de ferramentas de Processamento de Linguagem Natural no diagnóstico de *delirium* em pacientes idosos assistidos em unidades de cuidado intensivo. Trata-se de uma revisão sistemática da literatura, sem metanálise, orientada pela estratégia *Population or Problem, Variables and Outcomes* e que utilizou Biblioteca Virtual em Saúde, PubMed e SciELO como bases de dados para busca de artigos, realizada entre os dias 4 de março de 2025 e 21 de maio de 2025. Foram avaliados 5 estudos, com 18.647 pacientes, que demonstraram a utilidade da aplicação de estratégias de Processamento de Linguagem Natural, a partir de registros eletrônicos de saúde, no aprimoramento do diagnóstico e da estratificação de risco de *delirium* em pacientes idosos admitidos em Unidades de Terapia

Intensiva, corroborando uma avaliação automatizada, inovadora e que não impõe carga de trabalho adicional à equipe médica. A utilização do Processamento de Linguagem Natural, apesar de ter se mostrado eficiente e promissora no apoio ao diagnóstico de *delirium* em pessoas idosas no cenário de terapia intensiva, requer abordagens interdisciplinares, validação clínica e investimento em infraestrutura tecnológica.

Palavras-chave: Inteligência artificial; Delírio; Pessoas idosas.

Abstract

Delirium is an acute and potentially fatal brain condition with a fluctuating course in elderly individuals, which is associated with numerous complications, especially when underdiagnosed by health professionals. To this end, Natural Language Processing algorithms can be useful in converting unstructured data, available in electronic medical records, into structured formats for medical research, health surveillance and clinical decision support. This research aimed to analyze the repercussions and challenges of using Natural Language Processing tools in the diagnosis of delirium in elderly patients treated in intensive care units. This is a systematic literature review, without meta-analysis, guided by the Population or Problem, Variables and Outcomes strategy and which used the Biblioteca Virtual em Saúde, PubMed and SciELO as databases to search for articles, carried out between March 4, 2025 and May 21, 2025. Five studies were evaluated, with 18,647 patients, which demonstrated the usefulness of applying Natural Language Processing strategies, from electronic health records, in improving the diagnosis and risk stratification of delirium in elderly patients admitted to Intensive Care Units, corroborating an automated, innovative assessment that does not impose additional workload on the medical team. The use of Natural Language Processing, despite having proven efficient and promising in supporting the diagnosis of delirium in elderly people in the intensive care setting, requires

interdisciplinary approaches, clinical validation and investment in technological infrastructure.

Keywords: Artificial intelligence; Delirium; Elderly people.

INTRODUÇÃO

De acordo com o Estatuto da Pessoa Idosa (GOVERNO FEDERAL, 2022), define-se como população idosa o agrupamento de pessoas com idade igual ou superior a 60 anos. Este grupo etário é usuário regular dos serviços de saúde, em frequência que aumenta proporcionalmente ao avançar da idade, e constitui uma população singular no cuidado hospitalar. Quando comparados a adultos de 40 a 59 anos, as pessoas idosas são até duas vezes mais suscetíveis à hospitalização (RODRIGUES *et al.*, 2023), condição que, somada às particularidades desta faixa etária, corrobora a progressão da condição inicial de saúde para desfechos complexos e preveníveis, como o *delirium*.

O *delirium* é uma condição cerebral aguda e de curso flutuante, que se caracteriza por déficit de atenção e alteração do nível de consciência, com diminuição da capacidade de concentração e memória, desorientação e distúrbios de percepção, pensamento e comportamento (CECHINEL *et al.*, 2022). Potencialmente fatal em pessoas idosas, é uma disfunção grave e que está associada, sobretudo quando subdiagnosticada pelos profissionais da saúde, a maior tempo de hospitalização, dano funcional e cognitivo a longo prazo, perda de dispositivos invasivos, maior índice de mortalidade e aumento dos custos para o sistema de saúde (LIMA *et al.*, 2021).

Apesar de não totalmente elucidada, a fisiopatologia do *delirium* explica-se por mecanismos patológicos distintos, que compreendem desde alterações na concentração de neurotransmissores (acetilcolina, serotonina, dopamina, melatonina, noradrenalina e ácido gama-aminobutírico), até secreção de citocinas, com aumento da permeabilidade da barreira

hematoencefálica e diminuição do metabolismo oxidativo cerebral (CARVALHO *et al.*, 2022). Ademais, esta condição está intimamente relacionada a fatores predisponentes, como idade e presença de comorbidades, e fatores precipitantes, a exemplo do uso de medicamentos prévios e de alterações do ciclo sono-vigília (ALVES; MARQUES; VERNAGLIA, 2024).

Frequente em pessoas idosas hospitalizadas, a prevalência de *delirium* varia entre 21% e 35% no período pós-operatório e entre 60% e 85% em pacientes admitidos em Unidade de Terapia Intensiva (UTI) (FU *et al.*, 2022). Haja vista sua incidência, a quinta edição do Manual Diagnóstico e Estatístico de Transtornos Mentais (DSM-V) estabelece o diagnóstico de *delirium* a partir de cinco critérios, a saber: perturbação na atenção, déficit cognitivo, desenvolvimento da condição em curto período de tempo, inexistência de transtorno neurocognitivo preexistente e ausência de outra condição médica capaz de justificar o quadro (YOUNG *et al.*, 2022).

No entanto, o diagnóstico sistemático de *delirium* em pacientes assistidos em UTI, através da aplicação direta dos critérios supracitados, ainda é considerado um desafio (YOUNG *et al.*, 2022). Para isso, algoritmos de Processamento de Linguagem Natural (PLN) podem ser úteis na conversão de informações obtidas a partir de linguagem natural em formatos estruturados para pesquisa médica, vigilância em saúde e suporte à decisão clínica (KOLECK *et al.*, 2019).

O PLN é um subcampo da Ciência da Computação, desenvolvido na década de 1960, que se concentra no entendimento e na interpretação da linguagem humana, por meio de informações extraídas de dados não estruturados e do reconhecimento de padrões entre palavras em uma sentença (BERGE *et al.*, 2023). Dessa maneira, essa ferramenta pode ser utilizada para complementar campos de dados estruturados, como códigos de sintomas, medicamentos e achados laboratoriais, com ideias obtidas de notas não estruturadas (notas de

visita e relatórios de radiologia, por exemplo) de profissionais da saúde, disponíveis em registros médicos eletrônicos (HOMBURG *et al.*, 2023).

A mineração de informações de registros eletrônicos, portanto, pode tornar-se uma ferramenta útil no suporte à decisão clínica e no fornecimento de uma Medicina de precisão (LE GLAZ *et al.*, 2021). Apesar disso, as abordagens utilizadas para Processamento de Linguagem Natural na assistência à saúde esbarram, ainda, em falhas de desempenho, eficiência e transparência, uma vez que estão subordinadas à experiência médica humana no estabelecimento de regras determinísticas e na manutenção de dicionários especializados para a extração de dados clínicos (BERGE *et al.*, 2023).

Ainda assim, em serviços de saúde, a adoção de métodos de PLN é potencialmente eficaz e revolucionária na identificação de doenças (HOMBURG *et al.*, 2023) e pode ser empregada, em sistemas práticos, para aprimorar a fenotipagem clínica (HART *et al.*, 2021), ainda que obstáculos devam ser ultrapassados.

Dessa maneira, esta pesquisa se propõe a analisar e descrever as repercussões e os desafios da utilização do Processamento de Linguagem Natural no diagnóstico de *delirium* em pacientes idosos admitidos em Unidade de Terapia Intensiva.

METODOLOGIA

Bases de dados e critérios de pesquisa

Trata-se de uma revisão sistemática da literatura, sem metanálise, orientada pelo *Preferred Reporting Items for Systematic reviews and Meta-Analyses* (PRISMA) e registrada na plataforma PROSPERO (CRD420251014103). Através da estratégia PVO (*Population or Problem, Variables and Outcomes*) (Quadro 1), a revisão norteou-se na seguinte questão de pesquisa: quais os benefícios do PLN no diagnóstico de *delirium* em pacientes idosos admitidos em UTI?

Quadro 1: Estratégia PVO.

Acrônimo	Significado	Revisão
P	População/Pacientes	Pacientes idosos admitidos em UTI
V	Variável	Uso de PLN
O	Desfecho/ <i>Outcome</i>	Diagnóstico de <i>delirium</i>

Fonte: Elaborado pelos autores, 2025.

As bases de dados utilizadas para pesquisa dos artigos compreenderam Biblioteca Virtual em Saúde (BVS), *National Library of Medicine* (PubMed) e *Scientific Electronic Library Online* (SciELO). A busca de artigos estendeu-se entre os dias 4 de março de 2025 e 21 de maio de 2025, utilizando-se os descritores “Processamento de Linguagem Natural”, “Diagnóstico”, “*Delirium*”, “Idoso” e “Unidades de Terapia Intensiva”, delimitados a partir da plataforma DeCS/MeSH, nos idiomas português, inglês e espanhol. Além disso, as referências dos artigos elegíveis foram avaliadas, manualmente, em busca de outros estudos potencialmente relevantes. Em seguida, os artigos obtidos foram exportados para o *Rayyan Software* para análise posterior.

Critérios de seleção

Como critérios de inclusão, foram selecionados artigos publicados entre 2019 e 2025, disponíveis gratuitamente, na íntegra, escritos nos idiomas português, inglês e espanhol, e que utilizaram ferramentas de PLN no diagnóstico de *delirium* em amostras de pacientes, de ambos os sexos, com idade mediana, ao diagnóstico, igual ou superior a 60 anos. Além disso, foram incluídos pacientes admitidos em Unidades de Terapia Intensiva ou com avaliação prévia pelo CAM-UTI. Em contrapartida, foram excluídos estudos em modelo animal, artigos com resultados pouco claros, artigos de revisão, dissertações, teses, editoriais, painéis, relatórios e comentários.

Seleção dos estudos

Três pesquisadores independentes (CP, IS e MM) buscaram artigos de forma autônoma e, em seguida, identificaram os estudos relevantes, a partir da análise dos títulos e resumos. Posteriormente, os artigos selecionados foram avaliados na íntegra, considerando-se os critérios de inclusão. O coeficiente Kappa (K) para avaliação de concordância entre os pesquisadores foi calculado usando o pacote estatístico SPSS versão 17.0.

Qualidade metodológica

Na etapa de seleção, a qualidade metodológica dos artigos selecionados foi verificada, de maneira a analisar a confiabilidade dos métodos e resultados das pesquisas. Para tanto, a escala PEDro foi utilizada para elencar os estudos de boa e excelente qualidade metodológica, utilizando uma pontuação de 5 como ponto de corte. A escala PEDro compreende 11 itens pré-estabelecidos, a saber: 1) validade externa (item 1), que, apesar de avaliar a aplicabilidade do estudo clínico, não é incluído no escore total da escala; 2) validação interna (itens 2 a 9), que analisa o risco de viés; 3) descrição estatística (itens 10 e 11), que avalia se as informações estatísticas são suficientes para a interpretação dos resultados. Esses itens são categorizados em “aplicáveis” ou “não aplicáveis”, resultando em uma pontuação total que varia entre 0 e 10 pontos.

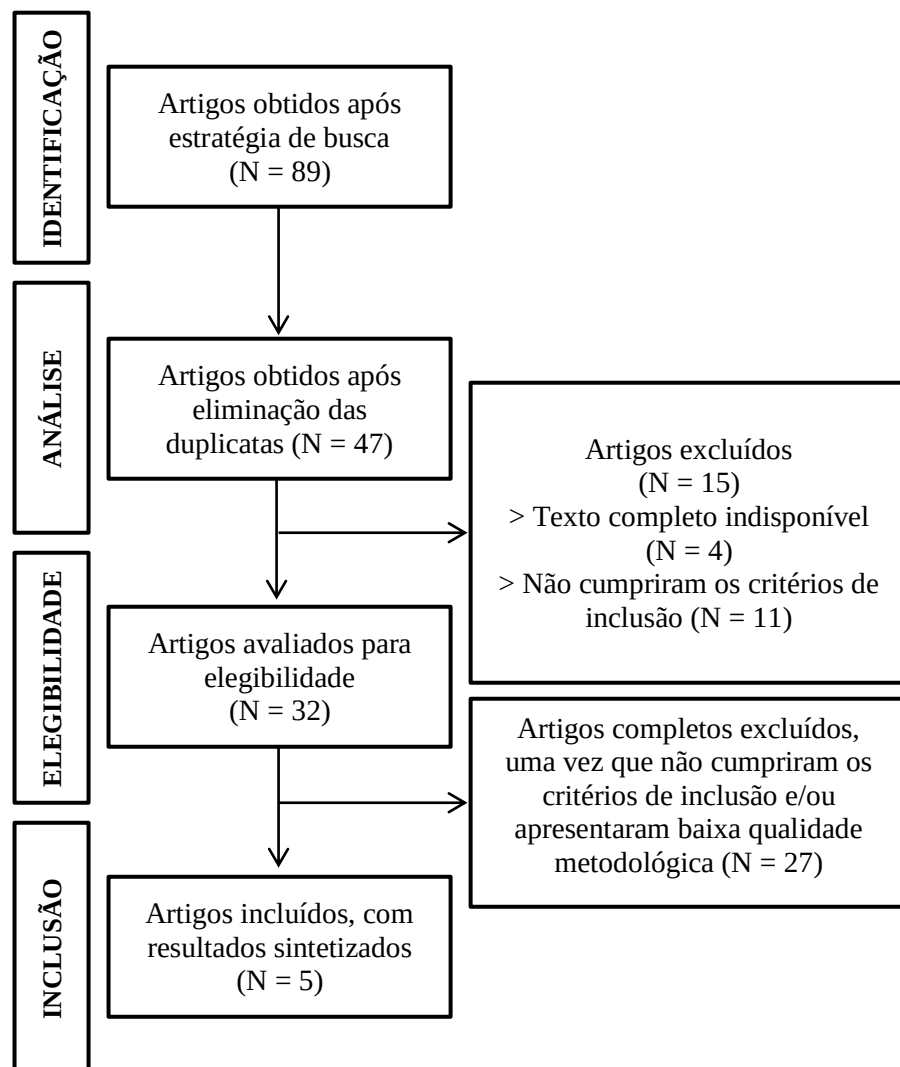
Coleta dos dados

A coleta de dados deu-se amparada na estratégia PVO e respeitando o campo de aplicação das pesquisas (UTI). Três pesquisadores independentes (CP, IS e MM) realizaram a busca, de forma a assegurar a relevância das informações coletadas, minimizar erros potenciais na transcrição e verificar a precisão dos dados colhidos.

RESULTADOS

Ao serem utilizadas as palavras-chave nas plataformas de busca, foram obtidos 89 artigos (45 na BVS, 43 na PubMed e 1 na SciELO). Após leitura do título, 33 artigos foram selecionados e, destes, 32 foram eleitos após leitura do resumo e encaminhados para avaliação de texto completo. Destes, no entanto, apenas 5 atenderam aos critérios de inclusão previamente descritos, com um total de 18.647 participantes (Figura 1). Os trabalhos apresentaram risco de viés moderado a baixo e boa qualidade metodológica pela escala PEDro (Tabela 1).

Figura 1: Fluxograma dos artigos analisados.



Fonte: Elaborada pelos autores, 2025.

Tabela 1: Características dos artigos analisados.

Autor/ Ano de publicação/ País	Grupo- controle	Grupo-teste	Nível de atenção	Ferramenta de avaliação	Resultado	PEDro
Castro <i>et al.</i> (2021) EUA	Coorte de treinamento (N = 2.152)	Coorte de teste (N = 755)	NR	Modelo estatístico de predição de <i>delirium</i> e PLN	AM e PLN aplicados a prontuários eletrônicos podem estratificar o risco de <i>delirium</i> em pacientes com COVID- 19	8
Chen <i>et al.</i> (2022) China	Grupo não <i>delirium</i> (N = 669)	Grupo <i>delirium</i> (N = 110)	NR	Instrumento de palavras- chave chinesas para triagem semiautomá- tica de <i>delirium</i> usando PLN	Melhora da detecção de <i>delirium</i> em idosos	8
Miyazawa <i>et al.</i> (2024) Japão	Grupo não <i>delirium</i> (N = 252)	Grupo <i>delirium</i> (N = 21)	UTI	PLN e AM	PLN e AM podem oferecer uma avaliação automatizada do risco de <i>delirium</i> , sem impor carga adicional de trabalho à equipe	9
Young <i>et al.</i> (2021) Austrália	Grupo negativo para PLN-Dx-BD (N = 7.267)	Grupo positivo para PLN-Dx-BD (N = 5.108)	UTI	PLN	PLN identificou rapidamente pacientes de UTI com alto risco para <i>delirium</i>	8
Young <i>et al.</i> (2022) Austrália	Grupo negativo para CAM-UTI e PLN-Dx-BD (N = 1.028)	Grupo positivo somente para CAM-UTI, somente para PLN-Dx-BD ou positivo para CAM- UTI e PLN- Dx-BD (N = 1.285)	UTI	CAM-UTI e PLN	Mais pacientes apresentaram resultados positivos para PLN- Dx-BD, em comparação ao CAM-UTI	8

NR: não reportado; AM: Aprendizado de Máquina; PLN-Dx-BD: Distúrbio Comportamental Diagnosticado por Processamento de Linguagem Natural; CAM-UTI: Método de Avaliação de Confusão para UTI.

Fonte: Elaborada pelos autores, 2025.

No estudo de Castro *et al.* (2021), um modelo estatístico foi desenvolvido para ser aplicado a dados de registros eletrônicos de saúde, de maneira a prever as chances de um paciente hospitalizado com COVID-19 desenvolver *delirium*. Para tanto, foram utilizadas duas coortes, uma para desenvolvimento do modelo (N = 2.152) e outra para validação independente (N = 755), com idade média de 62,9 anos e prevalência de homens (N = 1.536) e brancos (N = 1.503). Considerando apenas pacientes em cuidados intensivos (N = 655), na coorte agrupada, 26,4% (N = 173) apresentaram diagnóstico de *delirium*, enquanto os 73,6% (N = 482) dos casos restantes não apresentaram.

O modelo de Castro *et al.* (2021) baseou-se em características do paciente na admissão, sinais vitais, exames laboratoriais, histórico médico e, considerando a subnotificação de *delirium* em dados codificados, apoiou-se, também, em uma definição que incluiu PLN para identificar pacientes delirantes durante a internação. Assim, essa ferramenta mostrou-se eficaz na predição de *delirium* em pacientes hospitalizados com COVID-19 e demonstrou utilidade em uma gama de limiares de risco, ainda que um esforço de predição concentrado em UTI ofereça maior precisão, quando comparada à equivalência combinada encontrada neste estudo (CASTRO *et al.*, 2021).

Outro estudo, realizado por Chen *et al.* (2022), desenvolveu um instrumento para triagem de *delirium* em 779 pacientes, com idade média de 82,5 anos, a partir de palavras-chave, que podem ser obtidas em registros médicos eletrônicos. No estudo de Chen *et al.* (2022), os pacientes com documentos médicos e de enfermagem acessíveis foram alocados no grupo médico e enfermeiro (PN), enquanto àqueles sem documentos de enfermagem, foram classificados no grupo médico e não enfermeiro (PNN). A aplicação do modelo de triagem demonstrou que a precisão pode ser aprimorada substancialmente quando incorporados os documentos de enfermagem.

Quando comparado a outro instrumento de predição de *delirium* em unidades de

cuidado intensivo, o CHART-DEL-UTI (*Chart-based Delirium Identification Instrument* em Unidade de Terapia Intensiva), a validade do modelo de Chen *et al.* (2022) mostrou-se superior no grupo PN, além de que a confiabilidade teste-reteste foi maior, sugerindo mais estabilidade. Para Chen *et al.* (2022), apesar de, como instrumento de triagem, o desempenho dessa nova ferramenta ser inferior ao das escalas de cabeceira (DSM-V), o modelo de aprendizado de máquina adotado apresentou excelente viabilidade e aplicabilidade.

O estudo de Miyazawa *et al.* (2024) identificou fatores de risco capazes de prever o início de *delirium* em pacientes com COVID-19, através de técnicas de PLN e AM aplicadas a registros de enfermagem, considerando informações dos pacientes e comentários subjetivos dos enfermeiros. Para isso, foram incluídos pacientes internados no *Juntendo University Hospital* (N = 273), um serviço de cuidados intensivos, dos quais 7,7% (N = 21) apresentavam *delirium*. A idade média dos pacientes ao diagnóstico de *delirium* foi de 73,5 anos e 61,1% eram do sexo masculino.

A partir dos registros de enfermagem, Miyazawa *et al.* (2024) identificou fatores úteis na predição de *delirium*, a saber: 1) alterações nos comportamentos e nas condições relacionados à eliminação (diurese e evacuação) e 2) comportamentos e condições anormais como preditores de *delirium*. Este estudo concluiu que outros métodos de avaliação, como CAM-UTI, *Intensive Care Delirium Screening Checklist* (ICDSC), *Delirium Rating Scale-Revised-98* (DRS-R-98) e *Memorial Delirium Assessment Scale* (MDAS), são mais sensíveis quando comparados aos fatores identificados no modelo atual. No entanto, instrumentos de avaliação convencionais exigem esforço adicional da equipe médica, enquanto os preditores identificados neste estudo podem ser coletados de maneira objetiva e automática, sem trabalho adicional.

Outro estudo, de Young *et al.* (2021), utilizou técnica de PLN para identificar palavras sugestivas de distúrbios comportamentais, potencialmente associados ao *delirium*, a

partir de notas de progresso digitadas por profissionais de saúde. Os dados foram coletados de registros eletrônicos de saúde de 12.375 pacientes, internados na UTI do *Austin Hospital*, com idade média de 64 anos, dos quais 61,6% eram do sexo masculino e 57% estavam em ventilação mecânica.

Young *et al.* (2021) observaram que pacientes com PLN-Dx-BD apresentavam características clínicas consistentes com uma coorte de alto risco para *delirium*, como idade mais avançada, maior propensão a internações médicas ou não planejadas, diagnóstico neurológico, diagnóstico de doença crônica, maior predisposição a ventilação mecânica e terapia de substituição renal, além de sinais vitais e exames laboratoriais da admissão frequentemente alterados. Os resultados sugeriram, então, que o software de PLN pode ser utilizado para delimitar, a partir de palavras descritivas, uma população de pacientes gravemente enfermos, com risco aumentado para o desenvolvimento de distúrbios comportamentais, possivelmente *delirium*.

O estudo de Young *et al.* (2022), fundamentado em registros de pacientes do *Australian and New Zealand Intensive Care Society Adult ICU Patient Database* (N = 2.313), com idade mediana de 63 anos e prevalência do sexo masculino (60,2%), aplicou PLN, simultaneamente ao CAM-UTI, aos dados de progresso médico e de enfermagem e às anotações de cuidadores, com o objetivo de comparar prevalência, características clínicas, tratamento com antipsicóticos e desfecho de pacientes com PLN-Dx-BD e com avaliação de *delirium* baseada no CAM-UTI.

Segundo o estudo de Young *et al.* (2022), mais pacientes apresentaram resultados positivos para PLN-Dx-BD (N = 1.246), quando comparados àqueles com CAM-UTI positivo (N = 578). Os participantes foram classificados em quatro grupos: 1) negativo para ambas as avaliações (44,4%); 2) positivo somente para CAM-UTI (1,7%); 3) positivo somente para PLN-Dx-BD (30,6%) e; 4) positivo para ambas as avaliações (23,3%).

Para Young *et al.* (2022), a distribuição dos pacientes sugere que o PLN-Dx-BD é capaz de identificar mais facilmente comportamento anormal e estado cognitivo perturbado, uma vez que captura continuamente dados de comportamento, se comparado ao CAM-UTI. Dependendo dos fatores de risco e do fenótipo de *delirium*, este estudou corroborou descobertas anteriores, sugerindo que o uso de CAM-UTI pode atrasar o diagnóstico de *delirium*, quando comparado a avaliações não estruturadas, tendo em vista que a prevalência de distúrbios comportamentais detectados pelo CAM-UTI foi menor do que com a PLN-Dx-BD.

DISCUSSÃO

Nesta revisão sistemática sem metanálise, cinco pesquisas analisaram 18.647 pacientes, com o objetivo de avaliar os reflexos da utilização de técnicas de Processamento de Linguagem Natural no diagnóstico de *delirium* em pacientes idosos admitidos em unidades de cuidado intensivo. Os estudos considerados nesta revisão demonstraram que a análise de registros eletrônicos de saúde apresenta utilidade clínica e pode impactar, significativamente, os índices de morbidade e mortalidade a longo prazo.

O *delirium* é uma condição frequente em pacientes graves e desfechos clínicos negativos estão associados em curto e longo prazos (PAULINO; POVOA; SERAFIM, 2020), com uma incidência estimada entre 11% e 66%, respectivamente, em idosos hospitalizados, a depender do ambiente de internação e dos instrumentos de triagem (JAYANAMA; KAEWPONGSA; RUANGRITCHANKUL, 2025). O padrão-ouro para o diagnóstico de *delirium* compreende a avaliação por um profissional psiquiatra, utilizando-se dos critérios descritos no DSM-V (Quadro 2), o que não constitui uma alternativa exequível na prática clínica. Além disso, tendo em vista o curso flutuante dessa condição, avaliações seriadas são fundamentais, associadas à revisão de prontuários e à discussão com familiares

(HAYHURST; HUGHES; PANDHARIPANDE, 2016).

Quadro 2: Critérios diagnósticos de *delirium*, segundo o DSM-V.

Critério A	Perturbação da atenção (capacidade reduzida de direcionar, focalizar, manter e mudar a atenção) ou da consciência (orientação reduzida).
Critério B	Condição de início rápido e por curto período de tempo, em geral de horas a dias, com tendência a flutuar ao longo do dia.
Critério C	Alteração em, no mínimo, mais de um domínio cognitivo (memória e aprendizagem, desorientação, linguagem, percepção ou perceptomotor).
Critério D	As perturbações não podem ser justificadas por um transtorno neurocognitivo prévio ou em desenvolvimento e não ocorrem no contexto de coma.
Critério E	Há evidências na história, no exame físico ou em exames laboratoriais de que a perturbação é resultado de uma condição clínica subjacente, intoxicação ou abstinência de substância, exposição à toxina ou múltiplas etiologias.

Fonte: Adaptado da 5ª edição do Manual Diagnóstico e Estatístico de Transtornos Mentais, 2025.

Entretanto, a revisão manual de prontuários médicos é demorada e dispendiosa e, por isso, softwares de Processamento de Linguagem Natural podem ser ferramentas úteis na extração de dados clínicos disponíveis em registros eletrônicos (LI *et al.*, 2021). O modelo de predição de risco de *delirium* desenvolvido por Castro *et al.* (2021), por exemplo, obteve discriminação e calibração promissoras em uma coorte independente de pacientes diagnosticados com COVID-19 e com idade mediana de 62,9 anos.

Castro *et al.* (2021) avaliou o modelo desenvolvido em ambientes clínicos independentes e sua aplicação em prontuários eletrônicos de saúde, alimentados na admissão do paciente, mostrou-se calibrada para as taxas de ocorrência verificadas, com desempenho consistente com as características e o contexto clínico do paciente. Nesse sentido, o Aprendizado de Máquina aplicado a registros médicos pode prever o surgimento

subsequente de *delirium* em pacientes hospitalizados com COVID-19, demonstrando utilidade clínica em uma ampla gama de limiares de risco.

A abordagem de Castro *et al.* (2021), no entanto, pode minimizar a precisão preditiva de *delirium* em Unidades de Terapia Intensiva, de modo que um esforço de predição focado seja mais preciso do que a equivalência combinada encontrada. Além disso, a precoce estratificação de risco, que considera os dados clínicos disponíveis somente à admissão hospitalar, pode postergar o redirecionamento de indivíduos com alto risco para desenvolvimento de *delirium*.

Uma pesquisa ambientada em unidades de cuidado intensivo brasileiras demonstrou que os pacientes foram rastreados para *delirium* em, apenas, 37% das UTIs, o que corrobora o subdiagnóstico e a avaliação ineficaz (BISPO; SOUZA, 2024). O estudo de Chen *et al.* (2022), então, utilizou uma ferramenta de triagem semiautomática e em tempo real, de um minuto, baseado em palavras-chave para detectar *delirium* por computador, com excelente viabilidade e custo-eficiência.

A metodologia adotada por Chen *et al.* (2022) demonstrou ser promissora na identificação de *delirium* e apresentou alta confiabilidade teste-reteste, validação interna aceitável e excelente aplicabilidade e aceitabilidade. Nesta abordagem, pontuações mais altas sugerem probabilidade aumentada de *delirium* e a predição pode ser aprimorada significativamente quando incorporados dados de enfermagem e com o aprimoramento dos prontuários médicos. Como ferramenta de triagem, seu desempenho é inferior ao de escalas convencionais, mas apresenta potencial para o manejo preciso e padronizado do *delirium*.

O instrumento desenvolvido por Chen *et al.* (2022), todavia, apresenta várias limitações, a saber: 1) foi influenciado pela qualidade dos registros de saúde; 2) baseou-se, totalmente, em palavras-chave chinesas; 3) não incluiu muitos algoritmos relacionados ao início agudo e ao caráter flutuante do *delirium*; 4) apresentou lacunas quanto ao seu

desempenho em subtipos de *delirium*; e 5) foi aplicado em amostra com desequilíbrio de gênero (71,9% da população amostral era do sexo masculino).

Desenvolveram-se ferramentas (CAM-UTI e ICDSC), a partir do DSM-V, com o intuito de operacionalizar a triagem de *delirium* em ambientes de terapia intensiva e documentar os resultados obtidos em notas de progresso eletrônicas (MIRANDA *et al.*, 2023). Apesar de bem delimitadas e consolidadas, a aplicabilidade na identificação sistemática de *delirium* em pacientes idosos submetidos a cuidados intensivos ainda é conflituosa, haja vista que exigem trabalho adicional da equipe médica.

Como alternativa a esse entrave, Miyazawa *et al.* (2024) utilizou um modelo de aprendizado de máquina para reconhecer preditores de *delirium* de forma objetiva e automática, sem demandar esforço adicional da equipe e com potencial para detectar achados não percebidos pela análise de dados estruturados. Este estudo identificou fatores úteis na predição do início de *delirium*, a partir de registros de enfermagem de pacientes com COVID-19, que, no entanto, foram menos sensíveis quando comparados aos métodos de avaliação convencionais, sugerindo, assim, que este pode não ser um modelo imediatamente eficaz.

No estudo de Miyazawa *et al.* (2024), o “Grupo 1” abrangeu comportamentos e condições relacionados à eliminação, incluindo palavras como “banheiro”, “micção” e “constipação”, e considerou a falta de controle fisiológico como agente causal ou resultante do *delirium*. O “Grupo 2”, por sua vez, compreendeu comportamentos e condições anormais do paciente, a partir de termos como “sensor” e “chamada de enfermeira”, considerando restrição dos movimentos corporais como fator de risco para desenvolvimento do quadro. Por fim, o “Grupo 3” englobou a expressão subjetiva de dor ou gravidade, com expressões como “dor”, “engasgo”, “falta de ar”, “febre”, “insônia” e “estado catatônico” demonstrando fraca relação com o diagnóstico de *delirium*.

Quanto às limitações, os dados deste estudo foram obtidos em um único hospital, por

um curto período de tempo e a partir de registros de enfermagem disponíveis, apenas, em japonês, com resultados que podem não ser generalizáveis e que demandam validação externa.

Em casos de *delirium*, além de depender do observador e da frequência de avaliações, as metodologias diagnósticas convencionais baseiam-se, essencialmente, na observação de um distúrbio comportamental ou na incapacidade de o paciente responder satisfatoriamente a uma sequência de perguntas. O estudo de Young *et al.* (2021) demonstrou que técnicas de PLN podem ser uma opção a estas limitações, uma vez que, a partir da identificação de palavras em notas de progresso, são capazes de definir lógica, clínica e epidemiologicamente uma população de pacientes com alto risco de desenvolver *delirium* e que, por vezes, pode ser subdiagnosticada por métodos convencionais.

Young *et al.* (2021) utilizou técnicas de PLN para analisar registros médicos, de enfermagem e de saúde aliada e identificou mais de 5.000 pacientes com PLN-Dx-BD, promovendo uma avaliação em larga escala da epidemiologia dessa condição. Em contraponto, este estudo resguardou-se a um centro único, com resultados que podem não ser aplicáveis em outras UTIs. Ademais, tendo em vista que a equipe de cabeceira documenta mais prontamente estados agitados, esta técnica pode não detectar distúrbios comportamentais não agitados.

Young *et al.* (2022) comparou a prevalência de distúrbio comportamental diagnosticado por PLN e a positividade no CAM-UTI em pacientes assistidos em unidades de cuidado intensivo, com o PLN-Dx-BD identificando mais pacientes com probabilidade de utilizar medicação antipsicótica. Os desfechos deste estudo corroboraram descobertas anteriores de que, na prática clínica, o uso de CAM-UTI pode retardar a detecção de *delirium*. Além disso, o PLN-Dx-BD apresenta a vantagem de registrar características de triagem comportamental durante um período estendido, configurando-se como uma alternativa ao

CAM-UTI, que captura dados instantâneos.

O estudo de Young *et al.* (2022) foi conduzido em uma amostra expressiva e heterogênea de pacientes em cuidados intensivos, com validação externa aceitável. Porém, a pesquisa ambientou-se em uma Unidade de Terapia Intensiva terciária de um hospital universitário afiliado de um país desenvolvido, de modo que seus resultados podem não se aplicar em UTIs de países de média ou baixa renda.

A literatura acerca da utilização do Processamento de Linguagem Natural no diagnóstico de *delirium* ainda é escassa, especialmente quando considerados pacientes idosos assistidos em unidades de cuidado intensivo. Dessa maneira, os resultados aqui descritos devem ser confirmados ou refutados em estudos subsequentes, que explorem o potencial diagnóstico de técnicas de PLN.

CONCLUSÃO

As metodologias diagnósticas convencionais dependem do observador e da frequência de avaliações, corroborando o subdiagnóstico de *delirium*. A análise de registros eletrônicos de saúde promove intervenções mais rápidas e eficazes, com detecção precoce e mais precisa do *delirium*.

Dessa forma, a utilização de PLN tem se mostrado uma ferramenta útil e promissora para os profissionais da saúde no apoio ao diagnóstico de *delirium* em pessoas idosas no cenário de UTI, mas seu uso requer abordagens interdisciplinares, validação clínica e investimento em infraestrutura tecnológica. Espera-se que, com o avanço das pesquisas e o desenvolvimento de soluções específicas para o ambiente hospitalar, essa tecnologia possa ser incorporada à prática médica de maneira ética e eficaz, contribuindo para o aprimoramento do cuidado a pacientes idosos críticos, a partir de intervenções mais eficientes, e minimizando, assim, o risco de complicações e os custos para a saúde.

REFERÊNCIAS BIBLIOGRÁFICAS

- Alves, D. da S. B., Marques, D. M., & Vernaglia, T. V. C. (2024). Fatores predisponentes e precipitantes para o desenvolvimento de delirium pós-operatório em pacientes críticos de uma unidade intensiva universitária. **Revista Latino-Americana de Enfermagem**, 32. <https://doi.org/10.1590/1518-8345.7113.4234>
- Berge, G. T., Granmo, O.-C., Tveit, T. O., Ruthjersen, A. L., & Sharma, J. (2023). Combining unsupervised, supervised and rule-based learning: the case of detecting patient allergies in electronic health records. **BMC medical informatics and decision making**, 23 (1), 188. <https://doi.org/10.1186/s12911-023-02271-8>
- Bispo, M. M., & Souza, R. C. da S. (2024). Adherence to optimal delirium management practices in intensive care units in Brazil: a nationwide survey. **JBIG Implementation**. <https://doi.org/10.1097/XEB.0000000000000470>
- Carvalho, L. A. C. *et al.* (2022). Accuracy of delirium risk factors in adult intensive care unit patients. **Revista da Escola de Enfermagem da USP**, 56, e20210222. <https://doi.org/10.1590/1980-220X-REEUSP-2021-0222>
- Castro, V. M., Sacks, C. A., Perlis, R. H., & McCoy, T. H. (2021). Development and External Validation of a Delirium Prediction Model for Hospitalized Patients With Coronavirus Disease 2019. **Journal of the Academy of Consultation-Liaison Psychiatry**, 62 (3), 298-308. <https://doi.org/10.1016/j.jaclp.2020.12.005>
- Cechinel, C. *et al.* (2022). Frailty and delirium in hospitalized older adults: A systematic review with meta-analysis. **Revista Latino-Americana de Enfermagem**, 30, e3687. <https://doi.org/10.1590/1518-8345.6120.3687>
- Chen, L. *et al.* (2022). A novel semiautomatic Chinese keywords instrument screening delirium based on electronic medical records. **BMC Geriatrics**, 22(1), 779. <https://doi.org/10.1186/s12877-022-03474-w>
- Disponível em: <https://www.gov.br/mds/pt-br/noticias-e-conteudos/desenvolvimento-social/noticias-desenvolvimento-social/mds-lanca-diagnostico-sobre-envelhecimento-e-direito-ao-cuidado/Nota_Informativa_N_5.pdf>. Acesso em: 13 mar. 2025
- Fu, S. *et al.* (2022). Ascertainment of Delirium Status Using Natural Language Processing From Electronic Health Records. **The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences**, 77(3), 524-530. <https://doi.org/10.1093/gerona/glaa275>
- Hart, K. L. *et al.* (2021). Distribution of agitation and related symptoms among hospitalized patients using a scalable natural language processing method. **General Hospital Psychiatry**, 68, 46-51. <https://doi.org/10.1016/j.genhosppsych.2020.11.003>
- Hayhurst, C. J., Hughes, C. G., & Pandharipande, P. P. (2016). Intensive Care Unit Delirium: A Review of Diagnosis, Prevention, and Treatment. **Anesthesiology**, 125(6), 1229-1241. <https://doi.org/10.1097/ALN.0000000000001378>

Homburg, M. *et al.* (2023). A Natural Language Processing Model for COVID-19 Detection Based on Dutch General Practice Electronic Health Records by Using Bidirectional Encoder Representations From Transformers: Development and Validation Study. **Journal of Medical Internet Research**, 25, e49944. <https://doi.org/10.2196/49944>

Jayanama, K., Kaewpongsa, P., & Ruangritchankul, S. (2025). Risk factors and outcomes of hyperactive delirium in older medical in patients admitted to non-intensive care unit: a prospective cohort study. **BMC Psychiatry**, 25(1), 330. <https://doi.org/10.1186/s12888-025-06731-5>

Koleck, T. A., Dreisbach, C., Bourne, P. E., & Bakken, S. (2019). Natural language processing of symptoms documented in free-text narratives of electronic health records: a systematic review. **Journal of the American Medical Informatics Association: JAMIA**, 26(4), 364-379. <https://doi.org/10.1093/jamia/ocy173>

Le Glaz, A. *et al.* (2021). Machine Learning and Natural Language Processing in Mental Health: Systematic Review. **Journal of Medical Internet Research**, 23(5), e15708. <https://doi.org/10.2196/15708>

Li, M. D. *et al.* (2021). Analysis of Stroke Detection during the COVID-19 Pandemic Using Natural Language Processing of Radiology Reports. **American Journal of Neuroradiology**, 42(3), 429-434. <https://doi.org/10.3174/ajnr.A6961>

Lima, B. R. de, Nunes, B. K. G., Guimarães, L. C. da C., Almeida, L. F. de, & Pagotto, V. (2021). Incidence of delirium following hospitalization of elderly people with fractures: risk factors and mortality. **Revista da Escola de Enfermagem da USP**, 55, e20200467. <https://doi.org/10.1590/1980-220X-REEUSP-2020-0467>

Manual diagnóstico e estatístico de transtornos mentais [recurso eletrônico]: DSM-5 / [American Psychiatric Association] – 5. ed. – Porto Alegre: Artmed, 2014.

Miranda, F. *et al.* (2023). Confusion Assessment Method for the Intensive Care Unit (CAM-ICU) for the diagnosis of delirium in adults in critical care settings. **Cochrane Database of Systematic Reviews**, 11(11), CD013126. <https://doi.org/10.1002/14651858.CD013126.pub2>

Miyazawa, Y. *et al.* (2024). Identification of risk factors for the onset of delirium associated with COVID-19 by mining nursing records. **PloS One**, 19(1), e0296760. <https://doi.org/10.1371/journal.pone.0296760>

Paulino, M. C., Pova, P., & Serafim, R. B. (2020). What every intensivist needs to know about subsyndromal delirium in the intensive care unit. **Revista Brasileira de Terapia Intensiva**, 32(1), 14-16. <https://doi.org/10.5935/0103-507x.20200004>

Rodrigues, J. A. M. *et al.* (2023). Internação hospitalar e a ocorrência de delirium em idosos na condição de fragilidade física: estudo transversal. **Revista da Escola de Enfermagem da USP**, 57. <https://doi.org/10.1590/1980-220x-reeusp-2023-0156pt>

Young, M. *et al.* (2022). Natural language processing diagnosed behavioral disturbance vs confusion assessment method for the intensive care unit: prevalence, patient characteristics, overlap, and association with treatment and outcome. **Intensive Care Medicine**, 48 (5), 559-

569. <https://doi.org/10.1007/s00134-022-06650-z>

Young, M. *et al.* (2021). Natural language processing to assess the epidemiology of delirium-suggestive behavioural disturbances in critically ill patients. **Critical Care and Resuscitation: Journal of the Australasian Academy of Critical Care Medicine**, 23(2), 144-153. <https://doi.org/10.51893/2021.2.aa1>