META-ANÁLISE DE REVISÕES SISTEMÁTICAS DE LITERATURA DE MÉTODOS DE IMPUTAÇÃO DE VALORES AUSENTES

Julyanna Silva Cacau[1]

Fábio Manoel França Lobato[2]

Antonio Fernando Lavareda Jacob Junior[3]

INTRODUÇÃO

A quantidade e variedade de dados disponíveis extrapolaram a capacidade de análise manual destes. Com o avanço tecnológico, surgiram diversas técnicas para armazenar, processar e analisar esses dados (Badshar *et al*, 2024). No contexto de mineração de dados, os algoritmos atuam na extração e análise de padrões desses dados, sendo essenciais para a formação de conhecimento e para suporte à tomada de decisão tanto no meio empresarial quanto no meio acadêmico. (Macucule *et al*, 2024).

No entanto, o processo de aquisição de dados pode apresentar falhas, tais como: erros humanos, falhas de sensores, falta de respostas em formulários e desistência de participantes em estudos. (Lijuan Ren *et al*, 2023). Como resultado, bases de dados possuem campos com valores ausentes (VAs), e essa ausência impacta diretamente a análise dos algoritmos, resultando na possível perda de informações importantes. Devido a isso, surgiu a necessidade de investigar técnicas para preencher os valores ausentes.

De acordo com (Adnan *et al*, 2022), nas últimas três décadas houve um grande aumento nos estudos sobre valores ausentes e, como consequência, a diversificação de técnicas de imputação documentadas na literatura. Com uma quantidade considerável de estudos sobre os VAs, surge a necessidade de sintetizar os resultados desses estudos e avaliar o rigor metodológico das produções acadêmicas por meio de revisões sistemáticas de literatura (RSL) (Kolaski *et al*, 2023). Quando se tem uma quantidade considerável de RSL sobre um tema, utilizar um método para agrupá-las e analisá-las quantitativamente pode fornecer uma visão mais aprofundada. Sendo assim, as meta-análises podem auxiliar no melhor entendimento sobre um tópico ou tópico relacionado e podem fornecer listas de estudos que oferecem um bom ponto de partida para identificar e se familiarizar com a literatura relevante (Hansen *et al*, 2021).

Nesse contexto, o presente trabalho propõe a realização de uma meta-análise das RSLs sobre métodos de imputação de valores ausentes, publicadas entre 2015 e 2025, com o objetivo de identificar as abordagens de maior desempenho e apontar lacunas metodológicas no campo.

FUNDAMENTAÇÃO TEÓRICA

VALORES AUSENTES E MÉTODOS DE IMPUTAÇÃO

Little e Rubin (2019) defendem que entender o tipo de ausência de um dado é essencial para tratá-lo corretamente. Dentre os mecanismos de ausência mais conhecidos,

1 Discente em Engenharia da Computação, Universidade Estadual do Maranhão. E-mail: julyanna7899@gmail.com
2 Doutor em Engenharia Elétrica, docente do Instituto de Engenharia e Geociências da Universidade Federal do Oeste do Pará (UFOPA). E-mail: fabio.lobato@ufopa.edu.br
3 Doutor em Engenharia Elétrica, docente do Departamento de Engenharia da Computação da Universidade Estadual do Maranhão. E-mail: antoniojunior@professor.uema.br

estão: MCAR (*Missing Completely at Random*), MAR (*Missing at Random*) e MNAR (*Missing Not at Random*).

No MCAR, a ausência de dados é completamente aleatória e não apresenta nenhum padrão sistemático nem associação com os dados observados ou com quaisquer outras variáveis no conjunto de dados. No MAR, a ausência de dados pode ser prevista com base em outras variáveis do estudo, mas não diretamente a partir dos próprios dados faltantes. No caso do MNAR, os dados faltantes estão diretamente ligados ao valor da própria observação ausente (Zhou *et al*, 2024).

Para resolver esse problema, foram desenvolvidas técnicas para substituir os valores ausentes, de modo que se aproximassem dos valores disponíveis. Existem inúmeros métodos de imputação para lidar com os dados ausentes, como a remoção de instâncias incompletas ou a substituição por médias e medianas (Rubin, 1976), até técnicas mais sofisticadas baseadas em aprendizado de máquina, como *k-Nearest Neighbors Imputation* (KNN) (Troyanskaya *et al*, 2001), *Multiple Imputation by Chained Equations* (MICE) (Buuren & Groothuis-Oudshoorn, 2011) e imputação baseada em redes neurais (Gondara & Wang, 2018).

REVISÃO SISTEMÁTICA DE LITERATURA

A Revisão Sistemática de Literatura (RSL) é um método que visa avaliar produções acadêmicas por meio de protocolos bem definidos, a fim de garantir a reprodutibilidade e a imparcialidade dos resultados das pesquisas (Oliveira *et al*, 2023).

Ela é caracterizada pelo processo rigoroso, sistemático, explícito e reprodutível dos seus métodos, tendo como objetivo identificar se todos os estudos sobre um tópico seguem critérios previamente estabelecidos, como estratégia de pesquisa, critérios de pesquisa, métodos para extração e análise de dados (Sarmento *et al*, 2024).

No campo da imputação de valores ausentes, as revisões sistemáticas publicadas nas últimas décadas permitem uma análise detalhada e comparação das técnicas existentes. De acordo com Ismail *et al.* (2022), ela permite uma visão abrangente do desempenho dos algoritmos, da natureza dos dados, da identificação de pontos de melhoria em cada método e de novas tendências na área.

META-ANÁLISE

A meta-análise sintetiza os resultados de estudos científicos por meio de técnicas estatísticas, como a análise de variância entre os estudos primários. Dessa forma, fornece uma síntese quantitativa de um campo de pesquisa (Hansen *et al*, 2021).

Aplicando a meta-análise, é possível descobrir novas informações com base nas conclusões de estudos anteriores, comparar os resultados e definir qual deles apresenta maior precisão por meio de análises estatísticas (Chueke; Amatucci, 2022).

Existem diferentes métodos para realizar uma meta-análise. Destacam-se nas pesquisas em negócios e gestão: meta-análise univariada, meta-regressão, meta-análise por meio de modelagem de equações estruturais e meta-análise qualitativa (Hansen *et al.*, 2021).

Para garantir o rigor da análise, os autores dividem o processo em várias etapas. (Hansen *et al*, 2021) propõem um roteiro de oito passos para conduzir meta-análises de forma rigorosa, sistematizado na Tabela 1.

Tabela 1 – Passos para conduzir uma meta-análise

Passo	Descrição
1. Definir a pergunta de pesquisa	Estabelecer uma questão clara que guiará a meta-análise, considerando a relevância e a viabilidade da pesquisa.
2. Busca da literatura	Realizar uma busca sistemática e transparente em bases de dados relevantes para identificar estudos primários.
3. Escolha da medida de efeito	Determinar a métrica de efeito (ex.: coeficiente de correlação, diferença de médias padronizada) e converter medidas diferentes, se necessário.
4. Escolha do método analítico	Selecionar a abordagem estatística apropriada, como meta-análise univariada, meta-regressão ou meta-análise de modelagem estrutural.
5. Escolha do software	Utilizar programas estatísticos como R (metafor), Stata, SPSS, SAS ou softwares específicos para meta-análise.
6. Codificação dos tamanhos de efeito	Criar uma planilha de codificação para registrar os tamanhos de efeito, as características dos estudos e as possíveis variáveis moderadoras.
7. Análise dos dados	Verificar *outliers*, testar o viés de publicação e escolher entre modelos de efeito fixo e aleatório para a análise principal.
8. Relato dos resultados	Apresentar os achados de forma clara, incluindo tabelas, gráficos e uma discussão sobre as limitações e as implicações da meta-análise.

Fonte: Adaptado de Hansen *et al*, (2021)

METODOLOGIA

Considerando a natureza deste plano de trabalho, optou-se pelo uso do modelo *Cross-Industry Standard Process for Data Mining* (CRISP-DM) para o desenvolvimento das atividades. Ele é um modelo metodológico que oferece uma abordagem estruturada e cíclica para organizar e executar projetos de mineração de dados.

Sua estrutura contempla as fases e suas respectivas tarefas e saídas, formando um ciclo do projeto (Wirth, 2000). Dentro desse *framework*, as atividades foram organizadas em consonância com as oito etapas para meta-análise propostas por (Hansen *et al*, 2021), sumarizadas na Tabela 1.

Entendimento do Negócio e Formulação das Questões de Pesquisa

A pesquisa teve início na etapa de entendimento do negócio, conforme pré-estabelecido pelo modelo CRISP-DM. Nesta fase, foram analisados os objetivos e as necessidades do estudo, garantindo uma base sólida para a condução da investigação. Além disso, foi realizado um levantamento bibliográfico sistemático, com o objetivo de identificar os principais conceitos, metodologias e evidências empíricas relevantes para o estudo. Essa revisão da literatura permitiu direcionar o projeto conforme a Tabela 1, que orientou as demais etapas do estudo.

Dada a investigação, identificou-se uma lacuna na literatura quanto ao rigor do processo metodológico das revisões sistemáticas sobre métodos de imputação de valores ausentes. Dessa forma, definiu-se como objetivo central sintetizar os resultados das revisões sistemáticas de literatura sobre métodos de imputação de valores ausentes publicadas entre 2015 e 2025 por meio da meta-análise para analisar os estudos por meio de abordagem quantitativa, identificando os métodos de maior desempenho e apontando lacunas metodológicas nas pesquisas.

Para orientar a coleta e a análise dos dados, foi estruturado o *framework* PICOC (*Population, Intervention, Comparison, Outcome, Context*), conforme apresentado na Tabela 2, o qual orienta a formulação das questões de pesquisa.

Tabela 2 – Estrutura do PICOC

População	Intervenção	Comparação	Resultados	Contexto
pesquisadores, cientistas de dados, estatísticos, analistas de dados, engenheiros de software	Revisões sistemáticas de literatura, meta-análise, imputação de valores ausentes, métodos estatísticos, aprendizado de máquina, imputação múltipla, redes neurais, modelos probabilísticos	Métodos tradicionais, métodos modernos, algoritmos bioinspirados	Eficiência, precisão, redução de erros, eficiência computacional, robustez dos métodos	multirrótulo

Fonte: Elaborado pela autora (2025)

Por meio do PICOC, foi possível formular as questões de pesquisa apresentadas na Tabela 3 e realizar uma revisão ampliada dos métodos de imputação de valores ausentes para a identificação e comparação de temas emergentes na área.

Tabela 3 – Perguntas de pesquisa

Perguntas de pesquisa
• Quais são os estudos sobre métodos de imputação de valores ausentes publicados no período de 2015-2025?
• Quais são as características dos conjuntos de dados e tipos de análise usados nos estudos primários?
• Quais métodos obtiveram os melhores resultados na imputação de valores ausentes?
• A imputação por classificação multirrótulo com algoritmos bioinspirados apresenta resultados semelhantes aos dos estudos pesquisados?

Fonte: Elaborado pela autora (2025)

Busca da Literatura

A busca foi realizada com o apoio da ferramenta Parsifal (https://parsif.al/), uma plataforma que apoia pesquisadores no planejamento e na execução de revisões sistemáticas. Para realizar a consulta, inicialmente definiram-se palavras-chave para um mapeamento geral de pesquisas sobre revisões de valores ausentes nas seguintes bases de dados: *Google Scholar, IEEE Digital Library, ACM Digital Library, Scopus e Web of Science*. A seleção das *queries* foi realizada com base nos resultados que fornecem o maior número de estudos sobre revisões sistemáticas da literatura da abordagem desta pesquisa.

A *query* de busca adotada foi: *("imputation of missing values" OR "data imputation" OR "handling missing values") AND ("meta analysis" OR "quantitative synthesis" OR "statistical synthesis" OR "systematic literature review" OR "literature review" OR "research synthesis" OR "systematic review").* A seleção das *queries* considerou os resultados que forneciam o maior número de estudos sobre RSLs alinhados ao escopo desta pesquisa.

Seleção dos Estudos – Critérios de Inclusão e Exclusão

Com base nas questões de pesquisa, foram definidos critérios de inclusão e exclusão dos estudos, conforme descritos na Tabela 4. Além da exclusão por título e resumo, são necessárias a remoção de duplicatas e a triagem final após a leitura do texto completo (Pigott e Polanin, 2020).

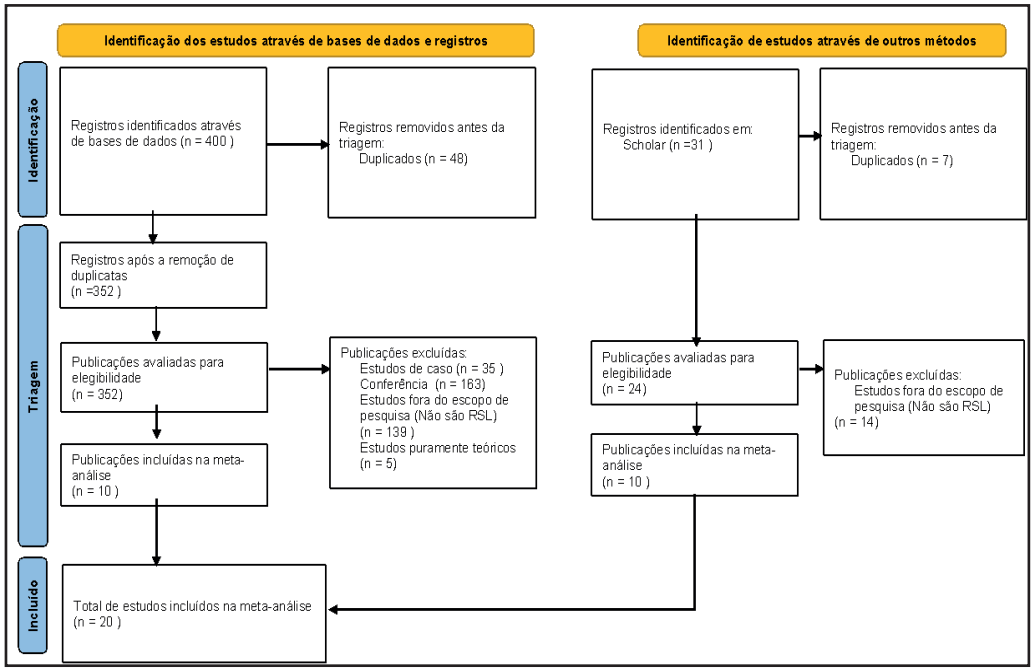
Tabela 4 – Critérios para inclusão e exclusão dos estudos

Inclusão	Exclusão
Artigos de 2015-2025	Estudos de caso, conferências.
Revisões sistemáticas de literatura e meta-análise dos métodos de imputação de valores ausentes	Artigos de opinião, editoriais e comentários sem dados empíricos.

Fonte: Elaborado pela autora (2025)

Após a filtragem, também foram selecionados manualmente estudos relevantes à pesquisa que não foram encontrados na etapa de busca. Todo esse processo foi documentado sistematicamente em cada etapa da seleção, conforme o fluxograma PRISMA apresentado na Figura 1.

Figura 1 – Diagrama de fluxo PRISMA.



Fonte: Elaborado pela autora (2025)

Análise de Qualidade – AMSTAR-2

A qualidade metodológica dos estudos selecionados foi avaliada por meio do instrumento AMSTAR-2, um sistema de classificação que avalia o nível de qualidade de todas as revisões, categorizando-as como muito baixa, baixa, média ou alta. Essa avaliação é categorizada da seguinte forma: i) alta: sem falhas ou com apenas uma falha não crítica: a revisão sistemática oferece um resumo claro e completo dos resultados; ii) média: mais de uma falha não crítica: a revisão sistemática oferece um resumo adequado dos resultados; iii) baixa: uma falha crítica,

ou com ou sem falhas não críticas: a revisão sistemática pode não fornecer um resumo claro e completo dos resultados; iv) criticamente baixa: com falhas críticas, ou com ou sem falhas não críticas: a revisão não pode ser considerada para fornecer um resumo claro e completo dos resultados.

Condução da Meta-análise

Antes da meta-análise, foi realizada uma avaliação da heterogeneidade dos estudos, que identificou grande variabilidade nas características dos dados (tamanho da amostra, métricas de avaliação, taxa de dados ausentes). Diante disso, optou-se pelo modelo de meta-análise de efeitos aleatórios, devido à grande variedade entre os estudos, como diferenças nas populações, métodos ou outros fatores (Deeks *et al*, 2023).

Para a análise estatística, foram selecionados os estudos que apresentaram dados quantitativos comparáveis. Foi utilizada a métrica de acurácia como medida de efeito, por ser a única comum aos dados dos artigos quantitativos disponíveis. Os dados foram estruturados com os seguintes elementos: identificador do estudo, autores, acurácia reportada e erro padrão (estimado com base no desvio padrão geral dos estudos). Com base nessa estrutura, foram calculados o efeito combinado (média ponderada das acurácias) e o intervalo de confiança de 95%.

RESULTADOS E DISCUSSÃO

Resultados da Busca e Seleção dos Estudos

A busca nas bases de dados, conforme as queries de busca, encontrou um total de 400 estudos sobre revisões sistemáticas de métodos de valores ausentes, resumidos na Tabela 5.

Tabela 5 – Resultado da pesquisa de queries nas bases de dados.

Base de dados	Resultado das buscas
ACM Digital Library	280 resultados
IEEE Digital Library	9 resultados
Web of Science	46 resultados
Scopus	65 resultados

Fonte: Elaborado pela autora (2025)

Esses resultados demonstraram que há um grande número de revisões sistemáticas no campo da imputação de valores ausentes. Dessa forma, a meta-análise consolidou-se como essencial nesse estudo, pois analisa, de forma estatística, grandes amostras, diferentemente da RSL, que é mais eficaz em amostras pequenas (Chueke; Amatucci, 2022).

Após a remoção de duplicatas e a aplicação dos critérios de inclusão e exclusão, conforme o protocolo PRISMA, foram selecionados 20 estudos para análise, listados na Tabela 6.

Tabela 6 – Lista de estudos incluídos

A Systematic Literature Review on Missing Values: Research Trends, Datasets, Methods and Frameworks	Setiawan *et al*, 2023
A systematic review of generative adversarial imputation network in missing data imputation	Zhang *et al*, 2023
A systematic review of machine learning-based missing value imputation techniques	Thomas *et al*, 2021
A Systematic Review of Recurrent Neural Network Adoption in Missing Data Imputation	Akbar *et al*, 2025
Applications of Missing Data Imputation Methods in Wastewater Treatment Plants; A Systematic literature Review using the Kitchenham Method	Chaoui *et al*, 2023
Computational Methods for Data Integration and Imputation of Missing Values in Omics Datasets	Schumann *et al*, 2025
Handling missing covariate data in clinical studies in haematology	Bonneville *et al*, 2023
Handling missing data in health economics and outcomes research (HEOR): a systematic review and practical recommendations	Kumar *et al*, 2023
Handling missing values in healthcare data: A systematic review of deep learning-based imputation techniques	Liu *et al*, 2023
Handling of missing outcome data in traumatic brain injury research: a systematic review	Sophie *et al,* 2019
Identify the most appropriate imputation method for handling missing values in clinical structured datasets: a systematic review	Afkanpou *et al*, 2024
Missing data handling in non-inferiority and equivalence trials: A systematic review	Brooke *et al*, 2018
Missing data in bioarchaeology I: A review of the literature	Wissler *et al*, 2022
Missing data is poorly handled and reported in prediction model studies using machine learning: a literature review	SWJ Nijman *et al*, 2022
Missing Value Imputation Designs and Methods of Nature-Inspired Metaheuristic Techniques: A Systematic Review	Chiu *et al*, 2022
Providing a comprehensive understanding of missing data imputation processes in evapotranspiration-related research: a systematic literature review	Basakin *et al*, 2023
Reporting and handling missing outcome data in mental health: a systematic review of Cochrane systematic reviews and meta-analyses	Loukia *et al*, 2015
Reporting and handling of missing data in predictive research for prevalent undiagnosed type 2 diabetes mellitus: a systematic review	Katya *et al*, 2015
Systematic Review of Using Machine Learning in Imputing Missing Values	Alabadla *et al*, 2022
The untold story of missing data in disaster research: a systematic review of the empirical literature utilising the Emergency Events Database (EM-DAT)	Jones *et al*, 2023

Fonte: Elaborado pela autora (2025)

Análise de Qualidade dos Estudos

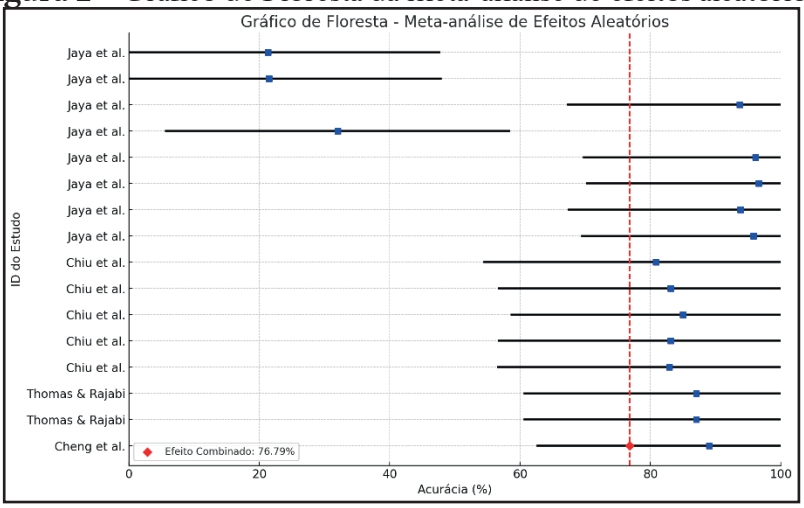
A avaliação pelo AMSTAR-2, resumida na Figura 3, indicou que todos os 20 artigos selecionados obtiveram pontuação baixa ou criticamente baixa. Isso se deve principalmente à falta da análise do Risco de Viés (RoB) de cada estudo, pois apenas 20% deles relataram e discutiram a problemática. Tal erro tem impacto significativo em toda a pesquisa, reduzindo a confiabilidade dos resultados, a validade das conclusões e a precisão das análises estatísticas. A sua presença pode levar a conclusões tendenciosas em diversas áreas, o que torna essencial o seu relato.

Também foi observado, na leitura completa dos artigos, que a maioria realizou apenas análises qualitativas, sem comparar quantitativamente o desempenho dos métodos. Sendo assim, essa pesquisa contém artigos com um RoB alto e com poucos artigos com valores quantitativos do desempenho dos métodos de imputação, e isso implicou resultados com grande risco de estarem enviesados.

Resultados da Meta-análise

Para a análise estatística, foram selecionados 4 artigos e adotou-se a métrica de acurácia, por ser comum a todos os artigos selecionados. Esses dados foram estruturados em uma tabela com os seguintes elementos: id do estudo: identificador único para cada estudo; autor(es): o nome dos autores de cada estudo; acurácia: a métrica de desempenho usada; erro padrão: a medida de variabilidade da acurácia, que foi estimada com base no desvio padrão geral dos estudos. Com base nesses dados, foram calculados o efeito combinado, que é a média ponderada das acurácias de cada método, e o intervalo de confiança.

Figura 2 – Gráfico de Floresta da meta-análise de efeitos aleatórios.



Fonte: Elaborado pela autora (2025).

A meta-análise de efeitos aleatórios mostrou uma média ponderada de acurácia de **76,79%**, com um intervalo de confiança de **95%** entre **63,89%** e **89,69%**. A heterogeneidade foi significativa, indicando que as diferenças nos resultados dos estudos podem ser atribuídas a fatores além do erro amostral. Conforme a figura 2, é possível perceber uma grande heterogeneidade dos resultados e também um elevado RoB dos dados. Isso implica a necessidade de uma análise mais robusta dos dados e de uma seleção de outros tipos de artigos.

CONSIDERAÇÕES FINAIS

A definição do protocolo de pesquisa orientado pelas oito etapas de (Hansen *et al,* 2021) e pelo fluxograma PRISMA permitiu que o estudo fosse padronizado, transparente e reprodutível. Dessa forma, foi possível descrever claramente todas as etapas do processo de realização da meta-análise.

Como resultado, por meio da análise de qualidade do AMSTAR-2, foi descoberta uma falha crítica nos artigos sobre revisões sistemáticas de métodos de imputação de valores ausentes: todos apresentaram uma qualidade baixa ou criticamente baixa. Isso ocorreu porque a maioria não realizou a avaliação do risco de viés e também não apresentou resultados quantitativos nas comparações entre os métodos, limitando-se a análises qualitativas.

A meta-análise de efeitos aleatórios evidenciou a necessidade de padronização das métricas de avaliação, das populações estudadas e das condições experimentais nos estudos sobre imputação de VAs, para reduzir a heterogeneidade e garantir maior consistência aos resultados.

Portanto, as revisões sistemáticas de literatura sobre métodos de imputação de valores ausentes precisam aprimorar seu rigor metodológico, especialmente no que se refere à análise quantitativa do desempenho dos métodos e à avaliação sistemática do risco de viés. Este trabalho

contribui para esse debate ao mapear as lacunas existentes e ao proporcionar, por meio da meta-análise, uma visão quantitativa do estado da arte da área.

AGRADECIMENTOS E USO DE IA GENERATIVA

Este trabalho foi apoiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), por meio da bolsa DT-303031/2023-9 e do PIBIC-134022/2024-6.

Declara-se que o modelo de IA generativa GPT-4 foi utilizado como ferramenta de apoio para a revisão gramatical e de conformidade com a ABNT, a tradução de artigos em inglês e a estruturação do modelo de artigo. A autoria e a responsabilidade integral pelo conteúdo final, incluindo a verificação de plágio e as correções, são de Julyanna Silva Cacau.

Figura 3 - Avaliação de qualidade dos artigos utilizando o AMSTAR-2

AMSTAR Questions	Jones et al., 2023	Alabadik et al.,2022	Katya et al., 2015	Loukia et al., 2015	Basavin et al.,2023	Chan et al., 2022	SWJ Nijman et al.,2022	Wissler et al., 2022	Brooke et al., 2018	Affonso et al., 2024	Sophie et al., 2019	Liu et al.,2023	Mamor et al., 2023	Bonnevi lle et al., 2023	Schuman n et al., 2025	Chaoui et al., 2023	Akbor et al.,2025	Thomas et al., 2021	Zhang et al.,2023	Setiava n et al., 2023
Did the research questions and inclusion criteria for the review include the components of PICO?	PY	PY	PY	Y	PY	Y	Y	Y	Y	Y	PY	PY	PY	N	N	PY	Y	Y	PY	Y
Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations	Y	Y	N	PY	PY	PY	PY	PY	PY	PY	Y	PY	PY	PY	PY	PY	PY	PY	PY	PY
Did the review authors explain their selection of the study designs for inclusion in t review?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Was a comprehensive literature search performed?	Y	Y	Y	PY	PY	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	PY	Y	Y	Y
Did the review authors perform data extraction in duplicate ?	Y	N	Y	N	N	Y	Y	N	Y	Y	Y	Y	N	N	N	N	N	N	N	N
Did the review authors perform study selection in duplicate ?	Y	Y	N	Y	N	Y	N	N	Y	Y	Y	Y	Y	N	N	N	N	N	N	N
Did the review authors provide a list of excluded studies and justify the exclusions	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
Did the review authors describe the included studies in adequate detail?	Y	Y	Y	Y	Y	Y	Y	PY	Y	Y	Y	Y	Y	Y	PY	PY	PY	PY	Y	PY
Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?	N	N	N	N	N	Y	Y	N	N	Y	Y	N	N	N	N	N	N	N	N	N
Did the review authors report on the sources of funding for the studies included in t review?	N	N	N	N	N	N	N	N	Y	N	N	N	N	N	N	N	N	N	N	N
If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
If meta-analysis was performed, did the review authors assess the potential impact RoB in individual studies on the results of the meta-analysis or other evidence synthesis?	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review?	N	N	N	N	N	N	Y	N	N	N	N	Não	N	N	N	N	N	N	N/A	N
Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?	N	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y
If they performed quantitative synthesis did the review authors carry out an adequat investigation of publication bias (small study bias) and discuss its likely impact on the results of	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?	Y	Y	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N
Resultado	Criticamente Baixa	Criticamente Baixa	Criticamente Baixa	Criticamente baixa	Criticamente Baixa	Criticamente Baixa	Baixa	Criticamente Baixa	Criticamente Baixa	Criticamente Baixa	Criticamente Baixa	Criticamente Baixa	Criticamente Baixa	Criticamente Baixa	Criticamente baixa	Criticamente baixa	Criticamente Baixa	Criticamente Baixa	Criticamente Baixa	Criticamente Baixa

Fonte: Elaborado pela autora (2025).

REFERÊNCIAS

ADNAN, F.A.; JAMALUDIN, K.R.; WAN MUHAMAD, W.Z.A. *et al.* **A review of the current publication trends on missing data imputation over three decades: direction and future research**. Neural Comput & Applic 34, 18325–18340, 2022. https://doi.org/10.1007/s00521-022-07702-7

AFKANPOUR, Marziyeh; HOSSEINZADEH, Elham; TABESH, Hamed. **Identify the most appropriate imputation method for handling missing values in clinical structured datasets: a systematic review**. BMC Medical Research Methodology, v. 24, n. 1, p. 188, 2024. DOI: 10.1186/s12874-024-02310-6.

AKBAR, Nur Aqilah Fadzil *et al.* **A Systematic Review of Recurrent Neural Network Adoption in Missing Data Imputation**. International Journal of Computing and Digital Systems, 2025. DOI: disponível em: https://iiict.uob.edu.bh/IJCDS.

ALABADLA, Mustafa *et al.* **Systematic Review of Using Machine Learning in Imputing Missing Values**. IEEE Access, v. 10, p. 44483–44502, 2022. DOI: 10.1109/ACCESS.2022.3160841.

BADSHAH, A.; DAUD, A.; ALHARBEY, R. *et al.* **Aplicações de Big Data: visão geral, desafios e futuro**. Artificial Intelligence Review, v. 57, 290, 2024. https://doi.org/10.1007/s10462-024-10938-5.

BAŞAKIN, Emre Egemen *et al.* **Providing a comprehensive understanding of missing data imputation processes in evapotranspiration-related research: a systematic literature review**. Hydrological Sciences Journal, v. 68, n. 14, p. 2089–2104, 2023. DOI: 10.1080/02626667.2023.2249456.

BONNEVILLE, Edouard F. *et al.* **Handling missing covariate data in clinical studies in haematology**. Best Practice & Research Clinical Haematology, v. 36, n. 2, p. 101477, 2023. DOI: 10.1016/j.beha.2023.101477.

CHAOUI, Abdellah *et al.* **Applications of Missing Data Imputation Methods in Wastewater Treatment Plants: a systematic literature review using the Kitchenham method**. International Journal of Advanced Computer Science and Applications, v. 14, n. 10, 2023. DOI: 10.14569/IJACSA.2023.0141049.

CHIU, Po Chan *et al.* **Missing Value Imputation Designs and Methods of Nature-Inspired Metaheuristic Techniques: a systematic review**. IEEE Access, v. 10, p. 61544–61566, 2022. DOI: 10.1109/ACCESS.2022.3172319.

CHUEKE, G. V.; AMATUCCI, M. **Métodos de sistematização de literatura em estudos científicos: bibliometria, meta-análise e revisão sistemática**. Internext, v. 17, n. 2, p. 284–292, 2022. https://doi.org/10.18568/internext.v17i2.704.

DEEKS, J. J.; HIGGINS, J. P. T.; ALTMAN, D. G. (eds.). **Chapter 10: Analysing data and undertaking meta-analyses**. In: HIGGINS, J. P. T. *et al.* (eds.). Cochrane Handbook for Systematic Reviews of Interventions version 6.4. Cochrane, 2023.

GONDARA, L.; WANG, K. MIDA: **Multiple imputation using denoising autoencoders**. In: Proceedings of the 2018 IEEE International Conference on Data Science and Advanced Analytics (DSAA), p. 260–269, 2018.

HANSEN, C.; STEINMETZ, H.; BLOCK, J. **How to conduct a meta-analysis in eight steps: a practical guide**. Management Review Quarterly, p. 1–19, 2021.

ISMAIL, Amelia Ritahani; ZAINAL ABIDIN, Nadzurah; MAEN, Mhd. **Systematic Review on Missing Data Imputation Techniques with Machine Learning Algorithms for Healthcare**. Journal of Robotics and Control (JRC), v. 3, p. 143–152, 2022. DOI: 10.18196/jrc.v3i2.13133.

JONES, Rebecca Louise; KHARB, Aditi; TUBEUF, Sandy. **The untold story of missing data in disaster research: a systematic review of the empirical literature utilising the Emergency Events Database (EM-DAT)**. Environmental Research Letters, v. 18, n. 10, p. 103006, 2023. DOI: 10.1088/1748-9326/acfd42.

KOLASKI, K.; LOGAN, L. R.; IOANNIDIS, J. P. A. **Orientações sobre as melhores ferramentas e práticas para revisões sistemáticas**. Systematic Reviews, v. 12, 96, 2023. https://doi.org/10.1186/s13643-023-02255-9.

LITTLE, R.; RUBIN, D. **Statistical Analysis with Missing Data**. 3. ed. Hoboken: John Wiley & Sons, 2019.

LIU, Mingxuan *et al.* **Handling missing values in healthcare data: a systematic review of deep learning-based imputation techniques**. Artificial Intelligence in Medicine, v. 142, p. 102587, ago. 2023. DOI: 10.1016/j.artmed.2023.102587.

MACUCULE, A. J.; VALENTIM, M. L. P. **Informação, conhecimento e inteligência organizacional: das bases conceptuais ao comportamento epistemológico na Ciência da Informação**. Transinformação, v. 36, 2024. Recuperado de https://periodicos.puc-campinas.edu.br/transinfo/article/view/8501.

MASCONI, Katya L. *et al.* **Reporting and handling of missing data in predictive research for prevalent undiagnosed type 2 diabetes mellitus: a systematic review**. EPMA Journal, v. 6, n. 1, p. 7, 2015. DOI: 10.1186/s13167-015-0028-0.

MUKHERJEE, Kumar *et al.* **Handling missing data in health economics and outcomes research (HEOR): a systematic review and practical recommendations**. PharmacoEconomics, v. 41, n. 12, p. 1589–1601, 2023. DOI: 10.1007/s40273-023-01297-0.

NIJMAN, S. W. J. *et al.* **Missing data is poorly handled and reported in prediction model studies using machine learning: a literature review**. Journal of Clinical Epidemiology, v. 142, p. 218–229, fev. 2022. DOI: 10.1016/j.jclinepi.2021.11.023.

OLIVEIRA, B. M. **Revisão sistemática de literatura (RSL): um guia da teoria à prática**. Ifpe.edu.br, 2023. Disponível em: https://repositorio.ifpe.edu.br/xmlui/handle/123456789/997.

PIGOTT, T. D.; POLANIN, J. R. **Methodological Guidance Paper: High-Quality Meta-Analysis in a Systematic Review**. Review of Educational Research, SAGE Publications, 2020. DOI: 10.3102/0034654319877153.

RABE, Brooke A. *et al.* **Missing data handling in non-inferiority and equivalence trials: a systematic review**. Pharmaceutical Statistics, v. 17, n. 5, p. 477–488, 2018. DOI: 10.1002/pst.1867.

REN, Lijuan; WANG, Tao; SEKLOULI-SEKHARI, Aicha; ZHANG, Haiqing; BOURAS,

Abdelaziz. **A review on missing values for main challenges and method**s. Information Systems, v. 119, p. 102268, 2023. DOI: 10.1016/j.is.2023.102268.

RICHTER, Sophie *et al.* **Handling of missing outcome data in traumatic brain injury research: a systematic review**. Journal of Neurotrauma, v. 36, n. 17, p. 2743–2752, 2019. DOI: 10.1089/neu.2018.6216.

RUBIN, D. B. **Inference and missing data**. Biometrika, v. 63, n. 3, p. 581–592, 1976.

SARMENTO, Hugo; ROSADO, António; SAAVEDRA, Néstor. **Revisão Sistemática da Literatura** - Ebook, 2024.

SCHUMANN, Yannis; GOCKE, Antonia; NEUMANN, Julia E. **Computational Methods for Data Integration and Imputation of Missing Values in Omics Datasets**. Proteomics, v. 25, n. 1-2, p. e202400100, jan. 2025. DOI: 10.1002/pmic.202400100.

SETIAWAN, Ismail; GERNOWO, Rahmat; WARSITO, Budi. **A Systematic Literature Review on Missing Values: research trends, datasets, methods and frameworks**. E3S Web of Conferences, v. 448, p. 02020, 2023. DOI: 10.1051/e3sconf/202344802020.

SPINELI, Loukia M.; PANDIS, Nikolaos; SALANTI, Georgia. **Reporting and handling missing outcome data in mental health: a systematic review of Cochrane systematic reviews and meta-analyses**. Research Synthesis Methods, v. 6, n. 2, p. 175–187, 2015. DOI: 10.1002/jrsm.1131.

THOMAS, Tressy; RAJABI, Enayat. **A systematic review of machine learning-based missing value imputation techniques**. Data Technologies and Applications, v. 55, n. 4, p. 558–585, 2021. DOI: 10.1108/DTA-12-2020-0298.

TROYANSKAYA, O. *et al.* **Missing value estimation methods for DNA microarrays**. Bioinformatics, v. 17, n. 6, p. 520–525, 2001.

VAN BUUREN, S.; GROOTHUIS-OUDSHOORN, **K. mice: Multivariate imputation by chained equations in R**. Journal of Statistical Software, v. 45, n. 3, p. 1–67, 2011.

WIRTH, R. CRISP-DM: **Towards a standard process model for data mining**. In: Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining, p. 29–39, 2000.

WISSLER, Amanda; BLEVINS, Kelly E.; BUIKSTRA, Jane E. **Missing data in bioarchaeology I: a review of the literature**. American Journal of Biological Anthropology, v. 179, n. 3, p. 339–348, 2022. DOI: 10.1002/ajpa.24609.

ZHANG, Yuqing; ZHANG, Runtong; ZHAO, Butian. **A systematic review of generative adversarial imputation network in missing data imputation**. Neural Computing and Applications, v. 35, n. 27, p. 19685–19705, 2023. DOI: 10.1007/s00521-023-08840-2.

ZHOU, Y.; ARYAL, S.; BOUADJENEK, M. R. **Review for handling missing data with a special missing mechanism**. arXiv preprint arXiv:2404.04905, 2024.