Unsupervised Learning for Identifying Clinical and Symptom-Based Profiles in Endometriosis: A Systematic Literature Review

Sabrina Ramos de Freitas[1], **Diana Francisca Adamatti**[2]

[1]Centro de Ciências Computacionais (C3) – Universidade Federal do Rio Grande (FURG)
Rio Grande – RS – Brazil

[2]Campus Serra – Universidade Federal do Rio Grande do Sul (UFRGS)
Caxias do Sul – RS – Brazil

sabrina.freitas@furg.br, dianaada@gmail.com

Abstract. *Endometriosis is a heterogeneous condition that is not fully represented by current staging systems. This Systematic Literature Review investigated the use of unsupervised machine learning to identify clinical symptom profiles in patients with endometriosis. Eight studies were included, using variables such as pain, comorbidities, and quality of life. The findings revealed distinct profiles related to symptom burden, multimorbidity, and biopsychosocial impact. These methods may complement anatomical classifications but still require external and longitudinal validation before clinical application.*

Resumo. *A endometriose é uma condição heterogênea que não é plenamente representada pelos atuais sistemas de estadiamento. Esta Revisão Sistemática da Literatura investigou o uso do aprendizado de máquina não supervisionado na identificação de perfis clínico-sintomáticos em pacientes. Foram incluídos oito estudos usando variáveis como dor, comorbidades e qualidade de vida. Os resultados evidenciaram perfis distintos relacionados à carga sintomática, à multimorbidade e ao impacto biopsicossocial. Conclui-se que esses métodos podem complementar as classificações anatômicas, mas ainda requerem validação externa e longitudinal antes da aplicação clínica.*

1. Introdução

A endometriose é caracterizada pela presença de tecido semelhante ao endométrio fora da cavidade uterina, mas essa definição anatômica não representa integralmente sua complexidade sintomática, patobiológica e multissistêmica [Zondervan et al. 2020]. Estima-se que a condição afete aproximadamente 10% das mulheres em idade reprodutiva, o que corresponde a cerca de 190 milhões de pessoas no mundo. Essa estimativa permanece imprecisa em razão da heterogeneidade clínica, da sobreposição dos sintomas com outras condições e da dependência histórica de confirmação por visualização cirúrgica [Zondervan et al. 2020, Taylor et al. 2021]. Entre as manifestações frequentemente relatadas estão dismenorreia, dor pélvica menstrual e não menstrual, dispareunia, infertilidade, fadiga e sintomas gastrointestinais e urinários, frequentemente acompanhados por outras comorbidades e por prejuízos à qualidade de vida e à produtividade [Nnoaham et al. 2011, Fuldeore and Soliman 2017].

Essa heterogeneidade, no entanto, não é plenamente contemplada pelos sistemas de classificação anatômicos e cirúrgicos utilizados na prática clínica. A classificação revisada da *American Society for Reproductive Medicine* (ASRM), empregada para descrever a extensão anatômica das lesões, apresenta correlação limitada com a intensidade da dor, a resposta ao tratamento e o prognóstico [Montanari et al. 2019, Zondervan et al. 2020]: pessoas classificadas em estágios iniciais podem apresentar elevada carga sintomática, enquanto aquelas em estágios avançados podem relatar sintomas menos intensos. A classificação ENZIAN amplia a descrição anatômica da endometriose profunda, mas, isoladamente, também não explica toda a diversidade das manifestações dolorosas e funcionais [Taylor et al. 2021]. Essas limitações indicam que a caracterização clínica da endometriose deve incorporar dimensões adicionais à localização e à extensão das lesões.

Nesse contexto, métodos de aprendizado não supervisionado oferecem uma abordagem complementar, ao permitirem investigar padrões nos dados sem depender de rótulos clínicos previamente definidos [Hastie et al. 2009, Bishop 2006]. Algoritmos de agrupamento, análise de classes latentes e modelos probabilísticos podem organizar pacientes segundo diferentes combinações de sintomas, dor, comorbidades, qualidade de vida e fatores psicossociais [Aggarwal and Reddy 2013]: enquanto algumas abordagens atribuem cada indivíduo a um único grupo, outras estimam graus de pertencimento a diferentes perfis latentes [Hastie et al. 2009]. Esses métodos podem, assim, contribuir tanto para a identificação de padrões clínicos quanto para a geração de hipóteses sobre novas formas de estratificação da endometriose.

Revisões anteriores já mapearam o uso de inteligência artificial na endometriose. [Sivajohan et al. 2022] conduziram uma revisão de escopo sobre aplicações clínicas de inteligência artificial na doença, reunindo principalmente modelos voltados ao diagnóstico e à predição a partir de diferentes tipos de dados, como variáveis clínicas, biomarcadores, dados genéticos e imagens. Mais recentemente, [Ramadan et al. 2025] discutiram o uso de aprendizado de máquina na detecção precoce da endometriose, contemplando tanto o agrupamento de sintomas quanto a análise de imagens; os autores destacam o potencial de métodos não supervisionados para identificar fenótipos a partir de sintomas autorrelatados e registros clínicos, mas ressaltam que essas abordagens permanecem menos exploradas do que os modelos voltados diretamente ao diagnóstico. Já revisões posteriores concentraram-se especificamente no campo das imagens: [Esmailzadeh et al. 2026] analisaram estudos baseados em ultrassonografia, enquanto [AlSaad et al. 2026] ampliaram o escopo para incluir também ressonância magnética e laparoscopia. Nessas revisões predominam as tarefas de diagnóstico, predição e análise de imagens, permanecendo a lacuna de uma síntese voltada especificamente aos trabalhos que empregam agrupamento e análise de classes latentes para caracterizar a heterogeneidade clínica da endometriose: ainda não está sistematizado como esses métodos têm sido empregados, quais fontes de dados e variáveis sustentam os agrupamentos, como o número de grupos é definido e quais características distinguem os perfis identificados.

Diante dessa lacuna, esta revisão busca responder à seguinte questão: *quais abordagens de aprendizado não supervisionado têm sido utilizadas para identificar, estratificar ou caracterizar agrupamentos de pacientes com endometriose, e quais características definem os perfis clínico-sintomáticos resultantes?*

2. Metodologia

A revisão foi organizada em três fases: planejamento, condução e apresentação dos resultados, seguindo as diretrizes propostas por [Kitchenham and Charters 2007] e as recomendações para revisões sistemáticas adotadas na literatura nacional [Scannavino et al. 2017, Mian et al. 2005]. Na fase de planejamento, foram definidos o objetivo, a questão de pesquisa, as fontes de informação, a estratégia de busca e os critérios de inclusão e exclusão. Na fase de condução, foram realizadas a identificação e a seleção dos estudos, seguidas da extração e da síntese dos dados. Por fim, os resultados foram organizados e apresentados por meio de uma síntese narrativa.

2.1. Fontes de informação e estratégia de busca

As buscas foram realizadas no PubMed[1], Scopus[2], ScienceDirect[3] e SciELO[4]. Essas fontes foram selecionadas por sua cobertura de publicações nas áreas de saúde, medicina, ciência de dados e inteligência artificial.

A estratégia de busca foi organizada em três blocos conceituais: população de interesse, manifestações clínicas e métodos de aprendizado não supervisionado. A sintaxe foi adaptada aos campos de busca e aos operadores disponíveis em cada fonte, preservando-se a seguinte estrutura:

```
(Endometriosis OR ``endometriosis-associated'')
AND (symptom* OR pain OR ``pelvic pain'' OR dysmenorrhea OR
dyspareunia OR fatigue OR ``quality of life'' OR ``symptom
profile*'')
AND (cluster* OR ``cluster analysis'' OR ``unsupervised
learning'' OR phenotyp* OR subgroup* OR profil* OR ``latent
class analysis'' OR ``latent profile analysis'')
```

A execução final da string nas quatro bases ocorreu no dia 15 de junho de 2026. Foram considerados estudos publicados em inglês, português ou espanhol. Não foi aplicada qualquer restrição quanto ao ano de publicação na etapa de busca; o recorte temporal observado nos estudos incluídos (2020 a 2025) decorre exclusivamente da disponibilidade de trabalhos elegíveis. Os registros recuperados foram exportados para um gerenciador de referências, no qual foram organizados para remoção de duplicatas e posterior triagem.

2.2. Critérios de elegibilidade

Foram incluídos estudos primários que atenderam aos seguintes critérios: (i) analisaram pacientes com endometriose; (ii) aplicaram pelo menos um método de aprendizado não supervisionado para identificar, estratificar ou caracterizar agrupamentos de pacientes; (iii) utilizaram variáveis clínicas, sintomáticas, funcionais, psicossociais, de comorbidades ou de utilização de serviços de saúde; (iv) descreveram a população, as variáveis de entrada, o método empregado e as características dos agrupamentos identificados; e (v) disponibilizaram o texto completo em inglês, português ou espanhol. Estudos primários disponibilizados como *preprints* foram elegíveis quando apresentaram informações suficientes para a extração dos dados e a avaliação metodológica.

[1] https://pubmed.ncbi.nlm.nih.gov/
[2] https://www.scopus.com/
[3] https://www.sciencedirect.com/
[4] https://www.scielo.org/

Foram excluídos estudos sem análise específica de pacientes com endometriose, que empregaram exclusivamente métodos supervisionados, que investigaram apenas dados genéticos, biomarcadores, histologia, imagens ou achados cirúrgicos sem caracterização clínica dos agrupamentos, que identificaram padrões predominantemente comportamentais sem relação direta com as manifestações da doença, ou que não forneceram informações suficientes sobre o método ou os perfis resultantes, além de revisões, editoriais, cartas, comentários, capítulos de livro e documentos sem acesso ao texto completo.

2.3. Seleção dos estudos, extração dos dados e síntese

A seleção foi conduzida em etapas sucessivas, detalhadas na Figura 1: remoção de duplicatas, análise dos títulos, leitura de títulos e resumos, avaliação preliminar das seções principais de cada trabalho e, por fim, leitura integral dos textos elegíveis segundo os critérios de inclusão e exclusão.

A triagem foi conduzida por um único avaliador, responsável pela aplicação sequencial dos critérios de elegibilidade. Os registros cuja elegibilidade permaneceu incerta, assim como todos os estudos avaliados em texto completo, foram revisados novamente para conferência. As divergências identificadas nessa releitura foram resolvidas por meio de nova análise dos critérios de elegibilidade até a obtenção de uma decisão consistente, registrada na planilha de controle da revisão, juntamente com o motivo de exclusão de cada estudo descartado após a leitura do texto completo.

Os dados dos estudos incluídos foram extraídos para uma planilha padronizada, que registrou a identificação e a procedência do estudo, o objetivo e o delineamento, as características e o tamanho da população, a fonte e a forma de coleta dos dados, os sintomas e as variáveis utilizadas na formação dos agrupamentos, o método não supervisionado e os procedimentos de pré-processamento, o critério de definição e o número de grupos, as características clínicas de cada perfil, as estratégias de validação ou avaliação e as limitações relatadas pelos autores.

Os estudos incluídos diferem quanto às populações analisadas, às fontes de dados, às variáveis de entrada e aos métodos de agrupamento empregados. Essa heterogeneidade inviabiliza a combinação estatística dos resultados, uma vez que os agrupamentos não são diretamente comparáveis entre si. Optou-se, portanto, por uma síntese narrativa, que abrange as características gerais dos estudos, incluindo os procedimentos de seleção, avaliação e validação das soluções de agrupamento, e os perfis identificados quanto à dor e à carga sintomática, às comorbidades e ao impacto funcional e biopsicossocial.

3. Resultados

3.1. Seleção dos estudos

A busca identificou 771 registros. Após a remoção de duplicatas e a aplicação dos critérios de elegibilidade, oito estudos foram incorporados à síntese narrativa, conforme apresentado na Figura 1.

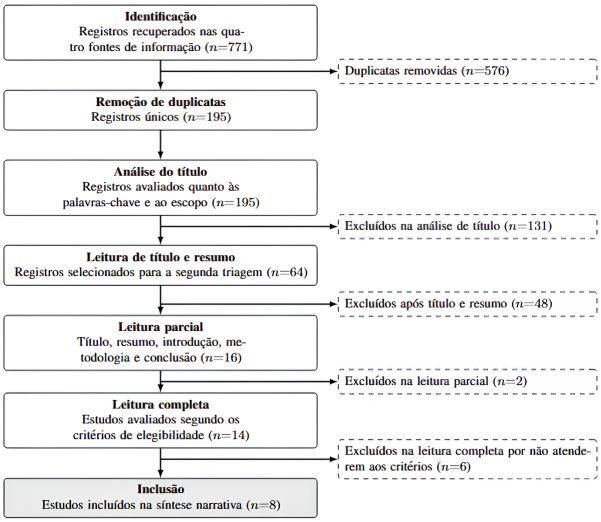


Figura 1. Fluxo de seleção dos estudos

3.2. Características gerais

Os oito estudos foram publicados ou disponibilizados entre 2020 e 2025 e apresentaram variações quanto ao tamanho amostral, às fontes de dados e ao delineamento. A Tabela 1 sintetiza suas principais características.

Tabela 1. Características dos estudos incluídos

Estudo	Amostra e fonte	Variáveis de entrada	Método	Nº de grupos
Urteaga et al. (2020)	4.368 participantes com diagnóstico autorrelatado; autorregistros do aplicativo *Phendo*	Localização, descrição e severidade da dor; sintomas gastrointestinais, geniturinários e sistêmicos; fluxo menstrual; impacto cotidiano; atividade sexual; uso de medicamentos	Modelo probabilístico de pertencimento misto	4
Sarria-Santamera et al. (2023)	4.055 mulheres; *Primary Care Clinical Database* do Sistema Nacional de Saúde da Espanha, 2013–2017	26 comorbidades com prevalência superior a 5% registradas na atenção primária	Agrupamento hierárquico (método de Ward)	6
Guare et al. (2024)	4.078 mulheres do conjunto não genotipado do *Penn Medicine Biobank*, vinculado a prontuários eletrônicos	Dezessete características derivadas de prontuários: sintomas, comorbidades e condições relacionadas à gestação com prevalência $\geq 5\%$	Clusterização espectral	5
Khan et al. (2025)	19.059 pacientes na UCSF e 24.453 em outros centros da *University of California* (UC-wide/UCHDW); prontuários estruturados segundo o OMOP	Diagnósticos estruturados representados por condições SNOMED	PCA, construção de grafo de vizinhança e algoritmo Leiden	21 e 26 (histórico completo); 31 e 41 (janela pré-diagnóstico)
Pereira & Kulak Junior (2025)	104 mulheres com indicação de tratamento cirúrgico atendidas em serviço terciário em Curitiba	Escores do FSFI e respostas às questões do EHP-30	*K-means*	2
Åkerblom et al. (2025)	113 participantes de clínica terciária de endometriose na Suécia	Intensidade da dor (NRS), extensão da dor (mapa de 36 áreas), ansiedade e depressão (HADS) e catastrofização (PCS)	Análise de classes latentes	2

Estudo	Amostra e fonte	Variáveis de entrada	Método	Nº de grupos
Kanti et al. (2024)	352 participantes de centro universitário de referência no Quebec	Frequência da dor pélvica, severidade de cinco sintomas álgicos (NRS) e qualidade sensorial/afetiva da dor (PQAS)	Análise de classes latentes	2
Vallée et al. (2025)	1.586 mulheres francesas; questionário on-line divulgado principalmente pelo Instagram	Nove manifestações: dispareunia, fluxo menstrual anormal ou intenso, infertilidade, disúria e disquezia durante a menstruação, sintomas digestivos, piora da dor ao longo do tempo, cólicas ou dor menstrual intensa e outros sintomas	PCA + *K-means*	8

Nota. PCA = análise de componentes principais. NRS = *Numeric Rating Scale*; PQAS = *Pain Quality Assessment Scale*; FSFI = *Female Sexual Function Index*; EHP-30 = *Endometriosis Health Profile*, versão de 30 itens; HADS = *Hospital Anxiety and Depression Scale*; PCS = *Pain Catastrophizing Scale*. OMOP = *Observational Medical Outcomes Partnership Common Data Model*; SNOMED = *Systematized Nomenclature of Medicine*; UCSF = *University of California, San Francisco*; UCHDW = *University of California Health Data Warehouse*.

De modo geral, os estudos relataram seus objetivos, as populações analisadas, as fontes de dados, as variáveis de entrada e os métodos de agrupamento. As principais diferenças metodológicas estiveram relacionadas à definição do número de grupos e às estratégias adotadas para examinar a estabilidade ou adequação das soluções. Nos estudos com análise de classes latentes, foram utilizados critérios de ajuste e separação probabilística, como AIC, BIC, testes de razão de verossimilhança e entropia. Entre as abordagens baseadas em *k-means* e agrupamento espectral, [Guare et al. 2024] compararam quatro algoritmos com base em silhueta, distorção e equilíbrio entre os tamanhos dos clusters; [Vallée et al. 2025] selecionaram a solução de oito grupos pelo maior valor do *Cubic Clustering Criterion* (CCC); e [Pereira and Junior 2025], embora o método do cotovelo e a silhueta média indicassem inicialmente três clusters, adotaram a solução de dois grupos por ser a única significativamente associada à classificação ASRM. No modelo probabilístico de [Urteaga et al. 2020], a seleção foi orientada principalmente pela interpretabilidade e esparsidade dos fenótipos, complementada por avaliação em dados retidos e análise de robustez a hiperparâmetros. Já o agrupamento hierárquico baseou-se na inspeção do dendrograma e na plausibilidade clínica dos grupos, enquanto a detecção de comunidades incluiu análise de sensibilidade ao parâmetro de resolução e comparações entre diferentes bases de dados e janelas diagnósticas.

A reprodução ou avaliação dos agrupamentos em amostras independentes foi pouco frequente. [Khan et al. 2025] compararam análises realizadas separadamente nas bases UCSF e UC-wide, examinando a concordância das associações de comorbidades e a correspondência entre clusters por meio da sobreposição de diagnósticos enriquecidos e da similaridade de Jaccard. [Urteaga et al. 2020] avaliaram o desempenho preditivo do modelo em dados retidos e compararam os fenótipos identificados com agrupamentos produzidos por especialistas e com respostas ao questionário clínico WERF EPHect, embora essas avaliações tenham sido realizadas dentro da mesma população de estudo. [Guare et al. 2024] combinaram métricas internas com revisão de prontuários em uma subamostra e, posteriormente, transferiram os fenótipos para diferentes biobancos para análises genéticas estratificadas. Nos demais trabalhos, a avaliação das soluções baseou-se principalmente em critérios internos de ajuste, associações com variáveis auxiliares da própria amostra ou análises de plausibilidade clínica.

3.3. Perfis clínico-sintomáticos identificados

Quatro estudos identificaram perfis associados a diferentes níveis de dor, carga sintomática ou comprometimento biopsicossocial [Urteaga et al. 2020, Pereira and Junior 2025, Åkerblom et al. 2025, Kanti et al. 2024]. [Kanti et al. 2024] encontraram dois fenótipos de baixa e alta carga de dor; o de alta dor apresentou pior qualidade de vida e associou-se ao uso de analgésicos e a condições como síndrome da bexiga dolorosa, fibromialgia, enxaqueca, dor lombar, síndrome do intestino irritável, ansiedade e depressão, enquanto a presença de endometrioma associou-se ao fenótipo de baixa dor. Em [Åkerblom et al. 2025], a classe de maior carga biopsicossocial apresentou valores mais elevados em todos os indicadores utilizados na formação das classes e, em análises posteriores, maior sensação de falta de controle e impotência e maior déficit de suporte social, avaliados junto à autoimagem pelo EHP-30. [Pereira and Junior 2025] também identificaram dois grupos entre mulheres com indicação de tratamento cirúrgico: o Cluster B apresentou dispareunia e disúria mais intensas, pior função sexual (FSFI) e maior comprometimento da qualidade de vida (EHP-30), com exceção do módulo de infertilidade, que não diferiu entre os clusters. Em [Urteaga et al. 2020], o fenótipo A, interpretado como o de maior gravidade, reuniu maior probabilidade de dor intensa, múltiplas localizações dolorosas, piores sintomas gastrointestinais e urinários, dificuldades nas atividades diárias, pior qualidade de vida e maior uso de analgésicos narcóticos, antidepressivos e medicamentos para dor neuropática; embora os registros fossem longitudinais, o modelo agregou as observações de cada participante, sem modelar explicitamente sua evolução temporal. Os demais fenótipos apresentaram menor gravidade e diferenciaram-se pelo tratamento: o B associou-se a terapias hormonais; o C, à ausência de tratamento médico; e o D, a uma combinação de terapias hormonais, narcóticos e antidepressivos.

Três estudos utilizaram prontuários eletrônicos ou registros de serviços de saúde para identificar padrões de diagnósticos e comorbidades [Sarria-Santamera et al. 2023, Guare et al. 2024, Khan et al. 2025]. [Sarria-Santamera et al. 2023] identificaram seis grupos: menor carga de comorbidades; múltiplas comorbidades; ansiedade e condições musculoesqueléticas; alergias tipo 1 ou hipersensibilidade imediata; anemia e infertilidade; e cefaleia ou enxaqueca. [Guare et al. 2024] identificaram cinco perfis: comorbidades dolorosas; distúrbios uterinos; complicações gestacionais; condições cardiometabólicas; e um grupo denominado pelos autores de "EHR-asymptomatic", com menor registro de sintomas e comorbidades no prontuário, sem que isso demonstre ausência clínica de sintomas. [Khan et al. 2025] identificaram dezenas de comunidades entre mais de 43 mil pacientes, com padrões relacionados a condições autoimunes, psiquiátricas e gestacionais; na comparação entre as duas bases destacaram-se agrupamentos correspondentes de gravidez e câncer no histórico completo e de gravidez e condições do trato urinário na janela pré-endometriose, e alguns subgrupos permaneceram consistentes entre as duas janelas em torno de hiperlipidemia, saúde mental, gravidez e anemia. A análise caso-controle identificou ainda associações significativas entre endometriose e condições como dor pélvica, enxaqueca e doenças gastrointestinais.

A qualidade de vida, a função sexual e os fatores psicossociais assumiram papéis distintos: integraram diretamente a formação dos grupos em [Pereira and Junior 2025] e [Åkerblom et al. 2025] e foram utilizados posteriormente para caracterizar os perfis em [Vallée et al. 2025], por meio do EHP-5, e em [Kanti et al. 2024]. Em [Vallée et al. 2025], maior presença de sintomas não correspondeu necessariamente a pior qualidade de vida:

o Cluster 5 reuniu alta prevalência de queixas como dispareunia e disquezia, mas todas as participantes foram classificadas na faixa de boa qualidade de vida segundo o critério adotado no estudo, enquanto o Cluster 6, que combinou múltiplas manifestações e piora da dor ao longo do tempo, teve todas as participantes na faixa de pior qualidade de vida.

4. Discussão

4.1. Síntese e interpretação dos achados

Esta revisão identificou diferentes aplicações de aprendizado não supervisionado na caracterização da heterogeneidade clínica da endometriose. Apesar da diversidade de métodos, populações e fontes de dados, os agrupamentos puderam ser organizados em três dimensões principais: carga dolorosa e sintomática, padrões de comorbidade e impacto funcional ou biopsicossocial.

Três aspectos foram recorrentes entre os estudos. O primeiro corresponde às características da dor – intensidade, frequência, localização ou extensão e diferentes manifestações dolorosas –, que participaram diretamente da formação dos grupos nos trabalhos baseados em dados autorrelatados [Urteaga et al. 2020, Åkerblom et al. 2025, Kanti et al. 2024, Vallée et al. 2025] e, em [Pereira and Junior 2025], foram comparadas posteriormente entre os clusters. O segundo é a presença de comorbidades associadas à dor, como síndrome do intestino irritável, enxaqueca ou cefaleia, dor lombar, fibromialgia e condições urinárias, que apareceram tanto nos agrupamentos derivados de prontuários eletrônicos quanto na caracterização posterior dos fenótipos de maior dor [Sarria-Santamera et al. 2023, Guare et al. 2024, Khan et al. 2025, Kanti et al. 2024]. O terceiro são as condições relacionadas à saúde mental, representadas como comorbidades nos registros clínicos [Sarria-Santamera et al. 2023, Khan et al. 2025], medidas por escalas psicométricas e utilizadas na formação das classes [Åkerblom et al. 2025] ou analisadas posteriormente como fatores associados ao fenótipo de maior dor [Kanti et al. 2024].

A qualidade de vida também esteve relacionada à diferenciação dos perfis, embora seu papel metodológico tenha variado. Ela acompanhou os perfis de maior gravidade em [Urteaga et al. 2020] e [Kanti et al. 2024] e, em [Pereira and Junior 2025], veio junto a piores escores de função sexual. O uso de medicamentos para dor associou-se de modo semelhante aos perfis mais sintomáticos nos dois primeiros estudos.

Essa correspondência, contudo, não foi universal. Em [Vallée et al. 2025], um dos grupos com melhor qualidade de vida apresentou frequências elevadas de diferentes sintomas, indicando que a presença de múltiplas manifestações e seu impacto sobre a vida cotidiana não variam necessariamente na mesma direção. A infertilidade também diferenciou pouco alguns perfis: não diferiu entre os clusters de [Pereira and Junior 2025] nem foi mais prevalente no fenótipo de alta dor de [Kanti et al. 2024], ao contrário das demais comorbidades avaliadas.

A composição dos agrupamentos variou conforme a natureza dos dados analisados. Estudos baseados em prontuários eletrônicos destacaram diagnósticos e comorbidades e produziram soluções mais subdivididas, enquanto questionários, escalas clínicas e dados gerados pelas próprias pacientes permitiram maior detalhamento da dor, da função sexual, da qualidade de vida e dos fatores psicológicos, em geral diferenciando perfis de menor e maior comprometimento. Assim, os perfis identificados devem ser interpretados como

representações condicionadas pelas variáveis disponíveis e pelo contexto de coleta, e não como subtipos clínicos universais e independentes da fonte dos dados.

4.2. Relação com classificações anatômicas

A relação entre os agrupamentos clínicos e os sistemas formais de classificação anatômica ou cirúrgica foi pouco investigada nos estudos incluídos. [Pereira and Junior 2025] foram os únicos a examinar diretamente a relação entre os agrupamentos identificados e um sistema formal de estadiamento, comparando-os com a classificação ASRM. Observou-se maior concentração de participantes em estágio IV no grupo com pior função sexual e qualidade de vida, com diferença significativa concentrada principalmente entre os estágios III e IV. A ASRM não foi utilizada como variável de entrada do *K-means*; contudo, a associação entre o estadiamento e as diferentes soluções de agrupamento foi utilizada para selecionar a solução final de dois clusters. Assim, essa associação não deve ser interpretada como validação externa independente dos agrupamentos, pois a própria classificação anatômica foi utilizada na seleção da solução final posteriormente caracterizada. Em [Guare et al. 2024], subtipos anatômicos de endometriose foram excluídos das variáveis utilizadas na formação dos clusters e analisados posteriormente para caracterizá-los. De modo semelhante, [Kanti et al. 2024] avaliaram características relacionadas ao tipo de lesão após a identificação dos fenótipos, observando associação entre a presença de endometrioma e o fenótipo de baixa dor. Nos demais estudos, classificações anatômicas formais não foram incorporadas à formação ou à avaliação dos agrupamentos.

Em conjunto, os achados indicam que os métodos não supervisionados podem complementar as classificações anatômicas ao representar dimensões clínicas que não são diretamente descritas pelo estadiamento das lesões. Enquanto os sistemas anatômicos caracterizam localização e extensão da doença, os agrupamentos clínico-sintomáticos podem representar como suas manifestações se combinam e afetam a vida das pacientes. Entretanto, os estudos disponíveis não demonstram que esses agrupamentos possam substituir as classificações existentes ou melhorar, de forma independente, decisões diagnósticas, prognósticas ou terapêuticas.

4.3. Implicações metodológicas e clínicas

A diversidade de abordagens metodológicas permitiu explorar diferentes dimensões da endometriose, mas também dificultou a comparação direta entre os resultados: o número e a composição dos grupos variaram conforme as características da população, as variáveis selecionadas, os procedimentos de padronização e redução de dimensionalidade, o algoritmo empregado e os parâmetros definidos em cada análise. Soluções com duas classes foram observadas principalmente em amostras clínicas menores e, em geral, distinguiram perfis de menor e maior comprometimento doloroso, funcional ou biopsicossocial, enquanto [Khan et al. 2025] identificaram dezenas de comunidades ao aplicar o algoritmo Leiden a extensos conjuntos de diagnósticos. Essa diferença mostra que o número de grupos não reflete necessariamente uma característica própria da endometriose, mas também depende do tipo e do nível de detalhamento dos dados, do objetivo da análise e das escolhas metodológicas adotadas.

A maioria dos estudos avaliou as soluções com base em critérios de ajuste, separação interna, estabilidade ou plausibilidade clínica, enquanto poucos verificaram sua reprodução em populações independentes. Além disso, diferenças estatísticas entre os

grupos não são suficientes, por si só, para demonstrar relevância clínica. Para que esses perfis possam apoiar estratégias de estratificação ou manejo personalizado, será necessário comprovar sua reprodutibilidade, estabilidade ao longo do tempo e relação com mudanças clinicamente relevantes.

A utilização de prontuários eletrônicos também exige cautela, pois os agrupamentos podem refletir não apenas manifestações da doença, mas também diferenças na codificação dos diagnósticos, no acesso ao cuidado, na frequência de utilização dos serviços e na completude dos registros. De forma semelhante, os dados autorrelatados permitem maior detalhamento dos sintomas, mas podem ser influenciados pelo perfil dos participantes, pela regularidade dos registros e pela percepção individual da experiência clínica. Esses fatores devem ser considerados na interpretação dos perfis identificados.

Estudos futuros devem buscar a integração de informações anatômicas, sintomas longitudinais, comorbidades, uso de medicamentos, fatores psicossociais e desfechos clínicos, descrevendo com transparência a seleção das variáveis, o tratamento de dados ausentes, a padronização, a redução de dimensionalidade, o ajuste de parâmetros e a definição do número de grupos. Recomenda-se ainda comparar soluções alternativas, realizar análises de sensibilidade e estabilidade e testar a reprodutibilidade em coortes externas, investigando possíveis vieses relacionados a características demográficas, ao acesso ao cuidado e aos padrões de utilização dos serviços de saúde.

4.4. Limitações da revisão

Esta revisão apresenta algumas limitações. A triagem dos estudos foi conduzida por um único avaliador, sem avaliação independente por um segundo revisor. Embora os registros de classificação não imediata e os estudos analisados em texto completo tenham sido reavaliados posteriormente, esse procedimento reduz, mas não elimina, a possibilidade de vieses na seleção dos trabalhos. Além disso, a estratégia de busca esteve restrita às bases e aos termos previamente definidos, de modo que estudos potencialmente relevantes podem não ter sido recuperados.

A heterogeneidade entre as populações, as fontes de dados, as variáveis utilizadas e os métodos de agrupamento dificultou a comparação direta entre os estudos e inviabilizou uma síntese quantitativa dos resultados. Também não foi realizada uma avaliação formal do risco de viés ou da qualidade metodológica por meio de um instrumento padronizado, embora aspectos relacionados à definição dos grupos, às estratégias de avaliação e à validação das soluções tenham sido considerados na extração e na discussão dos estudos. Por fim, o estudo de [Guare et al. 2024] foi incluído em sua versão de *preprint* e, portanto, ainda não havia sido publicado em periódico após revisão por pares no momento da realização desta revisão.

5. Conclusão

Os estudos analisados aplicaram métodos não supervisionados diversos – de análise de classes latentes e *k-means* a detecção de comunidades em grafos – a dados sobre sintomas, comorbidades, fatores psicossociais e qualidade de vida, provenientes de prontuários eletrônicos, questionários, escalas clínicas e dados gerados pelas próprias pacientes. De modo geral, os resultados revelaram perfis associados à carga sintomática, aos padrões de comorbidade e ao impacto funcional ou biopsicossocial, com recorrência de características

relacionadas à dor, a determinadas comorbidades e à saúde mental. Esses agrupamentos permitiram representar dimensões clínicas que não são diretamente contempladas pelos sistemas anatômicos de estadiamento. No entanto, ainda não configuram uma taxonomia comum nem subtipos clínicos consolidados.

A composição dos grupos variou conforme a população estudada, a fonte dos dados, as variáveis selecionadas e as decisões analíticas. Sua aplicação na estratificação clínica ainda requer reprodução em coortes independentes, avaliação da estabilidade ao longo do tempo e demonstração de utilidade prognóstica ou terapêutica. Até que essas evidências estejam disponíveis, as soluções de agrupamento devem ser interpretadas predominantemente como achados exploratórios e geradores de hipóteses.

Referências

Aggarwal, C. C. and Reddy, C. K., editors (2013). *Data Clustering: Algorithms and Applications*. CRC Press, Boca Raton.

Åkerblom, S., Jönsson, I. P., Ringqvist, Å., Nordengren, J., and Zhao, X. (2025). Identifying and characterizing clinical subgroups in individuals with endometriosis. *Frontiers in Pain Research*, Volume 6 - 2025.

AlSaad, R., Farrell, T., Elhenidy, A., Albasha, S., and Thomas, R. (2026). Artificial intelligence in endometriosis imaging: A scoping review. *AI*, 7(2).

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer, New York.

Esmailzadeh, A., Rashki Kemmak, A., Sezavar Dokhtfaroughi, S., Rasoulian, A., and Mazaheri Habibi, M. R. (2026). Investigating the role of artificial intelligence in the diagnosis and prediction of endometriosis using ultrasound images: a systematic review. *Reproductive Health*, 23(1):31.

Fuldeore, M. J. and Soliman, A. M. (2017). Prevalence and symptomatic burden of diagnosed endometriosis in the united states: National estimates from a cross-sectional survey of 59,411 women. *Gynecologic and obstetric investigation*, 82(5):453—461.

Guare, L., Humphrey, L. A., Rush, M., Pollie, M., Luo, Y., Weng, C., Wei, W.-Q., Kottyan, L., Jarvik, G., Elhadad, N., Penn Medicine Biobank, Regeneron Genetics Center, Zondervan, K., Missmer, S., Vujkovic, M., Velez-Edwards, D., Senapati, S., and Setia-Verma, S. (2024). Enhancing genetic association power in endometriosis through unsupervised clustering of clinical subtypes identified from electronic health records. *medRxiv*.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, New York, 2 edition.

Kanti, F. S., Allard, V., Métivier, A.-A., Lemyre, M., Arendas, K., and Maheux-Lacroix, S. (2024). Pain phenotypes in endometriosis: A population-based study using latent class analysis. *BJOG: An International Journal of Obstetrics & Gynaecology*, 132:492–503.

Khan, U., Oskotsky, T. T., Yilmaz, B. D., Roger, J., Gjoni, K., Irwin, J. C., Opoku-Anane, J., Elhadad, N., Giudice, L. C., and Sirota, M. (2025). Comorbidity analysis and clustering of endometriosis patients using electronic health records. *Cell Reports Medicine*, 6. doi: 10.1016/j.xcrm.2025.102245.

Kitchenham, B. and Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering. EBSE Technical Report EBSE 2007-001, Keele University and Durham University.

Mian, P. G., Conte, T., Natali, A. C. C., Biolchini, J., and Travassos, G. H. (2005). A systematic review process for software engineering. In *Proceedings of the 2nd Experimental Software Engineering Latin American Workshop (ESELAW'05)*, pages 1–6, Uberlândia, Brazil.

Montanari, E., Dauser, B., Keckstein, J., Kirchner, E., Nemeth, Z., and Hudelist, G. (2019). Association between disease extent and pain symptoms in patients with deep infiltrating endometriosis. *Reproductive BioMedicine Online*, 39:845–851. doi: 10.1016/j.rbmo.2019.06.006.

Nnoaham, K. E., Hummelshoj, L., Webster, P., d'Hooghe, T., de Cicco Nardone, F., de Cicco Nardone, C., Jenkinson, C., Kennedy, S. H., and Zondervan, K. T. (2011). Impact of endometriosis on quality of life and work productivity: a multicenter study across ten countries. *Fertility and Sterility*, 96:366–373.e8. doi: 10.1016/j.fertnstert.2011.05.090.

Pereira, J. O. and Junior, J. K. (2025). Exploring endometriosis before surgical treatment: unraveling pain, sexual function and quality of life patterns. *Revista Brasileira de Ginecologia e Obstetrícia*, 47.

Ramadan, Z. M., Mouazen, S. M., Khan, S. S., Khan, S., and Farag, N. S. (2025). Machine learning in the early detection of endometriosis: a literature review on symptom clustering and imaging integration. *Precision and Future Medicine*, 9(3):117–128.

Sarria-Santamera, A., Yemenkhan, Y., Terzic, M., Ortega, M. A., and del Barco, A. (2023). A novel classification of endometriosis based on clusters of comorbidities. *Biomedicines*, 11.

Scannavino, K. R. F., Nakagawa, E. Y., Fabbri, S. C. P. F., and Ferrari, F. C. (2017). *Revisão Sistemática da Literatura em Engenharia de Software: teoria e prática*. Elsevier.

Sivajohan, B., Elgendi, M., Menon, C., Allaire, C., Yong, P., and Bedaiwy, M. A. (2022). Clinical use of artificial intelligence in endometriosis: a scoping review. *npj Digital Medicine*, 5(109).

Taylor, H. S., Kotlyar, A. M., and Flores, V. A. (2021). Endometriosis is a chronic systemic disease: clinical challenges and novel innovations. *The Lancet*, 397:839–852. doi: 10.1016/S0140-6736(21)00389-5.

Urteaga, I., McKillop, M., and Elhadad, N. (2020). Learning endometriosis phenotypes from patient-generated data. *npj Digital Medicine*, 3:88.

Vallée, A., Arutkin, M., Ceccaldi, P.-F., and Ayoubi, J.-M. (2025). Quality of life identification by unsupervised cluster analysis: A new approach to modelling the burden of endometriosis. *PLOS ONE*, 20:e0317178.

Zondervan, K., M, C. B., and A, S. M. (2020). Endometriosis. *New England Journal of Medicine*, 382:1244–1256. doi: 10.1056/NEJMra1810764.