

ACURÁCIA DOS SCORES ALVARADO E AIR NO DIAGNÓSTICO DE APENDICITE AGUDA: UMA REVISÃO NARRATIVA DA LITERATURA RECENTE

KAROLINE COELHO NEDEL¹; THOMAZ DE AQUINO MORAES NETO²; PEDRO VITOR GODOY SILVA³; BIANCA KUHN HUBNER⁴; HIDYAN VICENZZO SILVA E LIMA⁵; CARLA ALBERICI PASTORE⁶;

¹*Universidade Federal de Pelotas - kcnedel@gmail.com*

²*Universidade Federal de Pelotas - thomazmoraes01@gmail.com*

³*Universidade Federal de Pelotas - pedrovitor.gsilva@gmail.com*

⁴*Universidade Federal de Pelotas - biancakuhnhubner@gmail.com*

⁵*Universidade Federal de Pelotas - hidyanvicenzzo15@gmail.com*

⁶*Universidade Federal de Pelotas - pastorecarla@yahoo.com.br*

1. INTRODUÇÃO:

A apendicite aguda desponta como uma das mais relevantes causas de abdome agudo atendidas em serviços de emergência e representa, assim, uma das principais indicações de cirurgia abdominal de urgência em todo o mundo. Destarte, o diagnóstico precoce é imprescindível para reduzir complicações como perfuração e abscesso, mas nem sempre é simples, especialmente diante de apresentações clínicas atípicas.

Em vista disso, a fim de dar ferramentas que possam embasar o raciocínio clínico, diversos escores foram desenvolvidos ao longo dos anos. O escore de Alvarado permanece como um dos mais difundidos por sua simplicidade e fácil aplicação, integrando sinais, sintomas e achados laboratoriais. Contudo, estudos recentes demonstraram que, apesar de sua alta sensibilidade (95%), a especificidade pode ser bastante baixa (7%), ambas de acordo com o estudo de FARAHBAKHSH et al., 2020. O que resultaria em risco elevado de falso-positivos e, consequentemente, em cirurgias desnecessárias.

Frente a esse desafio, novas ferramentas diagnósticas têm sido propostas. O *Appendicitis Inflammatory Response* (AIR) score, que incorpora parâmetros da resposta inflamatória, surge como alternativa para mitigar os problemas já expostos. Nesse contexto, torna-se necessário, do ponto de vista epidemiológico, avaliar a acurácia desses escores na prática clínica, a fim de verificar sua aplicabilidade e impacto no diagnóstico da apendicite aguda.

2. METODOLOGIA:

Foram utilizadas as bases de dados PubMed, SciELO e LILACS para a realização da pesquisa teórica. Empregaram-se os descriptores em inglês “appendix” OR “acute appendicitis” AND “alvarado” AND “accuracy”, bem como em português “apêndice” OR “apendicite aguda” AND “alvarado” AND “acurácia”. Foram incluídos artigos publicados nos últimos cinco anos que apresentavam, de forma numérica, os valores de acurácia dos escores AIR e Alvarado no diagnóstico de apendicite aguda. Identificaram-se 129 artigos, sendo 52 na PubMed, 76 na LILACS e 1 na SciELO. Após a aplicação dos critérios de elegibilidade, foram selecionados dois artigos, excluindo-se aqueles que não apresentavam dados numéricos de acurácia das respectivas escalas ou que se encontravam duplicados entre as bases pesquisadas. No presente artigo, não foi realizada distinção quanto ao grupo populacional ou condição clínica em que as escalas foram aplicadas, uma vez que o objetivo

consistiu em avaliar, de forma abrangente, a eficácia previamente descrita dessas ferramentas diagnósticas.

3. RESULTADOS E DISCUSSÕES:

Nos dois estudos incluídos, a acurácia dos escores variou conforme o cenário clínico e o desfecho analisado. No estudo A, realizado em um serviço de pronto-socorro no Irã (*A Comparative Study On The Diagnostic Validity Of Three Scoring Systems In The Diagnosis Of Acute Appendicitis In Emergency Centres*, 2020), o escore denominado Anderson, que é mesmo que escore AIR, apresentou acurácia de 64% (IC 95%: 54–71), sensibilidade de 77% (IC 95%: 69–82) e especificidade de 19% (IC 95%: 8–34). Nesse mesmo estudo, o escore de Alvarado apresentou uma acurácia maior, de 77% (IC 95%: 71–83), com sensibilidade de 95% (IC 95%: 91–98) e especificidade significativamente reduzida, de 7% (IC 95%: 1–19) (FARAHBAKHSH et al., 2020). Por outro lado, no estudo B nomeado *Diagnostic Value Of The MZXBTC Scoring System For Acute Complex Appendicitis* (2025), realizado por autores chineses, o escore AIR apresentou acurácia de 74%, sensibilidade de 68,37%, especificidade de 75,64% e *Area Under the Curve* (AUC) de 0,76, enquanto o escore de Alvarado alcançou acurácia de 57,9%, sensibilidade de 64,29%, especificidade de 55,64% e AUC de 0,66 (MA et al., 2025). É Importante destacar que o objetivo desse estudo foi prever casos de apendicite complexa em uma coorte cirúrgica e, diferentemente do estudo iraniano, os autores chineses não disponibilizaram intervalos de confiança dessas medidas, o que limitou a avaliação da precisão estatística e a comparabilidade.

A divergência dos resultados pode ser explicada por diferenças metodológicas e populacionais. No estudo iraniano, realizado em pronto-socorro e utilizando histopatologia como padrão-ouro, a coleta de PCR, uma variável fundamental no AIR/Anderson, foi restrita a determinados horários, o que pode ter prejudicado seu desempenho. Além disso, a avaliação clínica incluiu variabilidade entre examinadores, o que impacta na aplicação de itens semiológicos do escore (FARAHBAKHSH et al., 2020). Em contraste, o estudo chinês avaliou apenas pacientes submetidos a cirurgia, todos com apendicite confirmada, tendo como foco a complexidade da doença. Nesse cenário, a presença de marcadores inflamatórios como CRP, incorporados ao AIR, conferiu maior poder discriminatório, refletido em acurácia e AUC superiores (MA et al., 2025).

Do ponto de vista clínico, os achados demonstram que o escore Alvarado tende a apresentar uma sensibilidade muito alta, mas uma especificidade extremamente baixa, o que nos alerta para um possível grande número de falsos positivos e potenciais preparativos para apendicectomia desnecessários. Por outro lado, o AIR/Anderson apresenta desempenho mais equilibrado, variando conforme a logística laboratorial e o perfil populacional: em pronto-socorros com restrições para coleta rápida de CRP, sua performance pode ser inferior ao Alvarado, entretanto, em cenários cirúrgicos e de maior complexidade, o AIR mostra superioridade clara, com melhor acurácia e especificidade.

É fato que há necessidade de investimento em pesquisa capaz de avaliar prospectivamente ambos escores. É de suma importância ressaltar que mais investigações sobre essa temática são fundamentais, visto que podem orientar protocolos diagnósticos mais efetivos, acurados e com maior custo benefício.

ESCORES	ALVARADO	AIR/ANDERSON
Acurácia Estudo A	77%	64%
Acurácia Estudo B	57,9%	74%
Sensibilidade Estudo A	95%	77%
Sensibilidade Estudo B	64,29%	68,37%
Especificidade Estudo A	7%	19%
Especificidade Estudo B	55,64%	75,64%

5. CONCLUSÃO:

Segundo os dados expostos, pode-se inferir que a comparação entre os escores de Alvarado e AIR no diagnóstico de apendicite aguda revela que ambos apresentam vantagens e limitações distintas, dependendo do contexto clínico e da população estudada. Constatou-se que, embora o escore de Alvarado demonstre alta sensibilidade, sua baixa especificidade compromete a confiabilidade diagnóstica, elevando o risco de intervenções cirúrgicas desnecessárias. Em contrapartida, o escore AIR/Anderson apresentou desempenho mais equilibrado, com resultados superiores em cenários de maior complexidade clínica, especialmente quando associados à disponibilidade de exames laboratoriais como a dosagem de marcadores inflamatórios como CRP.

Diante desses achados, conclui-se que a escolha do escore ideal deve considerar as particularidades de cada serviço de saúde, incluindo a disponibilidade de exames complementares e o perfil epidemiológico dos pacientes. Além disso, evidencia-se que nenhum dos instrumentos avaliados é capaz, isoladamente, de substituir métodos complementares de diagnóstico, devendo ser considerados como ferramentas auxiliares na prática clínica.

Ademais, destaca-se a importância da realização de futuras pesquisas prospectivas e com diferentes perfis populacionais para legitimar esses escores em diferentes contextos clínicos, possibilitando a elaboração de protocolos diagnósticos mais acurados e de melhor custo-benefício. Dessa forma, será possível a realização de diagnósticos mais acurados e precisos, diminuindo a ocorrência de intervenções desnecessárias e complicações decorrentes da apendicite aguda.

6. REFERÊNCIAS

FARAHBAKHS, F.; TORABI, M.; MIRZAEI, M. A comparative study on the diagnostic validity of three scoring systems in the diagnosis of acute appendicitis in emergency centres. *African Journal of Emergency Medicine*, [s. l.], v. 10, n. 3, p. 132-135, 2020. DOI: <https://doi.org/10.1016/j.afjem.2020.04.009>

MA, Tianyi et al. Diagnostic value of the MZXBTC scoring system for acute complex appendicitis. *Scientific Reports*, v. 15, Article 1366, 2025. DOI: <https://doi.org/10.1038/s41598-025-85791-9>

