

ARTÍCULO ORIGINAL

Desempeño de las escalas AIR y Alvarado y los criterios de SIRS para predecir apendicitis aguda complicada en adultos: Estudio retrospectivo en La Dorada, Caldas (2021–2022)

Performance of the AIR and Alvarado scales and SIRS criteria in predicting complicated acute appendicitis in adults: A retrospective study in La Dorada, Caldas (2021–2022)

Santiago Orozco-Martínez, MD¹ , Manuela Orozco-Martínez, MD² ,
Wilmar Díaz-Quintero, MD³ , Luis Delgado-Arias, MD⁴ , Diego Castro-López, MD⁵ ,
Sara García-Chica, MD⁶ 

- 1 S.E.S Hospital Universitario de Caldas, Manizales, Colombia.
- 2 Programa de especialización en Medicina Interna, Universidad de Manizales, Manizales, Colombia.
- 3 Servicio de Cirugía general, E.S.E Hospital San Félix, La Dorada, Colombia.
- 4 Facultad de Medicina, Universidad de Manizales, Manizales, Colombia.
- 5 Programa de especialización en Cuidados Críticos, Universidad de Antioquia, Medellín, Colombia.
- 6 Hospital Alma Mater de Antioquia, Medellín, Colombia.

Resumen

Introducción. La apendicitis aguda complicada (AAC) representa una forma grave del abdomen agudo que requiere manejo oportuno. Las escalas clínicas como Alvarado y AIR y los criterios de SIRS han sido utilizados para su diagnóstico, pero su capacidad para predecir complicaciones no ha sido ampliamente evaluada en Colombia. El objetivo del estudio fue evaluar y comparar el desempeño diagnóstico de estas herramientas para identificar AAC en adultos.

Métodos. Estudio observacional, retrospectivo, de corte transversal, en adultos atendidos entre 2021 y 2022 en el Hospital San Félix, en La Dorada, Colombia. Se calcularon las puntuaciones en las escalas AIR y Alvarado y se determinó la presencia de SIRS. Se evaluaron sensibilidad, especificidad, valores predictivos, curvas ROC y el índice de Youden.

Fecha de recibido: 02/07/2025 - Fecha de aceptación: 13/09/2025 - Publicación en línea: 09/02/2026

Correspondencia: Santiago Orozco-Martínez, Avenida Alberto Mendoza # 77-40, Manizales, Colombia. Teléfono: +57 3052926057.

Dirección electrónica: santiom2001@gmail.com

Citar como: Orozco-Martínez S, Orozco-Martínez M, Díaz-Quintero W, Delgado-Arias L, Castro-López D, García-Chica S. Desempeño de las escalas AIR y Alvarado y los criterios de SIRS para predecir apendicitis aguda complicada en adultos: Estudio retrospectivo en La Dorada, Caldas (2021–2022). Rev Colomb Cir. 2026;41:329-38. <https://doi.org/10.30944/20117582.2986>

Este es un artículo de acceso abierto bajo una Licencia Creative Commons - BY-NC-ND <https://creativecommons.org/licenses/by-nc-nd/4.0/deed.es>

Resultados. Se incluyeron 365 pacientes mayores de 15 años con diagnóstico confirmado de apendicitis aguda. La escala AIR presentó el mejor desempeño diagnóstico (Área bajo la curva - AUC: 0,685), seguida de Alvarado (AUC: 0,621) y SIRS (AUC: 0,581). Los puntos de corte de AIR ≥ 7 y Alvarado ≥ 6 ofrecieron el mejor equilibrio entre sensibilidad y especificidad; los criterios de SIRS mostraron baja capacidad discriminatoria y no se recomendaron como herramienta diagnóstica aislada.

Conclusiones. La escala AIR es la herramienta más útil para predecir apendicitis complicada, mientras la escala de Alvarado conserva su utilidad en escenarios con recursos limitados. Se recomienda el uso de escalas clínicas estructuradas, complementadas con marcadores inflamatorios y estudios de imagen, para mejorar la toma de decisiones en el manejo de AAC.

Palabras clave: apendicitis; valor predictivo de las pruebas; síndrome de respuesta inflamatoria sistémica; medición de riesgo; sensibilidad y especificidad; curva ROC.

Abstract

Introduction. Complicated acute appendicitis (CAA) represents a severe form of acute abdominal pain that requires timely management. Clinical scales such as Alvarado and AIR, and the SIRS criteria have been used for diagnosis, but their ability to predict complications has not been widely evaluated in Colombia. The objective of this study was to evaluate and compare the diagnostic performance of these tools in identifying CAA in adults.

Methods. This was an observational, retrospective, cross-sectional study conducted in adults treated between 2021 and 2022 at San Félix Hospital in La Dorada, Colombia. AIR and Alvarado scores were calculated, and the presence of SIRS was determined. Sensitivity, specificity, predictive values, ROC curves, and the Youden index were evaluated.

Results. Three hundred sixty-five patients older than 15 years with a confirmed diagnosis of acute appendicitis were included. The AIR score showed the best diagnostic performance (AUC: 0.685), followed by Alvarado (AUC: 0.621) and SIRS (AUC: 0.581). AIR scores ≥ 7 and Alvarado scores ≥ 6 offered the best balance between sensitivity and specificity; the SIRS criteria showed low discriminatory capacity and were not recommended as a standalone diagnostic tool.

Conclusions. The AIR score is the most useful tool for predicting complicated appendicitis, while the Alvarado score retains its usefulness in resource-limited settings. The use of structured clinical scoring systems, complemented by inflammatory markers and imaging studies is recommended to improve decision-making in the management of acute appendicitis.

Keywords: appendicitis; predictive value of diagnostic tests; systemic inflammatory response syndrome; risk assessment; sensitivity and specificity; ROC curve.

Introducción

La apendicitis aguda (AA) es la causa más frecuente de abdomen agudo quirúrgico a nivel mundial, con un riesgo de presentación a lo largo de la vida del 8,6 % en hombres y del 6,7 % en mujeres¹. Su incidencia en países desarrollados se estima entre 90 y 100 casos por cada 100.000 habitantes por año, con un pico de presentación entre la segunda y tercera décadas de vida. En Colombia, entre 2017 y 2021 se registraron 345.618 casos, con una mayor prevalencia en mujeres (179.170) frente a los hombres (166.488)².

La AA se clasifica en formas no complicadas, cuando el proceso inflamatorio está confinado al apéndice (sin evidencia de perforación o necrosis), y en formas complicadas, que incluyen gangrena, perforación, formación de abscesos o peritonitis difusa¹. Esta distinción es fundamental, ya que determina el enfoque terapéutico y se asocia a diferentes desenlaces clínicos: mientras la forma no complicada puede manejarse en algunos casos con antibióticos, la forma complicada suele requerir intervención quirúrgica urgente, antibióticos de amplio espectro y, en ocasiones,

procedimientos adicionales como drenaje de abscesos o lavado peritoneal³. Se estima que entre un 20 % y un 30 % de los pacientes presentan apendicitis complicada, con mayor riesgo de morbilidad, estancia hospitalaria prolongada y sepsis. Por ello, su identificación precoz resulta clave para instaurar un manejo oportuno y mejorar los resultados clínicos⁴.

No obstante, el diagnóstico temprano y preciso de la AA, en especial de sus formas complicadas, sigue siendo un desafío. En la actualización de 2020 de las guías de Jerusalén¹ se señala que pocos estudios han logrado diferenciar con exactitud las formas complicadas de las no complicadas empleando únicamente criterios clínicos, lo que subraya la necesidad de integrar la evaluación clínica con marcadores inflamatorios y estudios de imagen.

Para apoyar esta labor diagnóstica, se han desarrollado diversas escalas que combinan hallazgos clínicos y parámetros de laboratorio (leucocitosis, neutrofilia y proteína C reactiva) para estimar la probabilidad de AA. Aunque la escala de Alvarado⁵ es ampliamente utilizada, presenta limitaciones en poblaciones específicas, como ancianos o inmunosuprimidos, y muestra baja especificidad para diferenciar las formas complicadas⁶. En contraste, el AIR score (*Appendicitis Inflammatory Response*), basado en variables más objetivas, ha demostrado un mejor desempeño para identificar formas avanzadas de apendicitis⁷. Según una revisión de Kularatna M, et al.⁸, esta escala alcanzó una sensibilidad del 92 % y una especificidad del 63 %.

En este contexto, también se ha explorado el papel del síndrome de respuesta inflamatoria sistémica (SIRS) como marcador de complicación, dado que refleja una respuesta inflamatoria más severa⁹. Si bien, ni las escalas clínicas ni los criterios de SIRS fueron concebidos originalmente para predecir la gravedad de la apendicitis, es plausible que una mayor puntuación o la presencia de SIRS se correlacionen con formas más avanzadas del cuadro, como perforación o sepsis abdominal. Su uso podría contribuir no solo al diagnóstico, sino también a la estratificación del riesgo, la

reducción de estudios innecesarios y la optimización del manejo quirúrgico¹⁰.

Sin embargo, en el contexto colombiano existe escasa evidencia sobre el desempeño real de estas herramientas en entornos clínicos. En particular, no encontramos que se haya establecido con claridad cuál de ellas predice con mayor precisión la presencia de apendicitis complicada en adultos. Por ello, el objetivo de este estudio fue evaluar y comparar la capacidad predictiva de las escalas AIR y Alvarado y los criterios de SIRS para identificar apendicitis aguda complicada en pacientes atendidos en el Hospital San Félix, en La Dorada, Colombia, entre enero de 2021 y diciembre de 2022.

Métodos

Tipo de estudio y población

Estudio observacional, retrospectivo y de corte transversal, en el que se revisaron 500 historias clínicas de pacientes con diagnóstico de apendicitis aguda. Se seleccionaron 365 pacientes mayores de 15 años, con diagnóstico confirmado mediante hallazgos quirúrgicos y patológicos, e información clínica completa. Se excluyeron los pacientes menores de 15 años, debido a que en esta población se recomienda el uso de otras escalas diagnósticas, así como aquellos casos con registros clínicos incompletos respecto a las variables de interés.

Variables

La variable dependiente fue la presencia de apendicitis complicada. Las variables independientes incluyeron las puntuaciones en AIR score y escala de Alvarado, además de la presencia de SIRS, que se calcularon de forma retrospectiva utilizando los signos y síntomas documentados en las historias clínicas al ingreso. Además, se describieron la edad, características clínicas como migración del dolor, emesis y fiebre subjetiva, así como comorbilidades (obesidad, tabaquismo y diabetes mellitus tipo 2). Los datos fueron anonimizados antes del análisis.

Análisis estadístico

Se usaron estadísticas descriptivas para caracterizar la población. Se calcularon sensibilidad, especificidad, valor predictivo positivo (VPP) y valor predictivo negativo (VPN) de las escalas para diferentes puntos de corte. Se construyeron curvas ROC (*Receiver Operating Characteristic*) y se calcularon las áreas bajo la curva (AUC), comparándolas mediante la prueba de DeLong. El punto óptimo de corte se determinó usando el índice de Youden. Un valor de $p < 0,05$ se consideró estadísticamente significativo. Todos los análisis se realizaron utilizando el Software Jamovi 2.6.44 (The jamovi project. Jamovi [Computer software]. Sydney: The jamovi project; 2021).

Resultados

Se incluyeron 365 pacientes mayores de 15 años con diagnóstico confirmado de apendicitis

aguda complicada, que fueron atendidos en el Hospital San Félix, en La Dorada, Colombia, entre enero de 2021 y diciembre de 2022. De ellos, 314 (86,0 %) eran mayores de 18 años y 51 (14,0 %) tenían entre 15 y 18 años; 171 (46,8 %) fueron hombres y 194 (53,2 %) mujeres (Tabla 1).

Entre las variables clínicas, se documentó dolor abdominal migratorio en 215 pacientes (58,9 %), emesis en 180 (49,3 %) y fiebre subjetiva en 118 (32,3 %). La presencia de SIRS se identificó en 103 pacientes (28,2 %). Respecto a las comorbilidades, se encontró que nueve (2,5 %) tenían diabetes mellitus tipo 2 y 45 (12,3 %) eran obesos. En la tabla 2 se presentan los resultados obtenidos al aplicar las escalas AIR, Alvarado y SIRS, de acuerdo con los síntomas y parámetros de laboratorio registrados al momento de la valoración inicial.

Los resultados del análisis de la curva ROC, incluyendo el área bajo la curva (AUC) con su

Tabla 1. Características clínicas y de laboratorios en los pacientes con apendicitis confirmada incluidos en el estudio

Variable	Frecuencia	%	Rango	Media (Desviación estándar)
Sexo				
Hombre	171	46,8	-	-
Mujer	194	53,2	-	-
Edad, años	-	-	15-90	34 (\pm 16,955)
Antecedentes				
Diabetes mellitus tipo 2	9	2,5	-	-
Tabaquismo	6	1,6	-	-
Obesidad	45	12,3	-	-
Síntomas				
Dolor migratorio	215	58,9	-	-
Emesis	180	49,3	-	-
Fiebre subjetiva	118	32,3	-	-
Examen físico				
Frecuencia cardiaca, latidos por minuto	-	-	56-176	87 (\pm 16,306)
Laboratorios				
Leucocitos (/ μ L)	-	-	1001-81160	13568 (\pm 5849,6)
Neutrófilos (%)	-	-	24-98	75 (\pm 1,24)
Proteína C reactiva (mg/L)	-	-	0,01-563	15,2 (\pm 44,9)

Fuente: Elaboración propia

intervalo de confianza del 95% y las estimaciones de desempeño diagnóstico para los diferentes puntos de corte evaluados (sensibilidad, especificidad, valores predictivos y precisión global), se presentan de manera resumida en la Tabla 3.

Para los criterios de SIRS, el análisis de la curva ROC mostró un área bajo la curva (AUC) de 0,578 (IC_{95%}: 0,525-0,632). La comparación con la línea de no discriminación mediante la prueba de DeLong evidenció que el AUC fue significativamente diferente de 0,5 (Z=2,87; p=0,004). El punto de corte identificado (SIRS < 0,5) presentó una sensibilidad de 39,6 %, una especificidad de 76,1 %, una precisión global de 40,4 % y un valor predictivo negativo (VPN) de 75,5 %.

El AIR score mostró un AUC de 0,685 (IC_{95%}: 0,629-0,741) en el análisis de la curva ROC. La prueba de DeLong confirmó que el rendimiento fue significativamente superior al azar (Z=6,44; p<0,001). Con un punto de corte ≤ 4, la sensibilidad fue del 95 %, la especificidad del 22 %, el valor predictivo positivo (VPP) del 38 % y el VPN del 92 %; mientras con un punto de corte ≥

9, la sensibilidad fue del 6 %, la especificidad del 98 %, el VPP del 50% y el VPN del 72 %. El índice de Youden identificó como punto de corte óptimo un valor cercano a 6,5, con una sensibilidad del 75 %, una especificidad del 54 %, un VPP del 40 % y un VPN del 84 %.

La escala de Alvarado presentó un AUC de 0,621 (IC_{95%}: 0,563-0,679). La prueba de DeLong mostró que el AUC fue significativamente superior a 0,5 (Z=4,08; p<0,001). Con un punto de corte ≤ 4, se obtuvo una sensibilidad del 99 %, una especificidad del 8 %, un VPP del 36 % y un VPN del 96 %; con un punto de corte ≥ 9, la sensibilidad fue del 5 %, la especificidad del 98 %, el VPP del 50 % y el VPN del 73 %. El índice de Youden identificó como punto de corte óptimo un valor cercano a 6,5, con una sensibilidad del 89 %, una especificidad del 33 %, un VPP del 35 % y un VPN del 88 %.

En la comparación de las curvas ROC, el AIR score presentó un AUC mayor que la escala de Alvarado (0,685 vs. 0,621). El rendimiento diagnóstico de los criterios de SIRS fue inferior, con un AUC de 0,581.

Tabla 2. Comparación de la clasificación de riesgo según Alvarado, AIR y SIRS de la cohorte

Alvarado	N (%)	AIR score	N (%)	SIRS	N (%)
Negativo (<4)	22 (6,0)	Bajo riesgo (0- 4)	62 (17,0)	No SIRS (<2 criterios)	262 (71,8)
Posible (5-6)	75 (20,5)	Riesgo intermedio (5-8)	253 (69,3)	SIRS (≥2 criterios)	103 (28,2)
Probable (7-8)	172 (47,1)	Riesgo alto (9-12)	50 (13,7)		
Muy probable (9-10)	96 (26,3)				

Fuente: Elaboración propia.

Tabla 3. Desempeño diagnóstico y puntos de corte de SIRS, AIR score y Alvarado (análisis ROC)

Escala	AUC (IC _{95%})	DeLong vs 0,5 (Z; p)	Punto de corte	S (%)	E (%)	VPP (%)	VPN (%)	Precisión (%)
SIRS**	0,578 (0,525–0,632)	Z=2,87; p=0,004	<0,500	39,6	76,1	—	75,5	40,4
AIR	0,685 (0,629–0,741)	Z=6,44; p<0,001	≤4	95	22	38	92	—
AIR	0,685 (0,629–0,741)	Z=6,44; p<0,001	≥9	6	98	50	72	—
AIR	0,685 (0,629–0,741)	Z=6,44; p<0,001	~6,5 (Youden)	75	54	40	84	—
Alvarado	0,621 (0,563–0,679)	Z=4,08; p<0,001	≤4	99	8	36	96	—
Alvarado	0,621 (0,563–0,679)	Z=4,08; p<0,001	≥9	5	98	50	73	—
Alvarado	0,621 (0,563–0,679)	Z=4,08; p<0,001	~6,5 (Youden)	89	33	35	88	—

* AUC: área bajo la curva; S: sensibilidad; E: especificidad; VPP: valor predictivo positivo; VPN: valor predictivo negativo.

** El punto de corte de SIRS se reporta según el umbral utilizado en el análisis ROC.

Fuente: Elaboración propia.

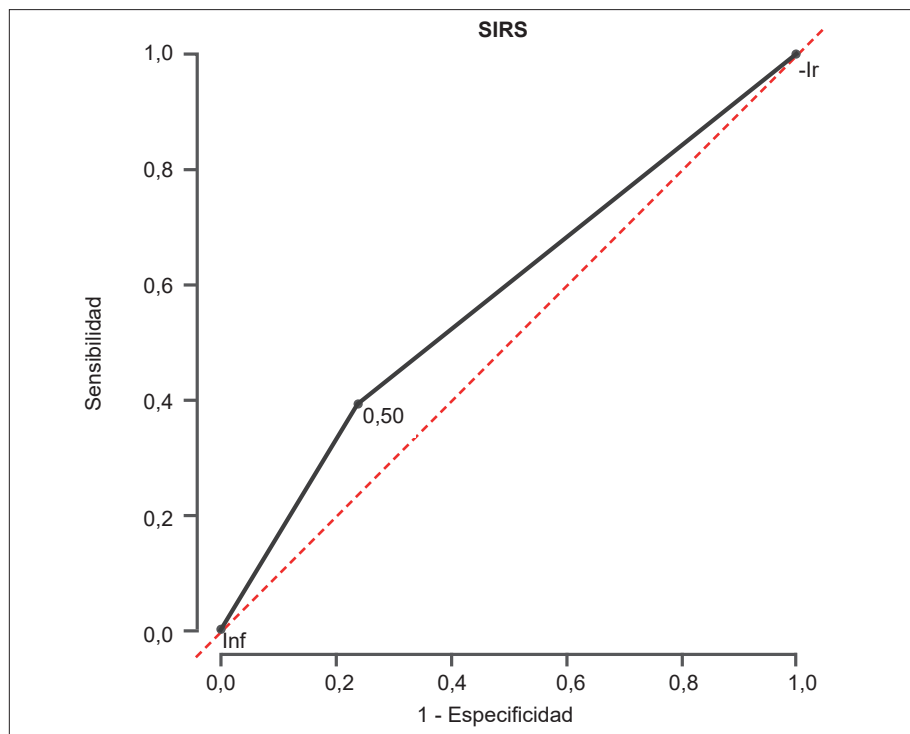


Figura 1. Curva ROC de SIRS

Fuente: Elaboración propia.

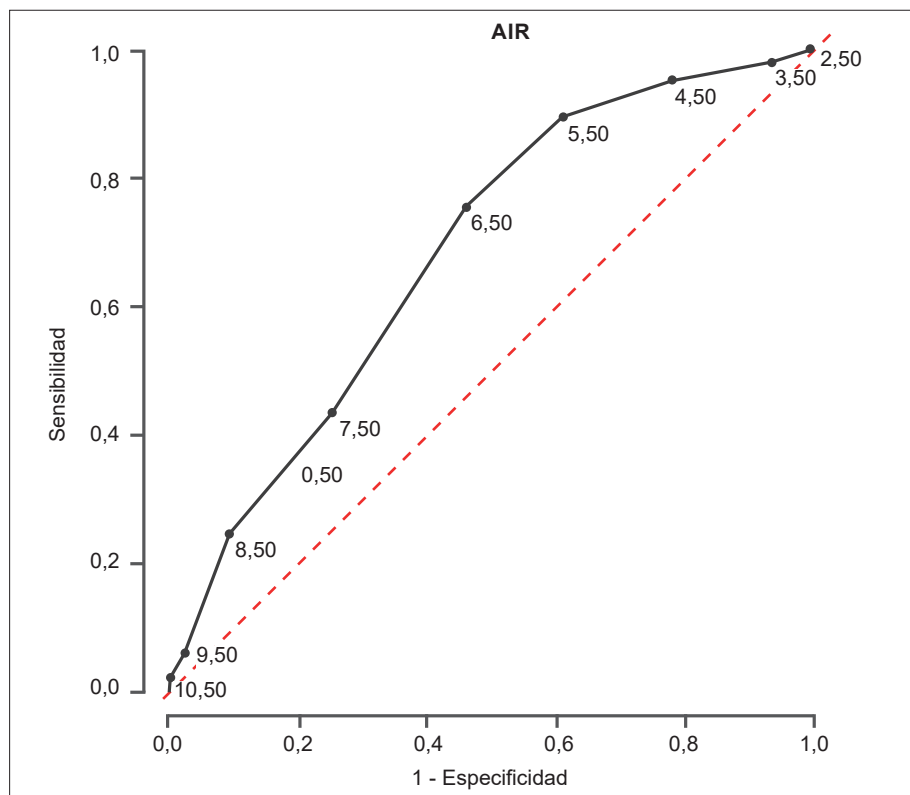


Figura 2. Curva ROC de AIR score

Fuente: Elaboración propia.

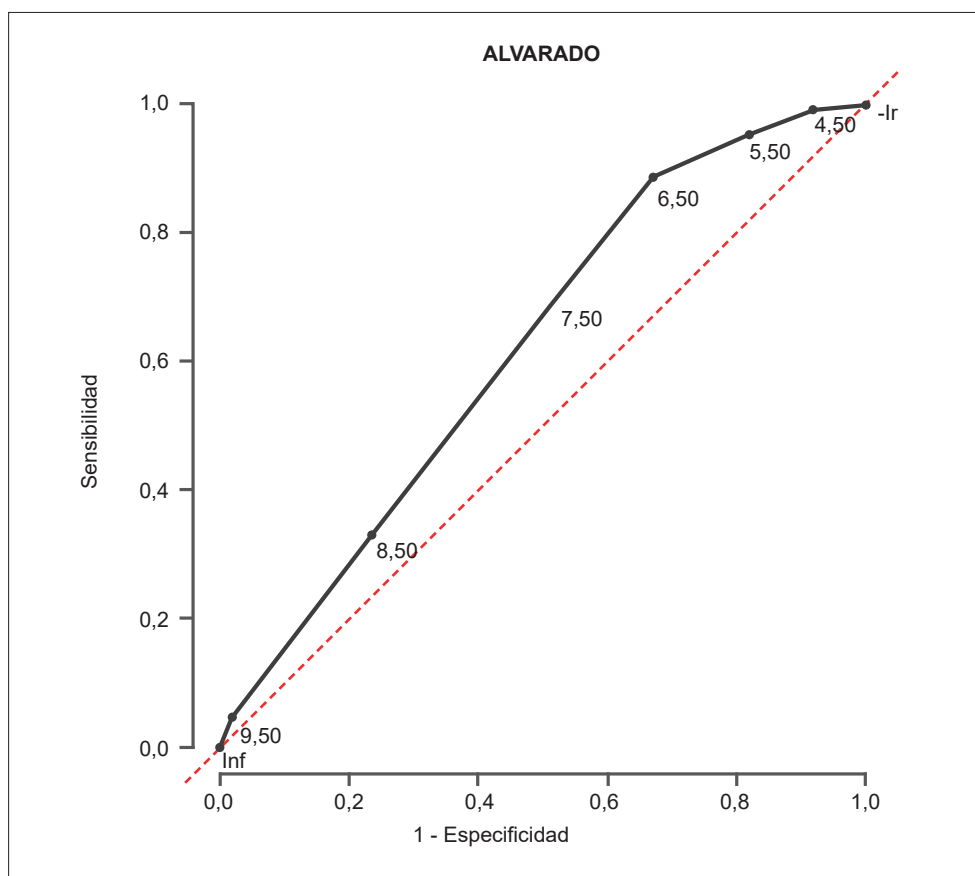


Figura 3. Curva ROC de escala de Alvarado

Fuente: Elaboración propia.

Discusión

Nuestro estudio evaluó la capacidad de las escalas AIR y Alvarado y los criterios SIRS para predecir apendicitis aguda complicada (AAC) en pacientes atendidos en el Hospital San Félix de La Dorada, Caldas. Los resultados mostraron que la capacidad de las escalas clínicas para discriminar AAC fue globalmente limitada. Los criterios de SIRS mostraron un desempeño pobre, cercano al azar ($IC_{95\%}$: 0,525-0,632), lo que confirma que no discrimina de forma eficaz entre casos complicados y no complicados, pese a alcanzar significación estadística ($Z=2,87$; $p=0,004$).

Por su parte, el AIR score evidenció una capacidad discriminativa moderada (AUC : 0,685; $IC_{95\%}$: 0,629-0,741), con un comportamiento

dependiente del punto de corte: valores bajos (≤ 4) permitieron descartar AAC con alta sensibilidad (95 %) y VPN (92 %), mientras que valores altos (≥ 9) mostraron elevada especificidad (98 %) y VPP (50 %), aunque con baja sensibilidad (6 %). La escala de Alvarado presentó un rendimiento inferior (AUC : 0,621; $IC_{95\%}$: 0,563-0,679), con utilidad principalmente en la exclusión de AAC a puntos de corte bajos (≤ 4 ; sensibilidad 99 %, VPN 96 %) y confirmación en valores elevados (≥ 9 ; especificidad 98 %, VPP 50 %). En conjunto, AIR score mostró mejor capacidad discriminativa que Alvarado (0,685 vs. 0,621), mientras que SIRS no resultó útil como predictor independiente (AUC 0,581), lo que resalta las limitaciones de las escalas clínicas cuando se emplean de forma aislada.

Estos hallazgos coinciden parcialmente con estudios regionales como el de Hernández-Cortez J, et al.¹¹, en México, donde el AIR score alcanzó un AUC de 0,78. Las diferencias observadas podrían deberse a variaciones en el acceso a pruebas de laboratorio, como la proteína C reactiva (PCR), o a diferencias poblacionales. Es importante destacar que en nuestra cohorte se evidenció una superposición significativa en los rangos de puntuación entre pacientes con apendicitis complicada y no complicada, lo cual reduce la capacidad discriminatoria de las escalas en los puntos de corte intermedios. Esta limitación fue documentada previamente por Haak F, et al.¹², quienes señalaron que para alcanzar una sensibilidad $\geq 90\%$ en la detección de AAC es necesario utilizar puntos de corte bajos (AIR ≥ 3 o Alvarado ≥ 5), lo que se traduce en especificidades $\leq 19\%$ y un alto número de falsos positivos.

Evidencia más reciente, como el metaanálisis de Andersson & Stark¹³ con más de 15.000 pacientes, respalda que el AIR score supera a la escala de Alvarado para identificar AAC (AUC: 0,93 vs. 0,88). Ese mismo estudio destaca que los puntajes extremos del AIR son altamente útiles: un AIR ≥ 9 ofrece una especificidad del 98-99 % para apendicitis complicada, mientras que un AIR < 4 presenta una sensibilidad del 95,99 % para descartarla. Sin embargo, en nuestro entorno, el rendimiento fue más modesto, posiblemente por factores clínicos y sociodemográficos como la alta prevalencia de comorbilidades, que podrían alterar la expresión clínica habitual y el valor predictivo de las escalas.

La escala de Alvarado, aunque ampliamente utilizada, mostró un desempeño diagnóstico inferior. Su principal fortaleza es la alta sensibilidad (99 % con punto de corte ≤ 4), lo que la hace útil para descartar AAC en contextos con recursos limitados. No obstante, su especificidad extremadamente baja (8 %) limita considerablemente su utilidad para confirmar la enfermedad, especialmente en poblaciones con presentación atípica, como adultos mayores o inmunocomprometidos.

Respecto al SIRS, aunque es una herramienta de cálculo simple, su bajo rendimiento diagnóstico (AUC: 0,581) y su limitada especificidad lo

hacen insuficiente como predictor aislado¹⁴. Sin embargo, nuestros resultados respaldan su asociación con formas más graves de apendicitis: observamos una mayor frecuencia de criterios de SIRS al ingreso en pacientes con AAC en comparación con los no complicados¹⁵. Esto concuerda con estudios como el de Bukhari & Mirza¹⁶, quienes reportaron una correlación significativa entre SIRS y hallazgos intraoperatorios de apendicitis perforada o gangrenosa ($p=0,001$) y concluyeron que unos criterios de SIRS positivos pueden tener valor predictivo para AAC, al estar presente en la mayoría de los casos complicados.

Nuestros hallazgos se alinean con esta evidencia, y sugieren que la evaluación inicial de pacientes con sospecha de AA debe incluir activamente la búsqueda de signos de respuesta inflamatoria sistémica (taquicardia, fiebre, taquipnea, leucocitosis), ya que su presencia podría alertar sobre un proceso más avanzado o una posible sepsis abdominal. No obstante, debe reconocerse que los criterios de SIRS carecen de especificidad etiológica, pues también pueden encontrarse en otras condiciones inflamatorias o infecciosas (como pancreatitis o neumonía). Por ello, su utilidad en el contexto de la apendicitis se limita a señalar gravedad potencial, más que a confirmar el diagnóstico de AAC.

Estos hallazgos tienen implicaciones clínicas relevantes. En instituciones con acceso a pruebas de laboratorio, recomendamos emplear la escala AIR complementada con marcadores como la PCR y, cuando sea posible, estudios de imagen (ecografía o tomografía), especialmente en pacientes con AIR ≥ 9 , donde la probabilidad de complicación es alta. En centros con recursos limitados, una estrategia costo-efectiva podría ser el uso de la escala de Alvarado con punto de corte ≤ 4 , combinada con ecografía abdominal, si está disponible, para orientar el triage inicial, aunque sin ignorar su baja especificidad.

Cabe destacar que las escalas clásicas no fueron diseñadas específicamente para predecir complicaciones, lo cual ha motivado el desarrollo de nuevos modelos que integran variables clínicas, de laboratorio y de imagen. Por ejemplo,

Düzgün AP, et al.⁴, en 2024 propusieron una escala modificada que incluye hiperbilirrubinemia, grosor apendicular en imágenes y sexo femenino, logrando una sensibilidad del 86 % y una especificidad del 90 % para AAC. Sin embargo, estos modelos aún requieren validación externa, especialmente en contextos como el nuestro, donde el acceso a recursos diagnósticos puede ser limitado y la carga de comorbilidad es elevada.

Al comparar nuestros hallazgos con los de la literatura reciente¹⁷, se confirma que la diferenciación preoperatoria entre cuadros de apendicitis simple y complicada sigue representando un desafío clínico considerable. Las escalas ampliamente difundidas, como Alvarado y AIR, ofrecen valor en la evaluación inicial de la sospecha de AA, pero muestran un rendimiento subóptimo para discriminar la presencia de complicaciones severas, especialmente si se utilizan de forma aislada. La incorporación de criterios de SIRS o de marcadores inflamatorios mejora la sensibilidad para detectar cuadros avanzados, aunque a expensas de la especificidad.

En la práctica clínica, la toma de decisiones debe seguir basándose en una integración cuidadosa de los hallazgos clínicos, los resultados de laboratorio (como PCR y leucocitos) y los estudios de imagen, más que en una puntuación aislada. En pacientes con alta probabilidad de AAC (por criterios clínicos, de laboratorio o de imagen), la conducta más segura sigue siendo la cirugía temprana, como lo recomiendan las guías actuales, para evitar la progresión del daño peritoneal. En contraste, en pacientes de bajo riesgo (clínica poco florida, puntajes bajos, sin criterios de SIRS ni hallazgos radiológicos de complicación), podría contemplarse un manejo no operatorio bajo estricta observación, con un umbral bajo para intervenir ante cualquier deterioro.

Conclusiones

El AIR *score* mostró la mejor capacidad para identificar apendicitis aguda complicada en nuestra población, con un desempeño superior al de la escala de Alvarado y la presencia de SIRS. Su utilidad es especialmente relevante en centros con

acceso a pruebas de laboratorio, que permite una estratificación de riesgo más precisa. La escala de Alvarado conserva valor en contextos con recursos limitados, principalmente para descartar enfermedad complicada a puntos de corte bajos (≤ 4), lo cual adquiere especial relevancia en Colombia, donde aproximadamente el 30 % de la población vive en zonas rurales¹⁸ con acceso limitado a laboratorios o estudios de imagen. En contraste, los criterios de SIRS no demostraron utilidad como predictor independiente.

En conjunto, ninguna escala es suficiente por sí sola para predecir apendicitis complicada, por lo que su uso debe integrarse con la evaluación clínica, marcadores inflamatorios y estudios de imagen. Estos hallazgos respaldan la necesidad de adaptar las herramientas diagnósticas al nivel de recursos disponible, con el objetivo de optimizar la identificación temprana de casos complicados y mejorar los desenlaces clínicos. Futuros estudios prospectivos deberán validar herramientas combinadas que integren variables clínicas, socio-demográficas, analíticas e imagenológicas.

Cumplimiento de normas éticas

Consentimiento informado: Este estudio se rigió por los principios éticos de la Declaración de Helsinki y por lo establecido en la Resolución 008430 de 1993 del Ministerio de Salud de Colombia, que regula la investigación en seres humanos. Fue clasificado como investigación sin riesgo, al basarse en fuentes documentales sin intervención directa sobre los pacientes, por lo que no se requirió diligenciamiento de consentimiento informado. El protocolo fue aprobado por el Comité de Ética Institucional del Hospital San Félix.

Conflictos de interés: Los autores declararon que no tienen ningún conflicto de interés.

Uso de Inteligencia Artificial: Los autores declararon que utilizaron herramientas de inteligencia artificial (ChatGPT, OpenAI) exclusivamente como apoyo para ajustes de redacción y edición de estilo. El contenido, análisis e interpretaciones del manuscrito fueron responsabilidad exclusiva de los autores.

Fuentes de financiación: Los autores declararon que no contaron con fuentes de financiación externa para la realización de este estudio.

Referencias

- 1 Di Saverio S, Podda M, De Simone B, Ceresoli M, Augustin G, Gori A, et al. Diagnosis and treatment of acute appendicitis: 2020 update of the WSES Jerusalem guidelines. *World J Emerg Surg.* 2020;15:27. <https://doi.org/10.1186/s13017-020-00306-3>
- 2 Bottia-Córdoba S, Lacouture I, Rugeles S, Rosselli D. Epidemiología de la apendicitis aguda en Colombia: Un análisis de las bases de datos administrativas del Ministerio de Salud. *Rev Colomb Cir.* 2024;39:245-53. <https://doi.org/10.30944/20117582.2455>
- 3 Bom WJ, Scheijmans JCG, Salminen P, Boermeester MA. Diagnosis of uncomplicated and complicated appendicitis in adults. *Scand J Surg.* 2021;110:170-9. <https://doi.org/10.1177/14574969211008330>
- 4 Düzgün AP, Özel HP, Şahingöz E, Dinç T. Identification of complicated and non-complicated appendicitis: A new Alvarado-based scoring system. *Ulus Travma Acil Cerrahi Derg.* 2024;30:101-6. <https://doi.org/10.14744/tjtes.2024.70979>
- 5 Alvarado A. A practical score for the early diagnosis of acute appendicitis. *Ann Emerg Med.* 1986;15:557-64. [https://doi.org/10.1016/s0196-0644\(86\)80993-3](https://doi.org/10.1016/s0196-0644(86)80993-3)
- 6 Andersson M, Kolodziej B, Andersson RE. Validation of the Appendicitis Inflammatory Response (AIR) score. *World J Surg.* 2021;45:2181-91. <https://doi.org/10.1007/s00268-021-06042-2>
- 7 Andersson M, Andersson RE. The Appendicitis Inflammatory Response score: A tool for the diagnosis of acute appendicitis that outperforms the Alvarado score. *World J Surg.* 2008;32:1843-9. <https://doi.org/10.1007/s00268-008-9649-y>
- 8 Kularatna M, Lauti M, Haran C, MacFater W, Sheikh L, Huang Y, et al. Clinical prediction rules for appendicitis in adults: Which is best? *World J Surg.* 2017;41:1769-81. <https://doi.org/10.1007/s00268-017-3926-6>
- 9 Snyder MJ, Guthrie M, Cagle S. Acute appendicitis: Efficient diagnosis and management. *Am Fam Physician.* 2018;98:25-33.
- 10 Wray CJ, Kao LS, Millas SG, Tsao K, Ko TC. Acute appendicitis: Controversies in diagnosis and management. *Curr Probl Surg.* 2013;50:54-86. <https://doi.org/10.1067/j.cpsurg.2012.10.001>
- 11 Hernández-Cortez J, De León-Rendón JL, Martínez-Luna MS, Guzmán-Ortiz DJ, Palomeque-López A, Cruz-López N, et al. Apendicitis aguda: Revisión de la literatura. *Cir Gen.* 2019;41:33-8.
- 12 Haak F, Kollmar O, Ioannidis A, Slotta JE, Ghadimi MB, Glass T, et al. Predicting complicated appendicitis based on clinical findings: The role of Alvarado and Appendicitis Inflammatory Response scores. *Langenbecks Arch Surg.* 2022;407:2051-7. <https://doi.org/10.1007/s00423-022-02533-5>
- 13 Andersson RE, Stark J. Diagnostic value of the appendicitis inflammatory response (AIR) score. A systematic review and meta-analysis. *World J Emerg Surg.* 2025;20:12. <https://doi.org/10.1186/s13017-025-00582-x>
- 14 Jose T, Rajesh PS. Appendicitis Inflammatory Response score in comparison to Alvarado score in acute appendicitis. *Surg J (N Y).* 2021;7:e127-e131. <https://doi.org/10.1055/s-0041-1731446>
- 15 Pogorelič Z, Mihanović J, Ninčević S, Lukšić B, Baloević SE, Polašek O. Validity of Appendicitis Inflammatory Response score in distinguishing perforated from non-perforated appendicitis in children. *Children (Basel).* 2021;8:309. <https://doi.org/10.3390/children8040309>
- 16 Bukhari HA, Mirza MA. Complicated appendicitis is predicted by the presence of systemic inflammatory response syndrome. *Saudi Surg J.* 2020;8:32. https://doi.org/10.4103/ssj.ssj_6_20
- 17 Virú-Flores HM, Roque-Quezada JC, Medina-Jimenez AN, Román-Yrigoín MJ, Barturen-Julca VA, Vega-Vega RR. Capacidad del Índice de Inflamación Sistémica Inmune (IISI) en el diagnóstico de la apendicitis aguda complicada en un hospital de referencia del Perú. *Rev Colomb Cir.* 2025;40:932-41. <https://doi.org/10.30944/20117582.2847>
- 18 Departamento Administrativo Nacional de Estadística (DANE). Estimaciones y proyecciones de población por área (cabecera y resto) 1985-2035. Bogotá: DANE; 2023. Disponible en: <https://www.dane.gov.co/index.php/estadisticas-por-tema/demografia-y-poblacion/proyecciones-de-poblacion/proyecciones-de-poblacion-bogota>