



REVISTA DE
GASTROENTEROLOGÍA
DE MÉXICO

www.elsevier.es/rgmx



ARTÍCULO ORIGINAL

Comparación de la capacidad predictiva de las escalas Alvarado y AIR en el diagnóstico de apendicitis aguda: un estudio prospectivo

L.M. Sanabria-Arévalo^{a,b}, L.C. Domínguez-Torres^{a,b,*}, J. Kock^a, J.D. Lotero^b, J.C. Gómez-Cáceres^a y E. Tuta-Quintero^{a,c}

^a Facultad de Medicina, Universidad de La Sabana, Chía, Colombia

^b Departamento de Cirugía General, Clínica Universidad de La Sabana, Chía, Colombia

^c Departamento de Epidemiología, Universidad de La Sabana, Chía, Colombia

Recibido el 7 de mayo de 2024; aceptado el 15 de julio de 2024

PALABRAS CLAVE

Apendicitis aguda;
Escalas de riesgo;
Estudio observacional

Resumen

Introducción y objetivo: La apendicitis aguda destaca como una de las afecciones quirúrgicas más frecuentes. El uso de escalas de riesgo para la apendicitis aguda, como la escala de Alvarado y la AIR, muestran un buen rendimiento diagnóstico. El objetivo del artículo es comparar la capacidad predictiva de las escalas de Alvarado y AIR en el diagnóstico de apendicitis aguda.

Métodos: Estudio de corte transversal con pacientes sometidos a apendicectomía por sospecha clínica de apendicitis aguda; se confirmó el diagnóstico por histopatología. Se evaluó la capacidad predictiva de las escalas Alvarado y AIR mediante análisis de ROC determinando el área bajo la curva (curva ROC). Este estudio utilizó la lista de verificación STROBE.

Resultados: Se incluyeron 358 pacientes con sospecha clínica de apendicitis aguda, el 51% eran hombres (183/358) y la mediana de edad fue de 36 años (RIC: 24-46). La curva ROC de la escala de Alvarado fue de 0.767 (IC 95%: 0.716-0.818), y con un punto de corte de 0 a 4 mostró una sensibilidad del 78% y una especificidad del 84%. La escala AIR mostró una curva ROC de 0.741 (IC 95%: 0.691-0.788), y con un punto de corte de 0 a 4 mostró una sensibilidad del 87% y una especificidad del 56%. Al comparar ambas escalas, no se observó una diferencia significativa ($p=0.266$).

* Autor para correspondencia. Kilómetro 7 Autopista Norte. Campus Universitario Puente del Común, Facultad de Medicina, Departamento de Cirugía, Edificio H, Universidad de La Sabana, Chía, Colombia.

Correo electrónico: carlos.dominguez@unisabana.edu.co (L.C. Domínguez-Torres).

<https://doi.org/10.1016/j.rgmx.2024.07.004>

0375-0906/© 2024 Asociación Mexicana de Gastroenterología. Publicado por Masson Doyma México S.A. Este es un artículo Open Access bajo la CC BY-NC-ND licencia (<http://creativecommons.org/licencias/by-nc-nd/4.0/>).

Cómo citar este artículo: L.M. Sanabria-Arévalo, L.C. Domínguez-Torres, J. Kock et al., Comparación de la capacidad predictiva de las escalas Alvarado y AIR en el diagnóstico de apendicitis aguda: un estudio prospectivo, Revista de Gastroenterología de México, <https://doi.org/10.1016/j.rgmx.2024.07.004>

Conclusión: Las escalas de Alvarado y AIR muestran una capacidad predictiva similar para la apendicitis aguda. Los puntos de corte bajos en las escalas de riesgo se relacionan con una mayor sensibilidad diagnóstica de la enfermedad.

© 2024 Asociación Mexicana de Gastroenterología. Publicado por Masson Doyma México S.A. Este es un artículo Open Access bajo la CC BY-NC-ND licencia (<http://creativecommons.org/licencias/by-nc-nd/4.0/>).

KEYWORDS

Acute appendicitis;
Risk scales;
Observational study

Comparison of the predictive capacity of the Alvarado and AIR scores in the diagnosis of acute appendicitis: A prospective study

Abstract

Introduction and aim: Acute appendicitis stands out as one of the most frequent surgically-treated diseases. Risk scales for acute appendicitis, such as the Alvarado and AIR scoring systems, show good diagnostic yield. The aim of our study was to compare the predictive capacity between the Alvarado and AIR scores in the diagnosis of acute appendicitis.

Methods: A cross-sectional study was conducted on patients that underwent appendectomy due to suspected acute appendicitis, confirmed by histopathology. The predictive capacity of the Alvarado and AIR scores was evaluated through an ROC curve analysis, determining the area under the ROC curve. The STROBE checklist was utilized.

Results: A total of 358 patients with clinical suspicion of acute appendicitis were included, 51% of whom were men (183/358). Median patient age was 36 years (IQR: 24-46). The ROC curve of the Alvarado score was 0.767 (95% CI: 0.716-0.818), and with a cutoff point of 0-4, had 78% sensitivity and 84% specificity. The AIR score had a ROC curve of 0.741 (95% CI: 0.691-0.788), and with a 0-4 cutoff point, 87% sensitivity and 56% specificity. There was no statistically significant difference between the 2 scores ($p=0.266$).

Conclusion: The Alvarado and AIR scores have a similar predictive capacity for acute appendicitis. The low cutoff points of the risk scales are related to greater diagnostic sensitivity of the disease.

© 2024 Asociación Mexicana de Gastroenterología. Published by Masson Doyma México S.A. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introducción y objetivo

La apendicitis aguda destaca como una de las afecciones quirúrgicas más frecuentes, con una incidencia que varía entre 90 y 100 casos por cada 100,000 habitantes, presentando un riesgo del 8.6% en hombres y del 6.7% en mujeres¹. La mortalidad presenta una tasa anualizada del 1 al 4%, mientras que la morbilidad puede alcanzar hasta un 30%^{1,2}. El enfoque diagnóstico inicial de la apendicitis aguda se realiza por medio de escalas de riesgo, que integran información clínica, resultados de exámenes de laboratorio e imágenes diagnósticas para mejorar la precisión predictiva de la enfermedad. A pesar de ello, la tasa de apendicectomías sin evidencia histológica de apendicitis o falsos positivos sigue siendo superior al 6%^{1,2}.

El uso de escalas de riesgo para la apendicitis aguda como la escala de Alvarado muestra una especificidad del 86% y una sensibilidad que oscila entre el 50 y el 72%³. Datos similares se reportan para la escala *Appendicitis Inflammatory Response* (AIR), la cual incorpora niveles de proteína C reactiva para mejorar su rendimiento, alcanzando una sensibilidad del 78.4% y una especificidad del 97%^{4,5}. Sin embargo, es limitada la evidencia médica que compare ambas escalas de riesgo^{6,7}. Por lo tanto, este estudio tiene como objetivo comparar respectivamente las escalas de Alvarado y AIR con

el análisis histopatológico para determinar sus respectivas precisiones en el diagnóstico de apendicitis aguda.

Métodos

Se llevó a cabo un estudio de corte transversal en un solo centro de atención de tercer nivel en Colombia, con el objetivo de evaluar la capacidad predictiva de las escalas de Alvarado y AIR para la apendicitis aguda, durante un período de estudio de 10 meses, desde noviembre de 2020 hasta septiembre de 2021. Este estudio utilizó la lista de verificación STROBE.

Criterios de elegibilidad

Los criterios de inclusión fueron pacientes mayores de 18 años, con sospecha clínica de apendicitis aguda y con un tiempo de evolución menor de 7 días desde el inicio de los síntomas. Se excluyeron del estudio los pacientes con otras enfermedades inflamatorias/infecciosas intraabdominales, así como aquellos sin informes histopatológicos del espécimen extirpado y las personas que experimentaban dolor abdominal crónico. Todos los datos fueron recuperados ana-

lizando la historia clínica de cada paciente y registrados en una base de datos segura.

Utilizamos una proporción de individuos sanos del 80%, una proporción de individuos enfermos del 20%, una especificidad y sensibilidad del 80%, y un margen de error alfa de 1.96 para calcular un tamaño de muestra estimado de 325 pacientes⁸.

Variables clínicas

Se calcularon tanto la escala de Alvarado como el AIR score. El manejo clínico y quirúrgico se llevó a cabo de acuerdo con el riesgo individual. Se realizó una tomografía computarizada o ultrasonografía para los pacientes con riesgo intermedio, y los pacientes con alto riesgo se sometieron a intervención quirúrgica de urgencia. El apéndice extirpado se envió para evaluación histopatológica con el fin de confirmar el diagnóstico basado en los hallazgos microscópicos para los pacientes que se sometieron a intervención quirúrgica. Consideramos un diagnóstico positivo con el análisis histopatológico final. La decisión de operar fue tomada por el cirujano.

Análisis estadístico

Todos los análisis se realizaron utilizando el programa Stata versión 16 (StataCorp LLC, College Station, EE. UU.). Los datos se obtuvieron automáticamente a través de un formulario de recogida electrónico, que posteriormente se ingresó en una hoja de cálculo Excel para su verificación por parte del grupo de investigación con el fin de identificar cualquier error de transcripción y corrección. Las variables cualitativas se informaron en frecuencias y porcentajes, mientras que las variables cuantitativas se resumieron utilizando la media y la desviación estándar o la mediana y el rango intercuartílico, según su distribución. Se calcularon la sensibilidad, la especificidad, el valor predictivo positivo (VPP) y el valor predictivo negativo (VPN) para las escalas de Alvarado y AIR en la identificación de apendicitis aguda confirmada por histopatología. Además, la validez predictiva de las escalas se evaluó mediante análisis de ROC determinando el área bajo la curva (curva ROC) de las escalas con un intervalo de confianza del 95% (IC 95%). Las curvas ROC calculadas se compararon utilizando la prueba DeLong⁹.

La interpretación de la curva ROC fue la siguiente: 0.50 indicaba una ausencia de capacidad discriminatoria; 0.51 a 0.60 indicaba casi ninguna capacidad discriminatoria; 0.61 a 0.69 indicaba una capacidad discriminatoria pobre; > 0.7 a 0.8 indicaba una capacidad discriminatoria aceptable; > 0.80 a 0.90 indicaba una capacidad discriminatoria excelente, y > 0.90 indicaba una capacidad discriminatoria sobresaliente⁹. La significación estadística se consideró cuando $p < 0.05$.

Consideraciones éticas

Nuestra investigación cumple con las regulaciones actuales sobre investigación bioética y fue autorizada por el Comité de Ética de la Facultad de Medicina de la Universidad de La Sabana (Código: MEDEsp-41-2020). Los autores decla-

Tabla 1 Análisis demográfico de la población

| | |
|--|------------|
| <i>Pacientes, n (%)</i> | 358 (100) |
| <i>Edad en años, mediana (RIC)</i> | 33 (24-46) |
| <i>Sexo masculino, n (%)</i> | 183 (51) |
| <i>Escala de Alvarado, n (%)</i> | |
| Riesgo leve (0-4 puntos) | 63 (18) |
| Riesgo intermedio (5-6 puntos) | 96 (27) |
| Alto riesgo (7-10 puntos) | 199 (56) |
| <i>Escala AIR, n (%)</i> | |
| Riesgo leve (0-4 puntos) | 91 (25) |
| Riesgo intermedio (5-8 puntos) | 198 (55) |
| Alto riesgo (9-12 puntos) | 69 (19) |
| <i>Hallazgos imagenológicos de apendicitis^a, n (%)</i> | |
| No | 86 (43) |
| Sí | 116 (57) |
| <i>Intervención quirúrgica, n (%)</i> | |
| No | 92 (26) |
| Sí | 266 (74) |
| <i>Confirmación de apendicitis por histopatología^b, n (%)</i> | |
| No | 97 (28) |
| Sí | 246 (72) |

AIR: *Appendicitis Inflammatory Response*; RIC: rango intercuartílico.

^a 202 pacientes.

^b 343 pacientes.

ran que este artículo no contiene información personal que pueda identificar a los participantes de nuestra investigación. Todos los pacientes incluidos aceptaron y firmaron el consentimiento informado, dada la naturaleza observacional prospectiva de este estudio.

Resultados

Se incluyeron 358 pacientes con sospecha clínica de apendicitis aguda, el 51% eran hombres (183/358) y la mediana de edad fue de 36 años (rango intercuartílico: 24-46). Las características generales de la población se describen en la [tabla 1](#).

La curva ROC de la escala de Alvarado fue de 0.767 (IC 95%: 0.716-0.818) para predecir el diagnóstico de apendicitis aguda ([fig. 1](#)). Para un riesgo bajo o un punto de corte de 0 a 4, se demostró una sensibilidad del 78%, una especificidad del 84%, un VPP del 95% y un VPN del 49%. En el caso de un riesgo alto o un punto de corte de 7 a 10, hay una sensibilidad del 71%, una especificidad del 83%, un VPP del 92% y un VPN del 48%.

La escala AIR mostró una curva ROC de 0.741 (IC 95%: 0.691-0.788) ([fig. 2](#)). Para un riesgo bajo o un punto de corte de 0 a 4, la sensibilidad es del 87%, la especificidad es del 56%, el VPP es del 61% y el VPN es del 16%. Para un riesgo alto o un punto de corte de 9 a 12, la sensibilidad es del 27%, la especificidad es del 97%, el VPP es del 96% y el VPN es del 67%.

Al comparar las 2 curvas ROC, no se observó una diferencia significativa entre las escalas de Alvarado y AIR ($p = 0.266$) ([fig. 3](#)).

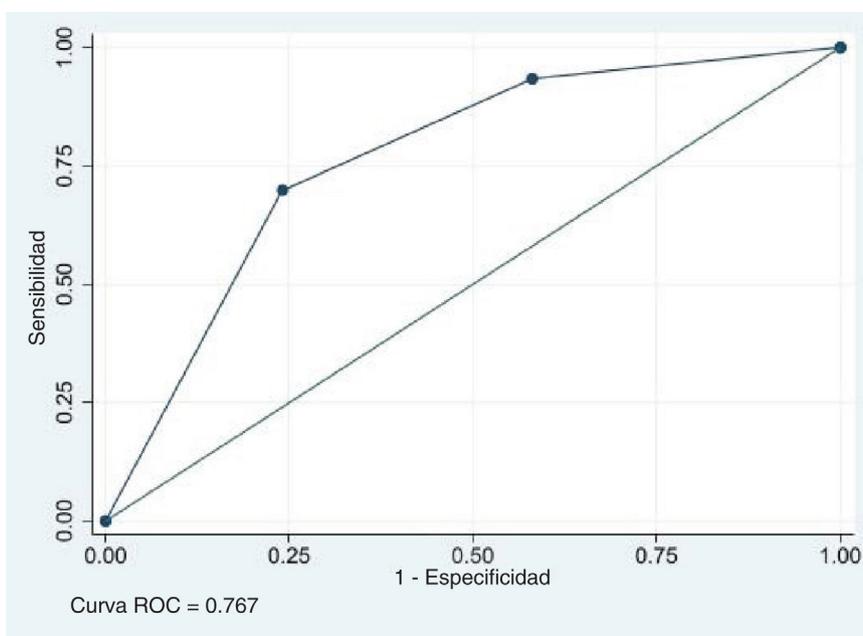


Figura 1 Rendimiento diagnóstico para la escala de Alvarado.

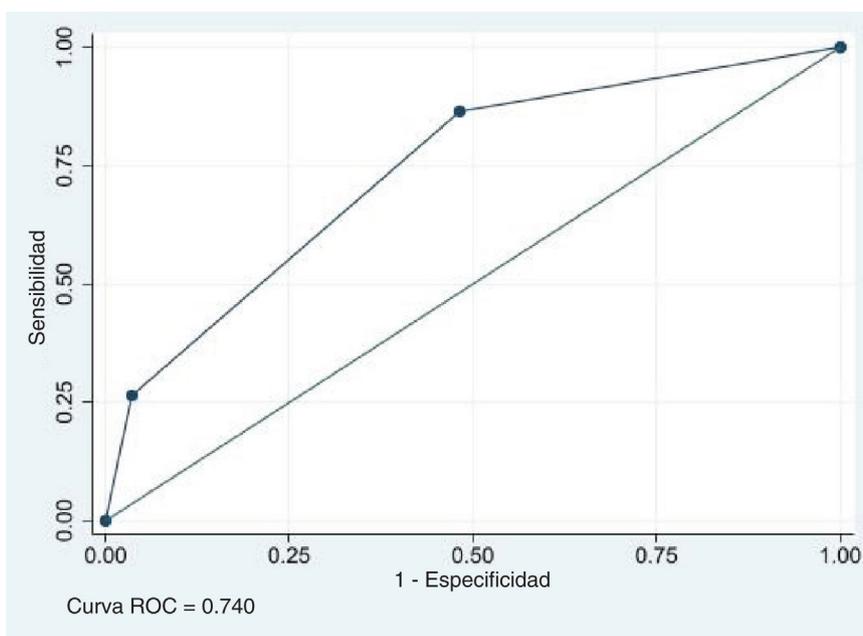


Figura 2 Rendimiento diagnóstico para la escala AIR.

Discusión

El diagnóstico de la apendicitis aguda representa un desafío. Se han desarrollado diversas escalas diagnósticas para evaluar el riesgo de presentar esta dolencia en el contexto de dolor abdominal agudo. Dentro de estas herramientas se encuentran las escalas de Alvarado y la AIR, las cuales tienen similitudes en su capacidad predictiva para la enfermedad^{4,5,10-14}. Los puntos de corte bajos en las escalas de riesgo se relacionan con una mayor sensibilidad diagnóstica de la enfermedad. En el presente estudio, la escala

AIR mostró una mayor especificidad y VPP frente a la escala de Alvarado con un punto de corte de riesgo alto. Además, en casos donde el riesgo de padecer apendicitis aguda es alto, el VPP se incrementó considerablemente, brindando una mayor confianza en el diagnóstico.

Estos hallazgos pueden discutirse a la luz de la evidencia disponible. Kollár et al.⁵ compararon la capacidad discriminadora de la escala AIR frente a la escala de Alvarado y la impresión clínica de un cirujano experimentado. Los resultados mostraron que las escalas de riesgo y el juicio clínico de un cirujano experto tienen una similar precisión

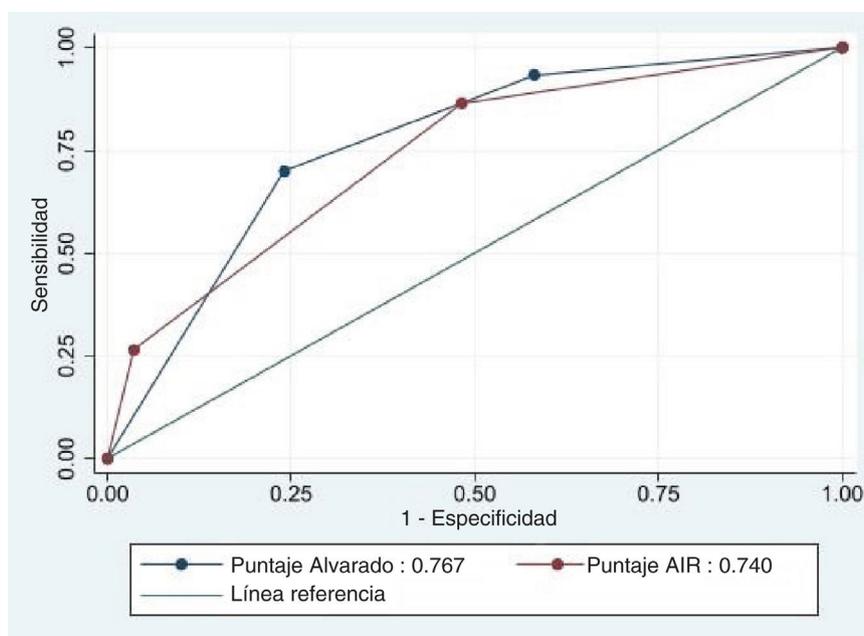


Figura 3 Comparación de los rendimientos diagnósticos para la escala de Alvarado y AIR.

para descartar apendicitis. Además, la escala AIR demuestra una excelente especificidad en pacientes con alto riesgo de presentar apendicitis aguda, superando tanto el juicio clínico como la escala de Alvarado, resultados que concuerdan con los hallazgos de nuestro estudio¹⁰⁻¹².

En estudios observacionales, incluido el nuestro, se ha observado una baja sensibilidad en pacientes con alto riesgo de apendicitis aguda según la escala AIR. Esto se debe a que una proporción significativa de pacientes son estratificados en el grupo de probabilidad media¹³. Por lo tanto, se hace necesario recurrir al uso de imágenes diagnósticas para confirmar y respaldar el diagnóstico clínico. La utilización de escalas de riesgo puede ser beneficiosa en la aproximación diagnóstica de la apendicitis aguda. No obstante, es fundamental emplearlos de manera complementaria a la evaluación clínica experta, el examen físico y las imágenes diagnósticas^{13,14}.

Aunque nuestro estudio no abarca el análisis de imágenes diagnósticas como resultado clínico, los hallazgos respaldan lo descrito en la literatura, donde se destaca que los VPP de ambas escalas son más precisos en pacientes clasificados como de alto riesgo de apendicitis aguda^{14,15}. Noori et al.⁴ compararon la capacidad discriminadora de las escalas AIR y Alvarado frente a la tomografía computarizada. Puntos de corte ≥ 4 a ≤ 6 en la escala de Alvarado mostraron una sensibilidad, una especificidad y unos valores predictivos menores frente a la imagen diagnóstica. Sin embargo, la tomografía computarizada tuvo un rendimiento diagnóstico comparable frente a puntuaciones de Alvarado y AIR de ≤ 4 y ≥ 7 , respectivamente.

Una limitación de este estudio fue la realización en un único centro y la exclusión de poblaciones pediátricas y gestantes, lo que puede representar un obstáculo para extrapolar nuestros resultados a la identificación de esta afección en dichos grupos poblacionales. La recopilación prospectiva de información clínica y el tamaño de la muestra

representativo para la población nos han permitido obtener resultados sólidos. Creemos que nuestros hallazgos pueden contribuir significativamente al enfoque y diagnóstico de la apendicitis aguda. Sin embargo, reconocemos la importancia de llevar a cabo estudios adicionales, así como de comparar nuestra escala con otras ya existentes para esta dolencia, con el fin de ampliar y fortalecer nuestros resultados¹²⁻¹⁶.

Conclusión

Las escalas de Alvarado y AIR muestran una capacidad predictiva similar para la apendicitis aguda. Sin embargo, la escala AIR demostró una mayor especificidad y VPP en comparación con la escala de Alvarado cuando se utilizó un punto de corte de riesgo alto. La utilización de estas escalas puede resultar útil en el servicio de urgencias para la identificación de esta afección.

Financiación

Este proyecto contó con la financiación de la Universidad de La Sabana (Código: MEDEsp-41-2020).

Contribución de los autores

Todos los autores participaron en la concepción y diseño del estudio, la adquisición de datos, el análisis e interpretación y la revisión de los resultados, la redacción y la revisión final del artículo.

Conflicto de intereses

No hay conflicto de intereses que impida la publicación de este trabajo.

Agradecimientos

A la Universidad de La Sabana.

Referencias

1. Snyder MJ, Guthrie M, Cagle S. Acute appendicitis: Efficient diagnosis and management. *Am Fam Physician*. 2018;98:25–33.
2. Bhangu A, Søreide K, di Saverio S, et al. Acute appendicitis: Modern understanding of pathogenesis, diagnosis, and management. *Lancet*. 2015;386:1278–87, [http://dx.doi.org/10.1016/S0140-6736\(15\)00275-5](http://dx.doi.org/10.1016/S0140-6736(15)00275-5).
3. Jose T, Rajesh PS. Appendicitis inflammatory response score in comparison to Alvarado score in acute appendicitis. *Surg J (NY)*. 2021;7:e127–31, <http://dx.doi.org/10.1055/s-0041-1731446>.
4. Noori IF, Jabbar AS, Noori AF. Clinical scores (Alvarado and AIR scores) versus imaging (ultrasound and CT scan) in the diagnosis of equivocal cases of acute appendicitis: A randomized controlled study. *Ann Med Surg (Lond)*. 2023;85:676–83, <http://dx.doi.org/10.1097/MS9.0000000000000270>.
5. Kollár D, McCartan DP, Bourke M, et al. Predicting acute appendicitis? A comparison of the Alvarado score, the Appendicitis Inflammatory Response Score and clinical assessment. *World J Surg*. 2015;39:104–9, <http://dx.doi.org/10.1007/s00268-014-2794-6>.
6. Gupta V, Gupta P, Gill CS, et al. Appendicitis inflammatory response score in acute appendicitis: A study at a tertiary care center in north India. *Int J Appl Basic Med Res*. 2022;12:234–8, <http://dx.doi.org/10.4103/ijabmr.ijabmr.287.22>.
7. Capoglu R, Gonullu E, Bayhan Z, et al. Comparison of scoring systems regarding the gender as a parameter with the traditional scoring systems for predicting appendicitis. *Updates Surg*. 2022;74:1035–42, <http://dx.doi.org/10.1007/s13304-022-01272-y>.
8. Machin D, Campbell MJ, Tan SB, et al. Sample size tables for clinical studies, 3rd edition. Oxford: Wiley Blackwell; 2009. p. 264.
9. Hosmer DW, Lemeshow S, Sturdivant RX. Special topics. En: Hosmer DW, Lemeshow S, Sturdivant RX, editores. *Applied logistic regression*. 3rd ed. New York, NY: John Wiley & Sons, Inc.; 2013. p. 401–8.
10. Al-Tarakji M, Zarour A, Singh R, et al. The role of Alvarado score in predicting acute appendicitis and its severity in correlation to histopathology: A retrospective study in a Qatar population. *Cureus*. 2022;14:e26902, <http://dx.doi.org/10.7759/cureus.26902>.
11. Karki OB, Hazra NK. Evaluation of the Appendicitis Inflammatory Response Score against Alvarado score in diagnosis of acute appendicitis. *Kathmandu Univ Med J (KUMJ)*. 2020;18:171–5.
12. Ebell MH, Shinholser J. What are the most clinically useful cutoffs for the Alvarado and Pediatric Appendicitis Scores? A systematic review. *Ann Emerg Med*. 2014;64:365–72.e2, <http://dx.doi.org/10.1016/j.annemergmed.2014.02.025>.
13. Andersson RE. The magic of an appendicitis score. *World J Surg*. 2015;39:110–1, <http://dx.doi.org/10.1007/s00268-014-2821-7>.
14. Chisthi MM, Surendran A, Narayanan JT. RIPASA and air scoring systems are superior to Alvarado scoring in acute appendicitis: Diagnostic accuracy study. *Ann Med Surg (Lond)*. 2020;59:138–42, <http://dx.doi.org/10.1016/j.amsu.2020.09.029>.
15. Naeem MT, Jamil MA, Anwar MI, et al. Diagnostic accuracy of Alvarado scoring system relative to histopathological diagnosis for acute appendicitis: A retrospective cohort study. *Ann Med Surg (Lond)*. 2022;81:104561, <http://dx.doi.org/10.1016/j.amsu.2022.104561>.
16. Ghali MS, Hasan S, Al-Yahri O, et al. Adult appendicitis score versus Alvarado score: A comparative study in the diagnosis of acute appendicitis. *Surg Open Sci*. 2023;14:96–102, <http://dx.doi.org/10.1016/j.sopen.2023.07.007>.