is so wide (95% CI = 0.2 -0.78) that the authors calculated an uncertainty of 0.4 (40%).

6. The sample is poorly representative of the target population of the RIPASA score, since it includes more than 85% of positive diagnoses, a proportion that magnifies the incidence observed in patients visiting the emergency departments with "right iliac fossa pain" (Chong reported 52.60% in his prospective study)[2], producing a distortion of the predictive values of the test, which are essential to estimate its performance in the healthcare scenario.
7. Newman et al.[3], among many other authors, ban the estimation of predictive values in non-consecutive samples.
8. The evaluation of patient records with a known final diagnosis by a single observer involves potential interpretation bias. To reduce the risk, a minimum of two independent blinded assessors are recommended with estimation of interobserver agreement[4].
9. The conclusion is typical of a review of the literature rather than an analysis of primary data, as it is based on studies by other authors.

The assessment of the performance of a score for the diagnosis of acute appendicitis offers opportunities for research that, unfortunately, were not considered:

1. There are data from previous studies to calculate, with reasonable approximation, the number of patients to obtain appropriate estimates per interval.
2. The disease is common, which would allow to reach the required sample size in a short recruitment period.
3. There is no need for prolonged follow-up, which ensures few dropouts and facilitates prospective studies.

Daniel E. Tripoloni
dtripoloni@hotmail.com
División "A" de Cirugía General. Hospital José María Ramos Mejía. Ciudad Autónoma de Buenos Aires. Argentina

Referencias bibliográficas | Reference

1. Gasque RA, Moreno WA, Vigilante GE. Correlación diagnóstica de la escala RIPASA en pacientes intervenidos quirúrgicamente por apendicitis aguda. Rev Argent Cirug. 2020; 112:178-84.
2. Chong CF, Thien A, Mackie AJ, Tin AS, Tripathi S, Ahmad MA, et al. Comparison of RIPASA and Alvarado scores for the diagnosis of acute appendicitis. Singapore Med J. 2011;52:340-5.
3. Newman TB, Browner WS, Cummings SR, Hulley SB. Diseño de estudios de pruebas médicas. En: Hulley SB, Cummings SR, Browner WS, Grady DG, Newman TB. Diseño de investigaciones clínicas. 4ª ed. Barcelona: Wolters Kluwer; 2014. Pp.171-91.
4. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas. 1960; 20:37–46.

Respuesta de los autores

Sr. Director:

Agradecemos al Dr. Tripoloni por haberse tomado el tiempo de leer nuestro artículo "Correlación diagnóstica de la escala RIPASA en pacientes intervenidos quirúrgicamente por apendicitis aguda" y habernos enviado sus valiosas apreciaciones. En respuesta a sus correcciones:

- Si bien es cierto que no fue calculado el tamaño muestral, decidimos incluir la mayor cantidad de pacientes en la muestra debido a que nuestro centro atiende principalmente adultos mayores (con un promedio de 63 años); por ende, no contamos con un número elevado de casos de apendicitis aguda.
- Para todas las pruebas estadísticas se utilizó una $p \leq 0,05$ (IC 95%) como estadísticamente significativa y se calcularon con las plataformas MedCalc v. 19.1.3® y Microsoft Excel 2016®. Utilizamos la prueba "t" de Student para contrastar la duración del cuadro clínico entre los pacientes con apendicitis aguda y sin ella en su estudio histopatológico.
- La desviación estándar de la duración del cuadro clínico para pacientes con apendicitis aguda fue de 8,81±3,41 horas y para aquellos sin apendicitis aguda fue de 7,95±4,32 horas.
- Una de las principales falencias de este estudio es su naturaleza retrospectiva y el hecho de que nos hayamos basado en las muestras de anatomía patológica para el diagnóstico de certeza. Igualmente, este trabajo fue presentado en el 89° Congreso Argentino de Cirugía y una de las críticas del jurado fue que la escala no fue aplicada de manera regular a todos los pacientes que se presentaron por guardia con "síndrome de fosa ilíaca derecha" como usted remarca en el punto 6; por ende, queda trunca su potencial función discriminatoria.
- Las historias clínicas electrónicas fueron analizadas por el autor principal del trabajo. Consideramos que el principal sesgo fue la subjetividad de cada cirujano tratante a la hora de la evaluación inicial del paciente. Este podría reducirse –como dijo usted– utilizando más de un evaluador y de manera prospectiva.
- La conclusión responde a los resultados y objetivos del estudio y no así a la discusión como usted propone.

Nuevamente, reiteramos la necesidad de realizar este estudio de manera prospectiva y con un tamaño muestral adecuado para la patología a fin de validar su utilidad en nuestro medio.

Rodrigo A. Gasque
Walter A. Moreno
Gabriel E. Vigilante

■ ENGLISH VERSION

Dear Sir,

We thank Dr. Tripoloni for taking the time to read our article "Diagnostic correlation of the RIPASA score in patients operated on for acute appendicitis" and for sending us his valuable observations. The responses to his corrections are listed below:

- Although the sample size was not calculated, we decided to include the largest number of patients because our center mainly treats older adults (with a mean age of 63 years) and therefore, we do not have a high number of cases of acute appendicitis.
- A p value ≤ 0.05 (95% CI) was considered statistically significant and all the calculations were performed using MedCalc version 19.1.3® statistical software package and Microsoft Excel 2016®. We used the Student's t test to compare the duration of symptoms between the patients with and without acute appendicitis and without appendicitis in the pathology report.
- The standard deviation of the duration of symptoms was 8.81 ± 3.41 hours for patients with acute appendicitis and 7.95 ± 4.32 for those without acute appendicitis.
- One of the main shortcomings of this study is its retrospective nature and the fact that we relied on pathological anatomy samples for the definitive diagnosis. This paper was presented at the 89th Argentine Congress of Surgery and the judging panel criticized the fact that the score was not systematically applied to all the patients presenting at the emergency department with "right iliac fossa pain" as you indicated in the point 6; therefore, the potential discriminating function of the score was not achieved.
- The electronic clinical records were analyzed by the lead author of the paper. We consider that the subjectivity of each treating surgeon at the time of the initial evaluation of the patient was the main bias. This bias could be reduced - as you said - by using more than one assessor in a prospective fashion.
- The conclusion is based on the results and objectives of the study and not to the discussion as you propose.

Again, we emphasize the need for a prospective study with a sample size appropriate to the disease to validate its usefulness in our environment.

Rodrigo A. Gasque
Walter A. Moreno
Gabriel E. Vigilante