

Desempeño comparativo de la escala de Alvarado y la escala de apendicitis pediátrica en escolares y adolescentes

Comparative performance of Alvarado scale and the pediatric appendicitis scale in schoolchildren and adolescents

Alejandro Ramírez Guirado^{1*} <https://orcid.org/0000-0003-0358-6062>

Ana Barbarita Navarro Sombert¹ <https://orcid.org/0000-0001-9843-2087>

Ramón Villamil Martínez² <https://orcid.org/0000-0002-3544-9604>

¹Hospital Pediátrico Docente Centro Habana. La Habana, Cuba.

²Hospital Pediátrico Universitario William Soler. La Habana, Cuba.

*Autor para correspondencia: alexramguirado@gmail.com

RESUMEN

Introducción: La escala de Alvarado y la escala de apendicitis pediátrica se utilizan para la estratificación de pacientes pediátricos con sospecha de apendicitis.

Objetivo: Comparar el desempeño diagnóstico de estas escalas en escolares y adolescentes, en general y según edad y sexo.

Métodos: Estudio observacional prospectivo de 452 pacientes entre 5 y 18 años de edad, ingresados en el Hospital Pediátrico Docente Centro Habana, por dolor abdominal o apendicitis aguda, entre 2016 y 2017.

Resultados: La escala de Alvarado tuvo a los <4 puntos: sensibilidad 99,14 %, razón de verosimilitud negativa 0,02; a los >7 puntos: especificidad 96,15 %, razón de verosimilitud positiva 10,3. La escala de apendicitis pediátrica presentó a los <5 puntos: sensibilidad 96,55 %, razón de verosimilitud negativa 0,07; a los >8 puntos: especificidad 96,15 %, razón de verosimilitud positiva 5,60. Las áreas bajo la curva fueron similares en general (0,851-0,858), siendo más altas y similares entre sí en escolares masculinos (0,918-0,923). En adolescentes femeninas, el área bajo la curva para la escala de apendicitis pediátrica (0,841) fue mayor que para la de Alvarado (0,802).

Conclusiones: La escala de Alvarado resulta más específica, mientras que la escala de apendicitis pediátrica es más sensible. Se deben considerar diferentes puntos de corte para cada escala al estratificar según probable apendicitis. El desempeño discriminativo se comporta similar en general para ambas escalas, y muy bien en escolares masculinos. La mayor diferencia en desempeño ocurre en adolescentes femeninas, a favor de la escala de apendicitis pediátrica.

Palabras clave: apendicitis; niños; escala de Alvarado; escala de apendicitis pediátrica.

ABSTRACT

Introduction: Alvarado scale and the pediatric appendicitis scale are used for the stratification of pediatric patients with suspected appendicitis.

Objective: Compare the diagnostic performance of these scales in schoolchildren and adolescents, in general and according to age and sex.

Methods: Prospective observational study of 452 patients between 5 and 18 years of age, admitted to Centro Habana Pediatric Teaching Hospital due to abdominal pain or acute appendicitis in the period 2016-2017.

Results: Alvarado scale had at <4 points: sensitivity 99.14 %, negative likelihood ratio 0.02; at >7 points: specificity 96.15 %, positive likelihood ratio 10.3. The paediatric appendicitis scale presented at <5 points: sensitivity 96.55 %, negative likelihood ratio 0.07; at >8 points: specificity 96.15 %, positive likelihood ratio 5.60. The areas under the curve had a similar overall (0.851 - 0.858), being higher and similar to each other in male schoolchildren (0.918 - 0.923). In female adolescents, the area under the curve for the pediatric appendicitis scale (0.841) was greater than for Alvarado's (0.802).

Conclusions: Alvarado scale is more specific, while the pediatric appendicitis scale is more sensitive. Different cut-off points should be considered for each scale when stratifying according to probable appendicitis. Discriminative performance behaves similar in general for both scales, and very well in male schoolchildren. The biggest difference in performance occurs in female adolescents, in favor of the pediatric appendicitis scale.

Keywords: appendicitis; children; Alvarado scale; pediatric appendicitis scale.

Recibido: 28/01/2021

Aceptado: 12/10/2021

Introducción

El diagnóstico de la apendicitis aguda en niños aún presenta desafíos. A pesar del refinamiento de las herramientas diagnósticas disponibles, se siguen realizando apendicectomías negativas, y cierta cantidad de casos tienen un diagnóstico tardío que resulta en más complicaciones y costos.

Varias escalas de predicción clínica se han utilizado para optimizar este diagnóstico en la población pediátrica. Algunas se desarrollaron para pacientes tanto adultos como pediátricos, como la escala de Alvarado (EA),⁽¹⁾ y otras se crearon específicamente para niños, como la escala de apendicitis pediátrica (EAP).⁽²⁾ Una revisión sistemática de las escalas diagnósticas para niños con posible apendicitis concluyó que estas dos fueron las mejor validadas.⁽³⁾ Ambas tienen un puntaje total posible de 10 puntos. De acuerdo a sus autores, la EA predice apendicitis con ≥ 7 puntos, mientras que la EAP lo hace con ≥ 6 puntos.^(1,2) Los componentes y sus valores se presentan en la tabla 1.

Tabla 1 - Componentes y puntaje para la escala de Alvarado y la escala de apendicitis pediátrica

Componente clínico o de laboratorio	EA*	EAP**
Migración del dolor	1	1
Anorexia	1	1
Náuseas/Vómitos	1	1
Dolor a la palpación en fosa ilíaca derecha	2	2
Dolor a la descompresión	1	-
Dolor en fosa ilíaca derecha al golpe de tos/salto/percusión	-	2
Temperatura elevada ($\geq 37,3$ °C)	1	1
Leucocitosis ($\geq 10\,000/\mu\text{l}$)	2	1
Desviación izquierda (neutrófilos ≥ 75 %)	1	1
Total	10	10

EA: escala de Alvarado; EAP: escala de apendicitis pediátrica.

*Adaptado de Alvarado;⁽¹⁾ **Adaptado de Samuel.⁽²⁾

Varios estudios brindan comparaciones entre estas escalas. En algunos, la EAP resulta la más efectiva,^(3,4,5) mientras que para otros, la EA tiene mejor desempeño,^(6,7,8) sin embargo, el consenso general es que ninguna por sí sola proporciona un diagnóstico positivo o negativo de apendicitis suficientemente certero.^(4,9,10,11) Su utilidad sí está probada, no obstante, para estratificar pacientes con dolor abdominal según su riesgo de apendicitis, por ejemplo, para enviar a casa a aquellos con riesgo bajo, observar o indicar estudios de imágenes para los de riesgo

intermedio, y para consultar con el cirujano en casos con riesgo elevado, con la ventaja potencial de disminuir la exposición a radiaciones y los costos.^(9,11,12,13,14,15,16,17,18,19,20)

Estas escalas, debido a su utilidad, se han integrado a algoritmos de decisión y a vías u organigramas clínico-radiológicos para niños con dolor abdominal en los servicios de urgencia pediátrica. La EAP^(15,16,17,18,19,20,21,22,23,24,25) se ha empleado más que la EA^(12,13,24,25,26) en este sentido, sin embargo, una encuesta reveló que la EA era más utilizada que la EAP para estratificar según riesgo a las pacientes pediátricas femeninas.⁽²⁷⁾ Las estrategias de conducta para la apendicitis pediátrica en general, probablemente incluirán una estratificación de riesgo ajustada a la edad y al sexo.⁽¹⁹⁾

La interpretación de la EA ha probado ser diferente en hombres, mujeres, niños y adolescentes,^(28,29,30) mientras que la EAP se ha descrito como más efectiva en adolescentes que en niños de edad escolar⁽³¹⁾ y más específica en niñas que en niños.^(23,32) Ninguna de las escalas se recomienda para niños menores de 5 años.^(33,34) Un estudio comparó la EA y la EAP en general y de acuerdo al sexo, pero no tomó en cuenta la edad.⁽⁴⁾ Parece relevante, por tanto, comparar el desempeño diagnóstico de estas escalas en niños de edad escolar y adolescentes, en general y según edad y sexo, lo cual constituye el objetivo de esta investigación.

Métodos

Se realizó un estudio observacional prospectivo en el Hospital Pediátrico Docente Centro Habana, entre enero de 2016 y enero de 2017. La población estudiada fueron los pacientes entre 5 y 18 años de edad, ingresados en el Servicio de Cirugía Pediátrica del citado hospital con diagnóstico presuntivo de apendicitis aguda o dolor abdominal. Se excluyeron los pacientes con dolor abdominal por más de 72 horas, los previamente apendicectomizados, y los diagnosticados con plastrón apendicular.

La EA y la EAP se calcularon al momento del ingreso. Los pacientes se dividieron, de forma independiente para cada escala, en dos grupos: apendicitis y no apendicitis, considerando el diagnóstico histopatológico como elemento discriminante. Estos dos grupos se subdividieron de la siguiente manera:

Apendicitis:

- $EA \geq 7 / EAP \geq 6$, que fueron operados, con diagnóstico histopatológico de apendicitis aguda (verdaderos positivos).
- $EA \leq 6 / EAP \leq 5$, que fueron operados, con diagnóstico histopatológico de apendicitis aguda, o que no fueron operados pero reingresaron durante las siguientes dos semanas con diagnóstico de apendicitis aguda o plastrón apendicular (falsos negativos).

No apendicitis:

- $EA \leq 6 / EAP \leq 5$, que no fueron operados, y no reingresaron durante las siguientes dos semanas con diagnóstico de apendicitis aguda o plastrón apendicular, o que fueron operados, con diagnóstico histopatológico distinto de apendicitis aguda (verdaderos negativos).
- $EA \geq 7 / EAP \geq 6$, que no fueron operados y no reingresaron durante las siguientes dos semanas con diagnóstico de apendicitis aguda o plastrón apendicular, o que fueron operados, con diagnóstico histopatológico distinto de apendicitis aguda (falsos positivos).

Se calcularon los siguientes parámetros: (VP: verdaderos positivos; FN: falsos negativos; VN: verdaderos negativos; FP: falsos positivos):

Sensibilidad: probabilidad de que la escala arroje un resultado positivo en los enfermos con apendicitis aguda:

$$VP / (VP + FN)$$

Especificidad: probabilidad de que la escala arroje un resultado negativo en los casos sin apendicitis aguda:

$$VN / (VN + FP)$$

Valor predictivo positivo: probabilidad de presentar apendicitis aguda si la escala ha predicho un diagnóstico positivo:

$$VP / (VP + FP)$$

Valor predictivo negativo: probabilidad de no presentar apendicitis aguda si la escala ha predicho un resultado negativo:

$$VN / (VN + FN)$$

Razón de verosimilitud positiva: cuánto mejora la escala la probabilidad de hacer un diagnóstico correcto positivo:

$$\text{Sensibilidad} / (1 - \text{especificidad})$$

Se interpreta su resultado como utilidad: >10 alta; 5-10 moderada; 2-4,9 baja; 1-1,9 muy baja.

Razón de verosimilitud negativa cuánto mejora la escala la probabilidad de hacer un diagnóstico correcto negativo:

$$(1 - \text{sensibilidad}) / \text{especificidad}$$

Se interpreta su resultado como utilidad: <0,1 alta; 0,1-0,2 moderada; 0,21-0,5 baja; 0,51-1,0 muy baja.

Desempeño discriminativo general de la escala: determinado por el área bajo la curva de características operativas del receptor (COR). Un área de 0,50 representa la distribución aleatoria y un valor de 1,0 la predicción perfecta.

La información concerniente a los datos clínicos y de laboratorio al ingreso se obtuvo directamente de los pacientes en el cuerpo de guardia (por parte de varios observadores; la coincidencia de los datos clínicos y de laboratorio con lo registrado en la historia clínica fue controlada sistemáticamente por el primer autor). La información registrada durante la estadía hospitalaria del paciente, así como el diagnóstico histológico, se obtuvo de las historias clínicas y los registros de anatomía patológica. Esta información se recogió en una base de datos del programa SPSS 25.0 (IBM Co. EE. UU.), que también se utilizó para el análisis estadístico, junto con el programa MedCalc19.1.3 (MedCalc Software, Bélgica).

Las variables utilizadas fueron: edad, escala de Alvarado, escala de apendicitis pediátrica, sexo, grupo de edad, migración del dolor, anorexia, náusea/vómitos, dolor a la palpación en fosa ilíaca derecha, dolor a la descompresión, dolor en fosa ilíaca derecha al golpe de tos/salto/percusión, temperatura elevada, leucocitosis, neutrófilos.

Se utilizó la prueba de *ji*-cuadrada (χ^2) para la comparación de variables cualitativas, y se evaluaron las diferencias entre las variables cuantitativas mediante la prueba no paramétrica U de Mann-Whitney. La medida de acuerdo entre las escalas se midió con el estadístico Kappa. Las diferencias entre las áreas bajo la curva de COR se determinaron mediante comparación pareada (método DeLong). Se calcularon intervalos de confianza de 95 %, y las diferencias entre los grupos se consideraron significativas con un valor de $p < 0,05$.

Se observaron los principios éticos para la investigación médica de la Declaración de Helsinki.⁽³⁵⁾ El protocolo de investigación fue aprobado por el Comité de Ética de la institución. Los datos de los pacientes se recogieron anónimamente. Los casos se ingresaron, diagnosticaron y trataron de acuerdo con el juicio clínico de los especialistas y residentes del servicio de cirugía pediátrica, independientemente de los datos de la investigación y los puntajes de las escalas calculadas.

Resultados

Se calcularon la EA y la EAP para 452 pacientes con diagnóstico al ingreso de dolor abdominal o apendicitis aguda, de los cuales 348 (77 %) tuvieron diagnóstico histopatológico confirmado de apendicitis. Las variables demográficas y de los elementos de las escalas, en general y de acuerdo al diagnóstico de apendicitis, se resumen en la tabla 2.

Tabla 2 - Características de los pacientes en general y según diagnóstico positivo o negativo de apendicitis

Variables	General (n=452)	Apendicitis (n=348)	No apendicitis (n=104)	p
Cuantitativas (rango)	Media (DT)	Media (DT)	Media (DT)	
Edad (5-18 años)	12,41 (3,61)	12,26 (3,62)	12,91 (3,56)	0,106 ^z
EA (0-10 puntos)	6,28 (2,08)	6,91 (1,68)	4,18 (1,92)	<0,001 [*]
EAP (0-10 puntos)	6,59 (2,07)	7,26 (1,46)	4,34 (2,19)	<0,001 [*]
Cualitativas	n (% del total)	n (% en apendicitis)	n (% en no apendicitis)	-
Sexo: masculino/femenino	226/226 (50,0/50,0)	200/148 (57,5/42,5)	26/78 (25,0/75,0)	<0,001 ^{**}
Grupo edad: escolares/adolesc.	176/276 (38,9/61,1)	142/106 (40,8/59,2)	34/70 (32,7/67,3)	0,137 ^{**}
Migración del dolor	197 (43,6)	177 (50,9)	20 (19,2)	<0,001 ^{**}
Anorexia	206 (45,6)	181 (52,0)	25 (24,0)	<0,001 ^{**}
Náusea/Vómitos	346 (76,5)	277 (79,6)	69 (66,3)	0,005 ^{**}
Dolor a la palpación en FID	418 (92,5)	346 (99,4)	72 (69,2)	<0,001 ^{**}
Dolor a la descompresión	276 (61,1)	266 (76,4)	10 (9,6)	<0,001 ^{**}
Dolor en FID al golpe de tos/salto/percusión	356 (78,8)	318 (91,4)	38 (36,5)	<0,001 ^{**}
Temperatura elevada	184 (40,7)	152 (43,7)	32 (30,8)	0,019 ^{**}
Leucocitosis $\geq 10\ 000/\mu\text{L}$	299 (66,2)	245 (70,4)	54 (51,9)	<0,001 ^{**}
Neutrófilos $\geq 75\ \%$	196 (43,4)	167 (48,0)	29 (27,9)	<0,001 ^{**}

DT: desviación típica; EA: escala de Alvarado; EAP: escala de apendicitis pediátrica; FID: Fosa ilíaca derecha; escolares: 5-11 años; adolescentes: 12-18 años; p: significación estadística al comparar los grupos Apendicitis y No apendicitis; *U de Mann-Whitney; ** χ^2 .

La edad media fue de 12,41 años (rango 5-18) y 61,1% de los pacientes fueron adolescentes (12-18 años). La edad no fue significativamente diferente entre los grupos apendicitis y no apendicitis. La distribución por sexos fue equitativa en general, pero 75 % de los pacientes en el grupo no apendicitis fueron femeninas ($p < 0,001$). La EA media para los casos con apendicitis fue de 6,91 puntos, y 4,18 puntos para los casos negativos ($p < 0,001$). La EAP media para los casos con apendicitis fue de 7,26, y de 4,34 puntos para los casos sin apendicitis ($p < 0,001$). El componente clínico o de laboratorio más frecuentemente encontrado fue el dolor a la palpación en fosa ilíaca derecha (92,5 %), y el menos observado fue la temperatura elevada (40,7 %).

La sensibilidad, especificidad, valor predictivo positivo (VP+), valor predictivo negativo (VP-), la razón de verosimilitud positiva (RV+), y la razón de verosimilitud negativa (RV-) fueron diferentes para cada escala en la mayoría de los criterios de corte (Tabla 3).

Tabla 3 - Sensibilidad y especificidad con intervalos de confianza de 95 %, razón de verosimilitud positiva y negativa, y valor predictivo positivo y negativo para cada criterio de corte de las escalas de Alvarado y de apendicitis pediátrica

Criterio	Escala	Sens.	95 % CI	Esp.	IC 95 %	RV+	RV-	VP+	VP-
≥ 0 puntos	EA	100,0	98,9 - 100,0	0,00	0,0 - 3,5	1,00	-	77,0	-
	EAP	100,0	98,9 - 100,0	0,00	0,0 - 3,5	1,00	-	77,0	-
> 2 puntos	EA	100,0	98,9 - 100,0	15,38	9,1 - 23,8	1,18	0,00	79,8	100
	EAP	100,0	98,9 - 100,0	19,23	12,2 - 28,1	1,24	0,00	80,6	100
>3 puntos	EA	99,14	97,5 - 99,8	35,58	26,4 - 45,6	1,54	0,02	83,7	92,5
	EAP	99,71	98,4 - 100,0	38,46	29,1 - 48,5	1,62	0,01	84,4	97,6
>4 puntos	EA	89,66	86,0 - 92,6	55,77	45,7 - 65,5	2,03	0,19	87,2	61,7
	EAP	96,55	94,1 - 98,2	52,88	42,8 - 62,8	2,05	0,07	87,3	82,1
>5 puntos	EA	76,72	71,9 - 81,1	75,00	65,6 - 83,0	3,07	0,31	91,1	49,1
	EAP	87,64	83,7 - 90,9	69,23	59,4 - 77,9	2,85	0,18	90,5	62,6
>6 puntos	EA	62,36	57,0 - 67,5	90,38	83,0 - 95,3	6,49	0,42	95,6	41,8
	EAP	69,83	64,7 - 74,6	83,65	75,1 - 90,2	4,27	0,36	93,5	45,3
>7 puntos	EA	39,66	34,5 - 45,0	96,15	90,4 - 98,9	10,3	0,63	97,2	32,3
	EAP	46,55	41,2 - 51,9	92,31	85,4 - 96,6	6,05	0,58	95,3	34,0
>8 puntos	EA	18,97	15,0 - 23,5	99,04	94,8 - 100	19,7	0,82	98,5	26,8
	EAP	21,55	17,3 - 26,2	96,15	90,4 - 98,9	5,60	0,82	94,9	26,8
>9 puntos	EA	4,02	2,2 - 6,7	100,0	96,5 - 100	-	0,96	100	23,7
	EAP	4,60	2,7 - 7,4	100,0	96,5 - 100	-	0,95	100	23,9
>10 puntos	EA	0,00	0,0 - 1,1	100,0	96,5 - 100	-	1,00	-	23,0
	EAP	0,00	0,0 - 1,1	100,0	96,5 - 100	-	1,00	-	23,0

Sens: sensibilidad; Esp: especificidad; IC: intervalo de confianza; RV+: razón de verosimilitud positiva; RV-: razón de verosimilitud negativa; VP+: valor predictivo positivo; VP-: valor predictivo negativo; EA: escala de Alvarado; EAP: escala de apendicitis pediátrica.

Utilizando >6 puntos como el criterio de corte para la EA, se obtuvieron los parámetros sensibilidad 62,36 %, especificidad 90,38 %, VP+ 95,6 %, VP- 41,8 %, RV+ 6,49, y RV- 0,42. En cuanto a la EAP, teniendo en cuenta > 5 puntos como criterio de corte, los resultados fueron sensibilidad 87,64 %, especificidad 69,23 %, VP+ 90,5 %, VP- 62,6 %, RV+ 2,85 y RV- 0,18. En la mayoría de los puntos de corte, la EA tuvo mayor especificidad, mientras que la EAP exhibió mayor sensibilidad.

Se construyeron las curvas COR para cada escala utilizando los valores de sensibilidad y especificidad (Fig.).

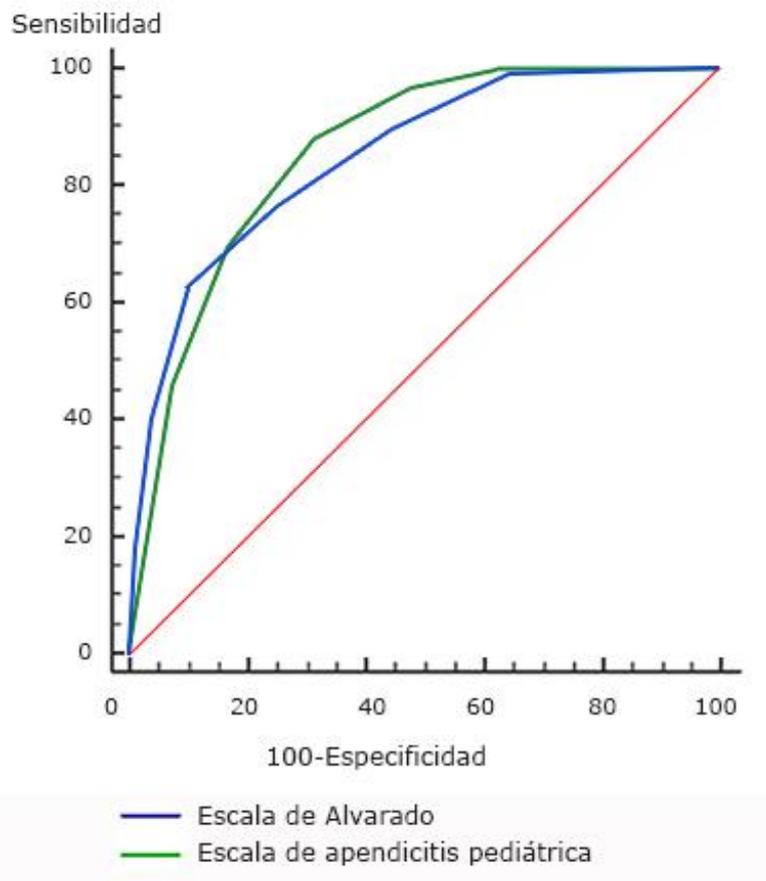


Fig. 1 - Curvas de características operativas del receptor para la escala de Alvarado y la escala de apendicitis pediátrica.

La tabla 4 muestra el análisis del área bajo la curva (ABC) para cada escala y su comparación.

Tabla 4 - Área bajo la curva de características operativas del receptor para la escala de Alvarado y la escala de apendicitis pediátrica, con resultados del análisis estadístico

Escala	ABC	IC95%	Nivel de significación independiente	Índice de Youden J	Criterio asociado	Nivel de significación comparativo *	Medida de acuerdo (kappa)
EA	0,851	0,815 - 0,883	$p < 0,0001$	0,5274	>6	$p = 0,5691$	0,458
EAP	0,858	0,823 - 0,889	$p < 0,0001$	0,5687	>5		

ABC: área bajo la curva; IC: intervalo de confianza; EA: escala de Alvarado; EAP escala de apendicitis pediátrica; *comparación pareada (método DeLong).

El desempeño discriminativo general de ambas escalas, interpretado del área bajo la curva COR, fue de 0,851 para la EA y de 0,858 para la EAP. La diferencia entre ambas no fue significativa ($p=0,5691$), sin embargo, en el análisis de las coincidencias de ambas herramientas para cada criterio de corte, solo se encontró un acuerdo moderado ($kappa=0,458$). El criterio de corte asociado resultó >6 puntos para la EA y >5 puntos para la EAP. Al dicotomizar estas variables partiendo del criterio de corte asociado para cada una, la medida de acuerdo mejoró ligeramente a $kappa=0,503$ ($p<0,001$), pero esto representa aún un acuerdo moderado.

La tabla 5 muestra el área bajo la curva COR para cada escala en las diferentes combinaciones de grupos de edad y sexo, así como la comparación entre ellos.

Tabla 5 - Área bajo la curva de características operativas del receptor y resultados del análisis estadístico para la escala de Alvarado y la escala de apendicitis pediátrica, según grupos de edad y sexo

Grupo de edad	Sexo	Escala	ABC	IC95%	Nivel de significación independiente	Nivel de significación comparativo*	Medida de acuerdo (kappa)
Escolares (5-11 años)	Masculino	EA	0,918	0,845-0,965	$p<0,0001$	$p=0,8877$	0,569
		EAP	0,923	0,850-0,967	$p<0,0001$		
	Femenino	EA	0,859	0,763-0,927	$p<0,0001$	$p=0,4330$	0,466
		EAP	0,838	0,739-0,911	$p<0,0001$		
Adolescentes (12-18 años)	Masculino	EA	0,828	0,752-0,888	$p<0,0001$	$p=0,5758$	0,484
		EAP	0,806	0,728-0,870	$p<0,0001$		
	Femenino	EA	0,802	0,728-0,863	$p<0,0001$	$p=0,0615$	0,340
		EAP	0,841	0,772-0,896	$p<0,0001$		

ABC: área bajo la curva; IC: intervalo de confianza; EA: escala de Alvarado; EAP: escala de apendicitis pediátrica; *comparación pareada (método DeLong).

La mayor ABC se observa para ambas escalas en el grupo de escolares masculinos (EA 0,918; EAP 0,923, $p=0,8877$ en la comparación pareada). En este grupo en particular, el acuerdo entre las escalas fue el mayor ($kappa=0,569$). El ABC para ambas escalas fue en general más alta en escolares que en adolescentes (excepto para la EAP en el sexo femenino). La EA en las adolescentes femeninas exhibió el ABC más bajo (0,802). Las diferencias entre el ABC de ambas escalas para cada combinación de grupo de edad y sexo no fueron significativas y la mayor fue la encontrada en las adolescentes femeninas, a favor de la EAP ($p=0,0615$). También en este grupo específico, la medida de acuerdo entre las escalas fue la más pobre ($kappa=0,340$).

Discusión

El porcentaje de pacientes con apendicitis confirmada fue de 77,0 %, similar a 74,4 % reportado por *Alvarado*,⁽¹⁾ quien también estudió solo pacientes hospitalizados, y superior al 63,0 % publicado por *Samuel* en el estudio que generó la EAP,⁽²⁾ quien también estudió casos con dolor abdominal sugestivo de apendicitis, operados o no, pero inferior al 85,2 % publicado por autores que solo incluyeron pacientes sometidos a apendicectomía,⁽¹⁰⁾ y por encima de otros (54 %⁽⁴⁾ y 36 %, ⁽⁸⁾) cuya población son pacientes con dolor abdominal tratados en el servicio de urgencias, pero no necesariamente hospitalizados.

Se ha demostrado que estas escalas diagnósticas son poco efectivas en niños menores de 5 años,^(33,34) y la atención pediátrica en Cuba incluye pacientes de hasta 18 años, por lo que el rango de edad utilizado fue 5-18 años. La edad media y el grupo de edad predominante que se observan en la tabla 2 se corresponden con la edad pico de incidencia de apendicitis en la población pediátrica.⁽¹¹⁾ Hubo una frecuencia alta del sexo femenino en el grupo de no apendicitis. Esto es una cuestión polémica. Algunos autores, aunque con predominio masculino en ambos grupos, informan una cantidad significativa de niñas en el grupo negativo para apendicitis.⁽⁵⁾ En contraste, otros, no encuentran diferencias importantes en cuanto a la distribución por sexo⁽²⁾ o encuentran que las apendicectomías negativas son más frecuentes en niñas mayores de 10 años.⁽³⁶⁾

Las diferencias entre los puntajes medios para casos con y sin apendicitis fue significativa, y en el caso de la EA, su autor originalmente obtuvo puntajes medios más elevados (7,71 vs. 5,24).⁽¹⁾ Otros han informado una EA media más alta con una diferencia aritmética inferior entre los puntajes medios de los casos con y sin apendicitis.^(4,10,25) En cuanto a la EAP, sus puntajes medios fueron ligeramente más altos en comparación con los de la EA. La diferencia aritmética entre los puntajes para los casos con y sin apendicitis fue mayor que la publicada en la literatura.^(4,10,17,20,37)

Todas las fuentes coinciden en informar el dolor a la palpación en la fosa ilíaca derecha como el componente más frecuentemente encontrado para ambas escalas. Algunas también registran la fiebre como el menos común⁽⁵⁾ y otros plantean la migración del dolor como la menos frecuente.⁽⁴⁾

Como muestra la tabla 3, en la mayoría de los criterios de corte la EA fue más específica, y la EAP más sensible. Algunos autores encuentran resultados similares,⁽⁴⁾ mientras otros no detectan diferencias significativas ni en sensibilidad ni especificidad entre las escalas.⁽¹⁰⁾ La EA y la EAP predicen la apendicitis con puntajes mayor de 6 y de 5, respectivamente.^(1,2) Considerando estos criterios de corte, la EA habría producido 37,64 % de apendicitis no diagnosticadas (sensibilidad 62,36 %), y 9,62 % de apendicectomías negativas (especificidad 90,38 %). Para la EAP, las apendicitis no diagnosticadas habrían sido 12,36 % (sensibilidad 87,64 %) y las apendicectomías negativas 30,77 % (especificidad 69,23 %) Ambas herramientas predictivas tuvieron razones de verosimilitud bajas a moderadas en estos puntos de corte. Estos hallazgos apoyan el consenso de que debido a la insuficiente sensibilidad y especificidad en cualquier criterio de corte, ninguna escala constituye un método diagnóstico efectivo de forma aislada.^(4,9,19,11)

Se ha demostrado que tanto la EA como la EAP son útiles para estratificar a los pacientes como riesgo bajo, intermedio o alto de apendicitis. ^(9,11,12,13,14,15,16,17,18,18,19,20)

Los valores de la tabla 3 muestran que una EA con puntaje menor que 4 puntos, donde la probabilidad de no diagnosticar la apendicitis es de 0,86 % (sensibilidad 99,14 %), y la probabilidad de hacer un diagnóstico negativo correcto es elevada (RV- 0,02), sería seguro descartar la apendicitis. La EAP, siendo más sensible, permitiría descartar la apendicitis con un criterio menor de 5 puntos, donde la apendicitis no diagnosticada sería de 3,45 % (sensibilidad 96,55 %) y la RV- es alta (0,07).

En cuanto a confirmar la apendicitis, o sea el posible límite entre los grupos de riesgo intermedio y alto, los valores de la tabla 3 muestran que una EA mayor de 7 puntos sería un buen criterio, en el cual las apendicectomías negativas serían 3,85 % (especificidad 96,15 %) y la probabilidad de hacer un diagnóstico positivo correcto es alta (RV+ 10,3). Como una escala menos específica, la EAP sería útil para confirmar la apendicitis con puntaje mayor de puntos, donde las apendicectomías negativas serían 3,85 % (especificidad 96,15 %), aunque la RV+ es solo moderada (5,60); no obstante, la EAP mayor de 7 puntos podría ser también una opción (especificidad 92,31 %, RV+ 6,05).

En una revisión sistemática,⁽⁷⁾ sus autores afirman que la razón de verosimilitud para la EA en niños fue de 0,02 (mayor de 4 puntos), 0,27 (4 a 6 puntos), y 4,2 (mayor o igual a 7 puntos). Para la EAP, las razones de verosimilitud fueron 0,13 (menor de 4 puntos), 0,70 (4 a 7 puntos), y 8,1 (mayor o igual a 8 puntos). Concluyen que en niños, una EA menor de 4 descarta la apendicitis, pero la EAP no identificó grupos de riesgo bajo o alto clínicamente significativos.⁽⁷⁾ Algunos autores consideran igualmente una EA menor de 4 para descartar la apendicitis,^(12,25,26) mientras que otros como el propio *Alvarado* recomiendan descartarla con puntajes menor de 5.^(1,13,29)

En cuanto a la EAP, algunos coinciden con los presentes resultados en descartar la apendicitis con puntaje menor de 5,^(14,17) pero otros emplean el criterio de menor de 4 puntos^(18,20) o incluso de 3 puntos.⁽²³⁾ Con respecto al rango de alto riesgo de apendicitis, la literatura tiene ejemplos de EA de 7-10 puntos,⁽¹²⁾ y de 9-10 puntos;⁽²⁶⁾ EAP de 8-10 puntos,^(14,17,18) y de 7-10 puntos.⁽²³⁾

En opinión de los autores, identificar un buen límite de descarte tiene más relevancia, ya que esto determina cuáles pacientes pueden enviarse a casa con seguridad. Establecer un límite de riesgo intermedio a alto también es importante, pero su variabilidad tiene consecuencias en cuanto a un mayor o menor tiempo de observación, la realización de estudios de imágenes o la consulta con el cirujano; mientras tanto ese paciente se encuentra en el hospital hasta que su caso esté bien definido.

El área bajo la curva representa el poder discriminativo de las herramientas diagnósticas, y como se observa en la figura 1 y en la tabla 4, el ABC para la EA y para la EAP fueron muy similares entre sí. En diferentes investigaciones se ha identificado valores parecidos de ABC (EA 0,847 vs. EAP 0,868),⁽⁵⁾ o (EA 0,87 vs. EAP 0,82).⁽⁸⁾ En diferentes momentos evolutivos el ABC para la EA es

siempre superior al ABC para la EAP (0,87-0,90 vs. 0,82-0,87);⁽⁶⁾ otros estudios, muestran valores de ABC más bajos, aunque similares entre sí (EA 0,771 vs. EAP 0,777),⁽⁴⁾ o valores inferiores de ABC para la EAP,^(24,37) o valores de ABC de 0,895 para la EAP.⁽¹⁴⁾

El hecho de que el desempeño discriminativo general de ambas escalas haya resultado similar entre sí, no significa que se comporten igualmente si se considera cada criterio de corte, como se deduce de los resultados de la tabla 3. De forma consistente con esto, la medida de acuerdo entre las escalas (kappa) fue solo moderada. El análisis del ABC arrojó un criterio asociado mayor de 6 puntos para la EA y de 5 puntos para la EAP, como se aprecia en la tabla 4, los mismos que fueron originalmente propuestos para cada una.^(1,2) La mayoría de los autores coinciden en encontrar un criterio mayor de 6 para la EA,^(6,10) y de >5 para la EAP,⁽¹⁴⁾ pero otros han identificado un criterio de corte mayor de 6 para la EAP.^(10,37) De cualquier manera, como se ha dicho ya, el diagnóstico mediante un criterio de corte único no constituye un modo de empleo válido para estas escalas.

En la tabla 5, el desempeño discriminativo general de la EA y la EAP no fue el mismo para cada combinación de grupo de edad y sexo, con mayores ABC para el sexo masculino y la edad escolar. La mayor diferencia en la comparación pareada así como la medida de acuerdo más pobre entre las ABC, se encontró en las adolescentes femeninas, aunque fue marginalmente no significativa, debido a un desempeño relativamente más pobre de la EA en este grupo específico de pacientes.

Se ha encontrado que ambas escalas son mejores para los varones que para las hembras, basados en los valores de sensibilidad, especificidad, VP+ y VP- obtenidos con el criterio de corte originalmente descrito para cada una.⁽⁴⁾ Para algunos autores, la EAP tuvo una mayor especificidad y VP- en niñas, que en niños. Las hembras también son objeto de más apendicectomías negativas, a pesar de haberseles hecho más estudios imagenológicos preoperatorios.⁽³²⁾ En estudios de EAP sin compararla con otra escala y con un criterio de corte mayor de 8 (en lugar de 7), diferentes autores demuestran una mejor especificidad entre las adolescentes femeninas en comparación con el resto de los pacientes.⁽²³⁾ Con un enfoque similar, otros, describen su precisión diagnóstica como más alta en el grupo de 13-18 años en comparación con los de 12 años o menos.⁽³¹⁾

Este estudio tiene algunas limitaciones. Aunque los autores supervisaron la recolección de la información primaria y su coincidencia con los registros clínicos, no se evaluó estadísticamente la posible variabilidad interobservador. Además, pudieran existir algunos falsos negativos debido al tiempo prolongado entre la recolección de los datos para calcular la escala y el diagnóstico definitivo de apendicitis en determinados casos. Por último, una gran proporción de pacientes tuvieron apendicitis, esto puede haber afectado la interpretación de las escalas en su rango bajo. Sin embargo, ambas se estudiaron en los mismos pacientes y bajo las mismas condiciones, por lo que es válida la interpretación de la comparación en su desempeño.

En conclusión, la EA fue relativamente más específica y la EAP fue relativamente más sensible, por lo que se deben utilizar diferentes criterios de corte para cada una al estratificar pacientes pediátricos según su probabilidad de apendicitis. Ambas escalas tuvieron un desempeño

discriminativo general similar, y se comportaron muy bien para los niños de edad escolar y sexo masculino. La mayor diferencia en desempeño se encontró en las adolescentes femeninas, favoreciendo la EAP.

Referencias bibliográficas

1. Alvarado A. A practical score for the early diagnosis of acute appendicitis. *Ann Emer Med*. 1986;15(5):557-64.
2. Samuel M. Pediatric appendicitis score. *J Pediatr Sur*. 2002;37(6):877-81.
3. Kulik DM, Uleryk EM, Maguire JL. Does this child have appendicitis? A systematic review of clinical prediction rules for children with acute abdominal pain. *J Clin Epidemiol*. 2012;66:95-104.
4. Mandeville K, Pottker T, Bulloch B, Liu J. Using appendicitis scores in the pediatric ED. *Am J Emer Med*. 2011;29:972-977.
5. Dokumcu Z, Toker Kurtmen B, Divarci E, Tamay PB, Kose T, Sezak M, *et al*. Retrospective Multivariate Analysis of Data from Children with Suspected Appendicitis: A New Tool for Diagnosis. *Emer Med Int*. 2018;2018:4810730. DOI: [10.1155/2018/4810730](https://doi.org/10.1155/2018/4810730)
6. Wu H-P, Yang W-C, Wu K-H, Chen C-Y, Fu Y-C. Diagnosing Appendicitis at Different Time Points in Children with Right Lower Quadrant Pain: Comparison Between Pediatric Appendicitis Score and the Alvarado Score. *World J Surg*. 2012;36(1):216-21.
7. Ebell MH, Shinholser J. What are the most clinically useful cutoffs for the Alvarado and Pediatric Appendicitis Scores? A systematic review. *Ann Emer Med*. 2014;64(4):365-72.e2.
8. Macco S, Vrouwenraets BC, de Castro SMM. Evaluation of scoring systems in predicting acute appendicitis in children. *Surgery*. 2016;160(6):1599-604. DOI: [10.1016/j.surg.2016.06.023](https://doi.org/10.1016/j.surg.2016.06.023)
9. 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 Emer Surg*. 2020;15(1):27. DOI: <https://doi.org/10.1186/s13017-020-00306-3>
10. Pogorelić Z, Rak S, Mrklič I, Jurić I. Prospective validation of Alvarado score and Pediatric Appendicitis Score for the diagnosis of acute appendicitis in children. *Pediatr Emer Care*. 2015;31(3):164-8. DOI: [10.1097/PEC.0000000000000375](https://doi.org/10.1097/PEC.0000000000000375)
11. Rentea RM, Peter SDS, Snyder CL. Pediatric appendicitis: state of the art review. *Pediatr Surg Int*. 2017;33(3):269-83. DOI: [10.1007/s00383-016-3990-2](https://doi.org/10.1007/s00383-016-3990-2)
12. Sincavage J, Buonpane C, Benyamen B, Benya E, Lautz T, Helenowski I, *et al*. Alvarado Scores Predict Additive Value of Magnetic Resonance Imaging in Workup of Suspected Appendicitis in Children. *J Surg Res*. 2019;244:42-9. DOI: [10.1016/j.jss.2019.06](https://doi.org/10.1016/j.jss.2019.06)
13. Blitman NM, Anwar M, Brady KB, Taragin BH, Freeman K. Value of Focused Appendicitis Ultrasound and Alvarado Score in Predicting Appendicitis in Children: Can We Reduce the Use of CT? *Am J Roentgenol*. 2015;204(6):W707-12. DOI: [10.2214/AJR.14.13212](https://doi.org/10.2214/AJR.14.13212)

14. Bhatt M, Joseph L, Ducharme FM, Dougherty G, McGillivray D. Prospective Validation of the Pediatric Appendicitis Score in a Canadian Pediatric Emergency Department. *Acad Emer Med*. 2009;16(7):591-6.
15. Fujii T, Tanaka A, Katami H, Shimono R. Usefulness of the pediatric appendicitis score for assessing the severity of acute appendicitis in children. *Pediatr Int*. 2020;62(1):70-3. DOI:[10.1111/ped.14032](https://doi.org/10.1111/ped.14032)
16. El-Gohary Y, Molina M, Chang J, Dodd A, Miller E, Harrell C, *et al*. The Use of Computed Tomography Versus Clinical Acumen in Diagnosing Appendicitis in Children: A Two-Institution International Study. *J Pediatric Surg*. 2021;56(8):1356-61. DOI: [10.1016/j.jpedsurg.2020.09.061](https://doi.org/10.1016/j.jpedsurg.2020.09.061)
17. Aydin D, Turan C, Yurtseven A, Bayindir P, Toker B, Dokumcu Z, *et al*. Integration of radiology and a clinical score in pediatric appendicitis. *Pediatr Int*. 2018;60(2):173-8. DOI: [10.1111/ped.13471](https://doi.org/10.1111/ped.13471)
18. Kobayashi E, Johnson B, Goetz K, Scanlan J, Weinsheimer R. Does the implementation of a pediatric appendicitis pathway promoting ultrasound work outside of a children's hospital? *Am J Surg*. 2018;215(5):917-20. DOI: [10.1016/j.amjsurg.2018.03.017](https://doi.org/10.1016/j.amjsurg.2018.03.017)
19. Lipsett SC, Bachur RG. Current Approach to the Diagnosis and Emergency Department Management of Appendicitis in Children. *Pediatr Emer Care*. 2017;33(3):198-203. DOI: [10.1097/PEC.0000000000001049](https://doi.org/10.1097/PEC.0000000000001049)
20. Planella Cornudella M, Pociello Almiñana N, Domingo Ruiz A, Bringue Espuny X, Rue Monne M, Sole Mir E. Utilidad del Pediatric Appendicitis Score y la ecografía abdominal en el proceso diagnóstico de la apendicitis aguda. *Cir Pediatr*. 2019 [19/03/2020];32(3):128-34. Disponible en: https://secipe.org/coldata/upload/revista/2019_32-3_128-134.pdf
21. Bachur RG, Callahan MJ, Monuteaux MC, Rangel SJ. Integration of ultrasound findings and a clinical score in the diagnostic evaluation of pediatric appendicitis. *J Pediatr*. 2015;166(5):1134-9. DOI: [10.1016/j.jpeds.2015.01.034](https://doi.org/10.1016/j.jpeds.2015.01.034)
22. Kharbanda AB, Madhok M, Krause E, Vazquez-Benitez G, Kharbanda EO, Mize W, *et al*. Implementation of Electronic Clinical Decision Support for Pediatric Appendicitis. *Pediatrics*. 2016;137(5):e20151745. DOI:[10.1542/peds.2015-1745](https://doi.org/10.1542/peds.2015-1745)
23. Scheller RL, Depinet HE, Ho ML, Hornung RW, Reed JL. Utility of Pediatric Appendicitis Score in Female Adolescent Patients. *Acad Emer Med*. 2016;23(5):610-5. DOI: [10.1111/acem.12916](https://doi.org/10.1111/acem.12916)
24. González del Castillo J, Ayuso FJ, Trenchs V, Martínez Ortiz de Zarate M, Navarro C, Altali K, *et al*. Diagnostic accuracy of the APPY1 Test in patients aged 2-20 years with suspected acute appendicitis presenting to emergency departments. *Emer Med J*. 2016;33(12):853-9. DOI: [10.1136/emered-2015-205259](https://doi.org/10.1136/emered-2015-205259)
25. Doniger SJ, Kornblith A. Point-of-Care Ultrasound Integrated Into a Staged Diagnostic Algorithm for Pediatric Appendicitis. *Pediatr Emer Care*. 2018;34(2):109-15. DOI: [10.1097/PEC.0000000000000773](https://doi.org/10.1097/PEC.0000000000000773)
26. Dhatt S, Sabhaney V, Bray H, Skarsgard ED. Improving the diagnostic accuracy of appendicitis using a multidisciplinary pathway. *J Pediatr Surg*. 2020;55(5):889-92. DOI: [10.1016/j.jpedsurg.2020.01.040](https://doi.org/10.1016/j.jpedsurg.2020.01.040)

27. Dahn CM, Milne WK. Hot Off the Press: Does This Adolescent Female Have Appendicitis? Can the Pediatric Appendicitis Score Help? *Acad Emer Med*. 2017;24(1):130-2. DOI: [10.1111/acem.13057](https://doi.org/10.1111/acem.13057)
28. Amini M, Zandbaf T, Alizadeh SA, Jand Y, Hosseini A, Eshrati B, *et al*. The significance of Alvarado score in the diagnosis of acute appendicitis based on age and sex. (English). *Arak Med University J*. 2011;14(1):12-8.
29. Ohle R, O'Reilly F, O'Brien KK, Fahey T, Dimitrov BD. The Alvarado score for predicting acute appendicitis: a systematic review. *BMC Med*. 2011;9:139.
30. Tan WJ, Acharyya S, Goh YC, Chan WH, Wong WK, Ooi LL, *et al*. Prospective Comparison of the Alvarado Score and CT Scan in the Evaluation of Suspected Appendicitis: A Proposed Algorithm to Guide CT Use. *J Am Coll Surg*. 2015;220(2):218-24. DOI: [10.1016/j.jamcollsurg.2014.10.010](https://doi.org/10.1016/j.jamcollsurg.2014.10.010)
31. Paramitha D, Fikri E, Nasution I. The accuracy of the Paediatric Appendicitis Score based on age group in cases of acute appendicitis. *Stem Cell Oncol*. 2018 [20/03/2019];289-92. Disponible en: <https://library.oapen.org/bitstream/handle/20.500.12657/30284/646791.pdf?sequence=1#page=305>
32. Salö M, Ohlsson B, Arnbjörnsson E, Stenström P. Appendicitis in children from a gender perspective. *Pediatr Surg Int*. 2015;31(9):845-53. DOI: [10.1007/s00383-015-3729-5](https://doi.org/10.1007/s00383-015-3729-5)
33. Song CW, Kang JW, Kim JY. Different Clinical Features and Lower Scores in Clinical Scoring Systems for Appendicitis in Preschool Children: Comparison with School Age Onset. *Pediatr Gastroenterol, Hepatol Nutrit*. 2018;21(1):51-8. DOI: [10.5223/pghn.2018.21.1.51](https://doi.org/10.5223/pghn.2018.21.1.51)
34. Pogorelić Z, Domjanović J, Jukić M, Peričić TP. Acute Appendicitis in Children Younger than Five Years of Age: Diagnostic Challenge for Pediatric Surgeons. *Surg Infect*. 2020;21(3):239-45. DOI: [10.1089/sur.2019.175](https://doi.org/10.1089/sur.2019.175)
35. Declaración de Helsinki de la Asociación Médica Mundial AMM - Principios éticos para las investigaciones médicas en seres humanos. New York: Representación AMM. 2017 [acceso 14/02/2020] Disponible en: <https://www.wma.net/es/policias-post/declaracion-de-helsinki-de-la-amm-principios-eticos-para-las-investigaciones-medicas-en-seres-humanos/>
36. Bachur RG, Hennelly K, Callahan MJ, Chen C, Monuteaux MC. Diagnostic imaging and negative appendectomy rates in children: effects of age and gender. *Pediatrics*. 2012;129(5):877-84.
37. Akgül F, Er A, Ulusoy E, Çağlar A, Çitlenbik H, Keskinoglu P, *et al*. Integration of Physical Examination, Old and New Biomarkers, and Ultrasonography by Using Neural Networks for Pediatric Appendicitis. *Pediatr Emerg Care*. 2019;10.1097/PEC.0000000000001904. DOI: [10.1097/PEC.0000000000001904](https://doi.org/10.1097/PEC.0000000000001904)

Conflictos de intereses

Los autores declaran no existencia de conflicto de intereses.

Contribución de los autores

Alejandro Ramírez Guirado: Conceptualización; curación de datos; análisis formal; metodología; supervisión; redacción – borrador original; redacción – revisión y edición.

Ana Barbarita Navarro Sombert: Curación de datos; supervisión; redacción – revisión y edición.

Ramón Villamil Martínez: Redacción – revisión y edición.