

ORIGINAL: ESCALAS DE PROBABILIDAD DIAGNÓSTICA EN LA EVALUACIÓN DEL ABDOMEN AGUDO DE FOSA ILÍACA DERECHA

Juan Ramón Gómez-López¹, Juan Carlos Martín-Del Olmo¹, Isabel Cristina López Mestanza², María Antonia Montenegro-Martín¹, Pilar Concejo-Cutoli¹, Eduardo Pablo Bordons-Mesonero¹, Miguel Toledano-Trincado³, Carlos Vaquero Puerta⁴.

1: Servicio de Cirugía General, Hospital Comarcal de Medina del Campo (Valladolid).
2: Servicio de Microbiología, Hospital Virgen de la Concha (Zamora). 3: Servicio de Cirugía General y Digestiva, Hospital Universitario Río Hortega (Valladolid). 4: Departamento de Cirugía, Universidad de Valladolid.

Correspondencia: para contactar al autor accionar [aquí](#) (Juan Ramón Gómez- López).

ABSTRACT:

Introducción: Para facilitar el diagnóstico de apendicitis aguda, se crearon escalas que objetivamente determinan la probabilidad de padecerla. Implementarlas reduciría los costes del tratamiento del dolor de fosa ilíaca derecha. El objetivo fue evaluar la fiabilidad diagnóstica de cuatro de las escalas en los pacientes de un Hospital de Nivel I y valorar su utilidad.

Material y métodos: Estudio observacional retrospectivo en un hospital de nivel I desde 2010 hasta 2016. Se recogieron datos de 458 pacientes intervenidos por sospecha de apendicitis aguda. Se seleccionaron las escalas de Alvarado, RIPASA (*Raja Isteri Pengiran Anak Saleha Appendicitis score*), AIR (*Appendicitis Inflammatory Response*) y AAS (*Adult Appendicitis Score*); calculándose las puntuaciones de los pacientes para cada escala, así como su estratificación. Se realizó un análisis mediante curvas ROC (*Receiver Operating Characteristic*),

calculando el Área Bajo la Curva (ABC) de cada una de las escalas. Los datos se analizaron mediante el programa estadístico SPSS versión 20.0.

Resultados: Todas las escalas fueron estadísticamente significativas al comparar los pacientes con apendicitis confirmada frente a aquellos con apendicectomía negativa, aunque fue la escala de Alvarado fue la que obtuvo el mayor ABC con 0,74 (IC 95%= 0,67–0,80, $p < 0,001$).

Discusión: Se ha demostrado que estas escalas proporcionan una evaluación objetiva y estandarizada de los pacientes con sospecha de apendicitis. La escala de Alvarado, la primera establecida, permite una buena discriminación cuando su puntuación es mayor o igual a 7; además, es la más sencilla de obtener y la más objetiva. La escala AIR en nuestra cohorte no mejora los resultados de Alvarado. La escala AAS tampoco mejora a la de Alvarado y es la más complicada de obtener. La escala de RIPASA es la peor escala para los pacientes de nuestro trabajo.

Conclusiones: Las escalas de probabilidad diagnóstica son herramientas útiles para la evaluación del abdomen agudo, pudiendo ahorrar tiempo y pruebas innecesarias. La escala de Alvarado, es la más sencilla, objetiva y fiable para nuestros pacientes.

Palabras clave: Apendicitis, Escalas de probabilidad, Abdomen agudo

ABSTRACT EN INGLÉS:

Introduction: To facilitate the diagnosis of acute appendicitis, scales were created that objectively determine the probability of suffering it. Implementing them would reduce the costs of treating right iliac fossa pain. The objective was to evaluate the diagnostic reliability of four of the scales in the patients of a Level I Hospital and to assess its usefulness.

Material and methods: A retrospective observational study was conducted in a Level I Hospital from 2010 to 2016. Data were collected from 458 patients operated on for suspected acute appendicitis. The scales of Alvarado, RIPASA (Raja Isteri Pengiran Anak Saleha Appendicitis score), the AIR scale (Appendicitis Inflammatory Response) and the AAS scale (Adult Appendicitis Score) were selected; the scores of the patients for each scale are calculated, as well as their stratification. An analysis was made using ROC curves (Receiver Operating Characteristic), calculating the Area under the Curve (ABC) of each of the scales. The data was analyzed using the statistical program IBM SPSS Statistics version 20.0.

Results: All scales were statistically significant when comparing patients with confirmed appendicitis vs patients with negative appendectomy. Although it was the Alvarado scale that obtained the highest ABC with 0.74 (CI95% = 0.67-0.80, $p < 0.001$).

Discussion: It has been shown that these scales provide an objective and standardized evaluation of patients with suspected appendicitis. Alvarado's scale, the first established, allows good discrimination when his score is greater than or equal to 7; In addition, it is the simplest to obtain and the most objective. The AIR scale in our cohort does not improve Alvarado's results. The AAS scale does not improve to Alvarado's and it is the most complicated to obtain. The RIPASA scale is the worst scale for the patients of our work.

Conclusions: Diagnostic probability scales are useful tools for the evaluation of acute abdomen; being able to save time and unnecessary tests. The Alvarado scale is the simplest, objective and reliable for our patients.

Key words: Appendicitis, Clinical Prediction Rules, Acute abdomen

ARTÍCULO ORIGINAL:

Introducción:

El cuadro típico de apendicitis aguda (AA) es muy sugestivo y permite sospechar fácilmente esta patología. Sin embargo, el correcto diagnóstico de AA sigue siendo un reto, ya que los signos típicos pueden no encontrarse hasta en un 55% de los casos (1). Se estima que no se diagnostican un 20-40% de casos y, por el contrario, se aportan datos de apendicectomías negativas o en blanco de entre el

10% - 34% (2,3). En la actualidad, se acepta hasta un 10 – 15% de apendicectomías en blanco para disminuir la incidencia de complicaciones (perforación, peritonitis...) (4). Esto se debe a que un retraso en el diagnóstico de una AA puede originar complicaciones graves para el paciente como perforación, abscesos, sepsis – shock séptico, etc. (5). Por otro lado, se sabe que una apendicectomía negativa (AN) no está exenta de morbilidad y puede estar acompañada de complicaciones postoperatorias, infecciones nosocomiales e incluso, se han descrito casos de muerte (6).

El diagnóstico de la AA debe ser eminentemente clínico y las actuales pruebas de imagen no deberán sustituir a una correcta anamnesis y exploración física. Aunque es cierto que éstas pueden ser de gran ayuda en los casos de presentación atípica y en los casos dudosos. En la literatura, existen artículos que defienden el uso rutinario de la ecografía en el diagnóstico de la AA (7) y artículos que defienden lo contrario (8). Lo que sí es cierto, es que la ecografía es una prueba observador-dependiente y existe gran variabilidad en cuanto a su sensibilidad y especificidad diagnósticas en la AA (9). En la actualidad, la prueba de imagen *gold-standard* para el diagnóstico de la AA es la Tomografía Computarizada (TC) pudiendo disminuir la tasa de AN por debajo del 10% (2,10). Sin embargo, es una prueba no exenta de riesgos (radiación ionizante, daño renal por contrastes, etc.), cara y no siempre disponible en todos los centros (11), por lo que quedará relegada a los casos en los que la ecografía sea dudosa o negativa y exista alta sospecha de patología quirúrgica (12).

En un intento de facilitar el diagnóstico de la apendicitis aguda, se han creado sistemas de escalas que objetivamente intentan determinar la probabilidad diagnóstica de padecer o no AA (5,13). Se han propuesto diversos tipos de escalas y

parece que la implementación de las mismas podría haber reducido los costes totales del tratamiento del dolor del cuadrante inferior derecho del abdomen (5).

El objetivo de nuestro estudio fue evaluar la fiabilidad diagnóstica de cuatro de las escalas reconocidas y aceptadas en los pacientes de un Hospital de Nivel I y valorar la utilidad de las mismas.

Material y métodos:

1.- CARACTERÍSTICAS DEL ESTUDIO:

Se realizó un estudio observacional retrospectivo en un Hospital de nivel I desde el 1 de enero de 2010 hasta el 31 de diciembre de 2016. Se revisaron las historias clínicas de un total de 500 pacientes que ingresaron en dicho hospital y que fueron intervenidos quirúrgicamente por sospecha de apendicitis aguda. Del total de historias clínicas revisadas, 458 pacientes cumplían los criterios de inclusión diseñados para el estudio: pacientes intervenidos quirúrgicamente por sospecha de AA (independientemente de si presentaban dicha entidad o de si se trató de apendicectomía negativa). Se consideró AA cuando en el análisis anatómico-patológico del apéndice se encontraba la presencia de células inflamatorias (leucocitos, linfocitos o células plasmáticas) (14,15). Mientras que se definió AN cuando dicho análisis mostraba la ausencia de datos de inflamación o patología en el espécimen (15).

Los datos recogidos de las historias clínicas fueron: datos demográficos y antecedentes personales de los pacientes, datos clínicos y analíticos al ingreso, datos radiológicos, datos de la intervención y datos de la evolución postoperatoria. Se seleccionaron cuatro de las escalas de probabilidad diagnóstica validadas y

publicadas en la literatura. Estas escalas fueron: escala de Alvarado, escala RIPASA (*Raja Isteri Pengiran Anak Saleha Appendicitis score*), escala AIR (*Appendicitis Inflammatory Response*) y escala AAS (*Adult Appendicitis Score*), con sus respectivos algoritmos diagnósticos.

Finalmente, se calcularon con dichos datos las escalas de probabilidad diagnóstica de apendicitis aguda Alvarado, RIPASA, AIR y AAS. También se calculó la estratificación de los pacientes en cada escala según su puntuación.

2.- ESTUDIO ESTADÍSTICO:

En el análisis descriptivo, las variables cuantitativas se representaron con la mediana y el rango intercuartílico (RIC). Las variables cualitativas se representaron según su distribución de frecuencias y porcentaje en relación al total de pacientes (N, %). Al mismo tiempo, se dividió la cohorte en dos grupos según el análisis histológico del apéndice (apendicitis aguda y apendicectomía negativa). La comparación de las variables cualitativas entre ambos grupos se realizó mediante el test Chi-cuadrado de Pearson (χ^2) y la comparación de las variables cuantitativas se realizó mediante las pruebas no paramétricas U de Mann Whitney y Kruskal Wallis, según si el número de grupos a comparar eran dos o mayor. Además se realizó un análisis descriptivo de la puntuación de cada una de las escalas a estudio en los pacientes de nuestra cohorte global.

Para predecir el valor diagnóstico de las escalas a estudio, se realizó un análisis mediante curvas ROC (*Receiver Operating Characteristic*), calculando el Área Bajo la Curva (ABC) de cada una de las escalas.

Para analizar los datos se usaron los programas estadísticos IBM SPSS Statistics versión 20.0 para Windows. Aquellos valores de $p < 0,05$ fueron considerados estadísticamente significativos.

3.- ASPECTOS ÉTICOS:

La metodología del estudio fue aprobada por la Comisión de Investigación y el Comité Ético de Investigación Clínica (CEIC) del Área de Salud Valladolid Este y por la Dirección del Hospital de Medina del Campo.

Resultados:

1.- DESCRIPTIVO:

Se analizaron 458 pacientes intervenidos quirúrgicamente por sospecha de AA. De éstos, 404 (88,2%) tuvieron confirmación histológica de apendicitis y 54 (11,8%) presentaron un AN. La mediana de edad del total de pacientes fue de 31 años (RIC= 18,0 – 48,0 años). En la distribución por sexo, se observó un predominio masculino con 266 (58,1%) varones afectados. El 60,9% del total de pacientes con AA confirmada histológicamente eran varones frente al 37% de AN ($p < 0,001$). No se encontraron diferencias estadísticamente significativas cuando se compararon los antecedentes personales en el grupo de AA confirmada frente a los AN.

El tipo de intervención quirúrgica realizada, fue fundamentalmente la apendicectomía laparoscópica, empleada en 443 casos (94,5%). La conversión a cirugía abierta se realizó en 12 de estos (en 5 de ellos existió dificultad técnica para realizar pneumoperitoneo, 2 casos de dilatación de asas de intestino secundarias a íleo paralítico, 2 casos de peritonitis purulenta difusa, 2 casos de síndrome

adherencial y un caso, con apéndice normal pero con la presencia de perforación intestinal); lo que supuso una tasa de conversión del 2,7%.

Las complicaciones postoperatorias, de acuerdo a la clasificación de Clavien – Dindo, tuvieron lugar en 59 pacientes (12,9%), predominando las de Grado I y II en 47 pacientes (80%,) que son aquellas que no precisan de tratamiento invasivo (quirúrgico, radiológico intervencionista ni endoscópico) para solventarlas. Los otros 10 pacientes, precisaron reintervención (2,2% de toda la serie), constituyendo el 16,9% de los pacientes con complicaciones Clavien - Dindo Grado IIIb. De estos 10 pacientes reintervenidos, uno tenía AN y fue reoperado por sospecha de absceso postquirúrgico, encontrando sangre en Douglas. Los otros 9 pacientes tuvieron AA y volvieron a ser operados por: absceso en 5 casos, hemoperitoneo en 2, una dehiscencia de muñón y una oclusión intestinal. Por último, la estancia hospitalaria de los pacientes tuvo una mediana de 2 días (RIC=1,0 – 4,0), con rápida recuperación e inicio de tolerancia oral, con una mediana de 12 horas tras la cirugía (RIC=8,0 – 24,0) (ver **Tabla 1**).

	Total = 458	CONFIRMACIÓN HISTOLÓGICA		P
		AN = 54 (11,8%)	AA = 404 (88,2%)	
Edad (años)	31,00 [18,00 - 48,00]	30,00 [18,00 - 46,00]	31,00 [17,50 - 48,00]	NS
Sexo (masculino)	266 (58,1%)	20 (37,0%)	246 (60,9%)	< 0,001
ASA	1 [1 - 2]	1 [1 - 2]	1 [1 - 2]	NS
Laparoscópica	443 (94,5%)	52 (96,3%)	381 (94,3%)	NS
Duración cirugía (minutos)	50,00 [35,00 - 60,00]	40,00 [35,00 - 60,00]	50,00 [40,00 - 60,00]	NS
Complicaciones	59 (12,9%)	4 (7,4%)	55 (13,6%)	NS
Reintervención	10 (2,2%)	1 (1,9%)	9 (2,2%)	NS
Tolerancia postoperatoria (horas, RIQ)	12 [8 - 24]	12 [8 - 24]	12 [8 - 24]	NS
Analgesia de 2º Escalón (Días)	0 [0 - 1]	0 [0 - 1]	0 [0 - 1]	NS
Estancia hospitalaria (Días)	2 [1 - 4]	2 [2 - 4]	2 [1 - 4]	NS

Tabla 1. Análisis descriptivo en función de la confirmación histológica.

2.- ANÁLISIS DE LAS ESCALAS

Se aplicaron las escalas a estudio: Alvarado (ver **Tabla 2**), RIPASA (ver **Tabla 3**), AIR (ver **Tabla 4**) y AAS (ver **Tabla 5**) a los pacientes de nuestra cohorte, resultando las cuatro estadísticamente significativas, al comparar su aplicación en los pacientes con AA frente a los pacientes con AN; siendo la mediana de puntuación mayor en el subgrupo de pacientes con apendicitis verdaderas.

	Total = 458	CONFIRMACIÓN HISTOLÓGICA	
		AN = 54 (11,8%)	AA = 404 (88,2%)
Anorexia (1 Pto)	139 (30,4%)	10 (18,5%)	129 (32,0%)
Migración del dolor a FID (1 Pto)	177 (38,6%)	13 (24,1%)	164 (40,6%)
Náuseas/Vómitos (1 Pto)	282 (61,6%)	30 (55,6%)	252 (62,4%)
Dolor FID (2 Ptos)	369 (80,6%)	40 (74,1%)	329 (81,4%)
Blumberg (+) (1 Pto)	341 (74,5%)	35 (64,8%)	306 (75,7%)
Tª > 37,3 °C (1 Pto)	110 (24,0%)	17 (31,5%)	93 (23,0%)
Leucocitos > 10000/μL (2 Pto)	363 (79,3%)	25 (46,3%)	338 (83,7%)
Neutrofilia > 75% (1 Pto)	351 (76,6%)	25 (46,3%)	326 (80,7%)

Tabla 2. Resultado para la escala de Alvarado.

	Total = 458	CONFIRMACIÓN HISTOLÓGICA	
		AN = 54 (11,8%)	AA = 404 (88,2%)
Hombre (1 Pto)	266 (58,1%)	20 (37,0%)	246 (60,9%)
Edad ≤ 39 años	300 (65,5%)	39 (72,2%)	261 (64,6%)
Extranjero (1 Pto)	38 (8,3%)	4 (7,4%)	34 (8,4%)
Náuseas/Vómitos (1 Pto)	285 (62,2%)	30 (55,6%)	255 (63,1%)
Anorexia (1 pto)	135 (29,5%)	10 (18,5%)	125 (30,9%)
Dolor FID (0,5 Pto)	369 (80,6%)	40 (74,1%)	329 (81,4%)
Migración dolor a FID (0,5 Pto)	177 (38,6%)	13 (24,1%)	164 (40,6%)
Evolución < 48 horas (1 Pto)	350 (76,4%)	30 (55,6%)	320 (79,2%)
Hipersensibilidad FID (1 Pto)	0 (0,0%)	0 (0,0%)	0 (0,0%)
Rovsing (+) (2 Pto)	25 (5,5%)	3 (5,6%)	22 (5,4%)
Blumberg (+) (1 Pto)	341 (74,5%)	35 (64,8%)	306 (75,7%)
Defensa (2 Ptos)	257 (56,1%)	27 (50,0%)	230 (56,9%)
Tª > 37,0°C y < 39°C (1 Pto)	159 (34,7 %)	25 (46,3%)	134 (33,2%)
Leucocitos > 10000/μL (1 Pto)	363 (79,3%)	25 (46,3%)	338 (83,7%)
S. de Orina (-) (1 Pto)	177 (55,8%)	23 (62,2%)	154 (55,0%)

Tabla 3. Resultado para la escala de RiPASA.

	Total = 458	CONFIRMACIÓN HISTOLÓGICA	
		AN = 54 (11,8%)	AA = 404 (88,2%)
Vómitos (1 Pto)	184 (40,2%)	17 (31,5%)	167 (41,3%)
Dolor FID (1 Pto)	369 (80,6)	40 (74,1%)	329 (81,4%)
Defensa/Rebote LIGERO (1 Pto)	163 (35,6%)	18 (33,3%)	145 (35,9%)
Defensa/Rebote MODERADO (2 Ptos)	220 (48,0%)	17 (31,5%)	203 (50,2%)
Defensa/Rebote SEVERO (3 Ptos)	28 (6,1%)	7 (13,0%)	21 (5,2%)
Tª > 38,5°C (1 Pto)	18 (3,9%)	2 (3,7%)	16 (4,0%)
Leucocitos 10000/μL - 14999/μL (1 Pto)	210 (45,9%)	14 (25,9%)	196 (48,5%)
Leucocitos > 15000/μL (2 Ptos)	157 (34,3%)	11 (20,4%)	146 (36,1%)
Neutrofilia 70% - 84,9% (1 Pto)	207 (45,2%)	19 (35,2%)	188 (46,5%)
Neutrofilia > 85% (2 Ptos)	179 (39,1%)	11 (20,4%)	168 (41,6%)
PCR 10 - 49 (1 Pto)	51 (11,4%)	8 (14,8%)	43 (10,9%)
PCR > 50 mg/dL (2 Ptos)	1 (0,2%)	0 (0,0%)	1 (0,3%)

Tabla 4. Resultado para la escala de AIR.

	Total = 458	CONFIRMACIÓN HISTOLÓGICA	
		AN = 54 (11,8%)	AA = 404 (88,2%)
Dolor FID (2 Ptos)	369 (80,6%)	40 (74,1%)	329 (81,4%)
Migración del dolor a FID (2 Pto)	175 (38,2%)	13 (24,1%)	162 (40,1%)
Edad > 50 años (3 Ptos)	101 (22,1%)	10 (18,5%)	91 (22,5%)
Mujer de 16 a 49 años (1 Pto)	114 (24,9%)	21 (38,9%)	93 (23,0%)
Rebote/Defensa LEVE (2 Ptos)	163 (35,6%)	18 (33,3%)	145 (35,9%)
Rebote/Defensa MODERADO (3 Ptos)	220 (48,1%)	17 (31,5%)	203 (50,4%)
Rebote/Defensa SEVERO (4 Ptos)	28 (6,1%)	7 (13,0%)	21 (5,2%)
Leucocitos 7200/μL - 10900/μL (1 Pto)	88 (19,2%)	20 (37,0%)	68 (16,8%)
Leucocitos 10900/μL - 14999/μL (2 Pto)	183 (40,0%)	11 (20,4%)	145 (35,9%)
Leucocitos > 15000/μL (3 Ptos)	157 (34,3%)	11 (20,4%)	146 (36,1%)
Neutrófilos 62% - 75% (2 Ptos)	81 (17,7%)	18 (33,3%)	63 (15,6%)
Neutrofilia 75% - 84,9% (3 Ptos)	170 (37,1%)	14 (25,9%)	156 (38,6%)
Neutrofilia > 85% (4 Ptos)	179 (39,1%)	11 (20,4%)	168 (41,6%)
PCR 4 - 11 + Evolución < 24 h (2 Ptos)	51 (11,4%)	6 (11,1%)	45 (11,4%)
PCR 11 - 25 + Evolución < 24 h (3 Ptos)	10 (2,2%)	1 (1,9%)	9 (2,3%)
PCR 25 - 83 + Evolución < 24 h (5 Ptos)	0 (0,0%)	0 (0,0%)	0 (0,0%)
PCR > 83 + Evolución < 24 h (1 Pto)	0 (0,0%)	0 (0,0%)	0 (0,0%)
PCR 12 - 53 + Evolución > 24 h (2 Ptos)	32 (7,1%)	5 (9,3%)	27 (6,8%)
PCR 53 - 152 + Evolución > 24 h (2Pts)	0 (0,0%)	0 (0,0%)	0 (0,0%)
PCR > 152 + Evolución > 24 h (1 Pto)	0 (0,0%)	0 (0,0%)	0 (0,0%)

Tabla 5. Resultado para la escala de AAS.

El análisis para detectar las diferencias estadísticas entre el grupo de AN y AA para las puntuaciones obtenidas en las 4 escalas evaluadas en este trabajo se

muestra en la **tabla 6**. Al calcular el ABC de cada escala en función de la probabilidad para predecir el diagnóstico de apendicitis, todas las escalas fueron estadísticamente significativas al comparar los pacientes con AA vs los pacientes con AN. Aunque fue la escala de Alvarado la que obtuvo el mayor ABC con 0,74 (IC 95%= 0,67 – 0,80, $p < 0,001$). Las escalas de AIR y AAS obtuvieron ambas el mismo ABC= 0,70 (IC 95%= 0,62 – 0,78, $p < 0,001$). Por último, la escala de RIPASA alcanzó un ABC= 0,63 (IC 95%= 0,56 – 0,71, $p < 0,001$) (ver **tabla 7** y **figura 1**).

	Total = 458	CONFIRMACIÓN HISTOLÓGICA		P
		AN = 54 (11,8%)	AA = 404 (88,2%)	
Escala de Alvarado	6,0 [5,0 - 8,0]	5,0 [4,0 - 6,0]	6,0 [5,0 - 8,0]	< 0,001
Escala de RIPASA	7,5 [6,5 - 9,0]	7,0 [5,5 - 8,0]	7,5 [6,5 - 9,0]	0,001
Escala AIR	5,0 [4,0 - 7,0]	4,0 [2,0 - 5,0]	5,0 [4,0 - 7,0]	< 0,001
Escala AAS	11,0 [9,0 - 13,0]	9,0 [7,0 - 11,0]	11,0 [9,0 - 13,5]	< 0,001

Tabla 6. Análisis inferencial de las puntuaciones en las 4 escalas en los grupos del estudio.

ABC DE LAS ESCALAS A ESTUDIO			
Escalas	ABC	Intervalo de confianza al 95%	P
Alvarado	0,74	0,67 - 0,80	<0,001
RIPASA	0,63	0,56 - 0,71	<0,001
AIR	0,70	0,62 - 0,78	<0,001
AAS	0,70	0,62 - 0,78	<0,001

Tabla 7. ABC de las escalas a estudio para predecir el diagnóstico de AA en la cohorte.

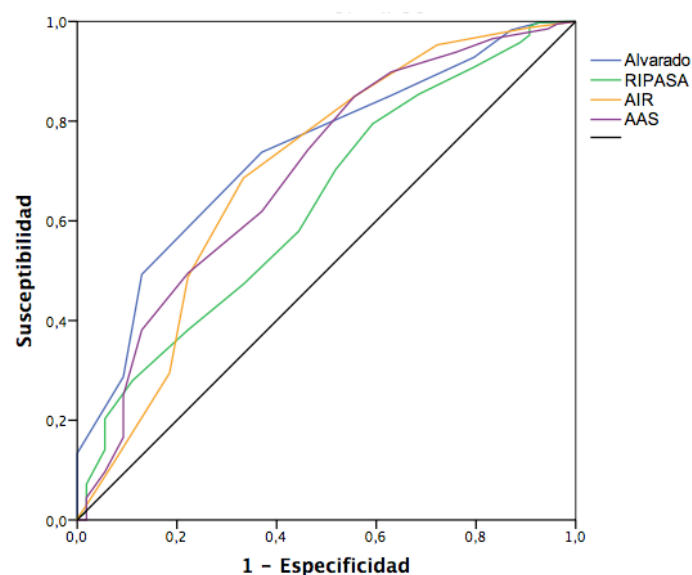


Figura 1. Curvas COR de las escalas a estudio para predecir el diagnóstico de AA.

Discusión:

La AA es la principal causa de abdomen agudo quirúrgico (16). Se estima que entre el 5% y el 15% de la población padecerá a lo largo de su vida una AA (17). Presenta un pico de incidencia entre la segunda y la tercera décadas de la vida (17-19), siendo poco común en menores de 5 años y por encima de los 70; donde el riesgo de sufrirla es inferior al 1% (18,20,21). Este riesgo es, aproximadamente, del 8,6% en el varón y del 6,7% en la mujer (17) y el ratio hombre:mujer es de 1,3:1, aunque en la edad adulta parece igualarse (18,22).

En nuestra cohorte de pacientes encontramos un ligero predominio de varones (58,1% vs 41,9%) con un ratio 1,35:1. La mediana de edad que se obtuvo fue de 31 años RIC [18,0 – 48,0], lo que concuerda con la mayor incidencia de esta patología en las primeras décadas de la vida (23).

El diagnóstico de AA debe ser eminentemente clínico y, las pruebas complementarias de laboratorio y de imagen, no deben sustituir a una correcta anamnesis y exploración física (24). En nuestro trabajo los pacientes podían referir dolor en FID en la anamnesis y no presentar defensa o rebote en la exploración física y, al contrario. Con esto recalamos que la presentación típica de la AA puede no encontrarse hasta en un 55% de los casos (1,2).

Un retraso en el diagnóstico incrementará el riesgo de complicaciones quirúrgicas. Abu-Foul *et al.* demostraron que un retraso en el diagnóstico incrementa las complicaciones, recomendando la apendicectomía temprana; puesto que el tiempo transcurrido desde el inicio de los síntomas hasta la intervención, se relaciona con el estadio evolutivo del apéndice y con la tasa de complicaciones peroperatorias (25). Por otro lado, como confirmaron Lu *et al.* en su estudio (26), un

diagnóstico inadecuado conducirá a un aumento de operaciones innecesarias no exentas de complicaciones. A lo largo del tiempo en la literatura se han aceptado tasas de AN de hasta el 20% (27). En nuestra cohorte hemos registrado una tasa de AN del 11,8%, permaneciendo por debajo del 15-20% de apendicectomías negativas que son aceptadas por la literatura actual (27). Al igual que en el estudio holandés de Lahaye *et al.* (7), en nuestro hospital la realización de pruebas de imagen está establecida en el proceso diagnóstico de todos los pacientes con sospecha de AA; lo que creemos que debería estar asociado a una menor tasa de AN de la registrada. Sin embargo, la ecografía tiene una limitada sensibilidad y especificidad para el diagnóstico de AA, ya que cuenta con limitaciones como observador-dependencia, experiencia del radiólogo, constitución del paciente (obesidad, cicatrices de cirugías previas...) y localización del apéndice (retrocecal) (4,12,23). Además, como se reflejó en el estudio de Kryzauskas *et al.*, la ecografía incrementaría el número de apendicectomías negativas (28). En cuanto a la TC podría disminuir la tasa de apendicectomías negativas por debajo del 10% (2,23), pero su elevado coste y la exposición a radiación ionizante (11) desaconsejan su uso rutinario y la relegan a aquellos casos con ecografía dudosa o sin visualización apendicular pero con alta sospecha de AA y también en pacientes obesos y en aquellos que vuelven a acudir al Servicio de Urgencias con sospecha de AA en un espacio de tiempo de pocas horas o días (4,12). Estas mismas consideraciones se tuvieron en cuenta en nuestros pacientes a la hora de realizar la TC diagnóstica. Al igual que Kartal *et al.* (29), creemos que la combinación de pruebas de laboratorio y ecografía inicial estandarizada, debería usarse en los pacientes con cuadro sospechoso de apendicitis aguda, seguido del TC en los casos dudosos sin diagnóstico confirmado.

Las escalas de probabilidad diagnóstica son una herramienta de aproximación diagnóstica en la evaluación del abdomen agudo de fosa ilíaca derecha. Se ha demostrado que estas escalas proporcionan una evaluación objetiva y estandarizada de los pacientes con sospecha de AA (30) y seleccionan a los pacientes en función de la probabilidad de padecer AA en: pacientes con probabilidad baja (podrían ser dados de alta con observación domiciliaria), pacientes con probabilidad alta (directamente deberían ser intervenidos quirúrgicamente) y pacientes con probabilidad intermedia (aquellos que realmente se beneficiarían de pruebas de imagen complementarias) (31). Se han creado numerosos *scores* y no está del todo claro cuáles de ellos son los mejores. Como demostraron Kularatna *et al.* (30), la heterogeneidad de la metodología y la calidad de los estudios disponibles, impiden la realización de un meta-análisis. No obstante, hay 12 escalas validadas para el diagnóstico de los pacientes con sospecha de AA. En nuestro estudio hemos seleccionado 4 de ellas.

La escala de Alvarado: que fue de las primeras establecidas y está presente en la mayoría de los estudios de validación (30,32). Es tomada como referencia en los estudios de este tipo (33). Al igual que en el estudio de Kariman *et al.* (34) en el que se aplicó la escala de Alvarado a 300 pacientes, obtenemos que un score mayor o igual a 7 otorga una gran probabilidad de padecer apendicitis (93% en el estudio mencionado y 96,6% en el nuestro). El área bajo la curva de esta escala en nuestros pacientes fue de 0,74, considerándose una escala aceptable para predecir el diagnóstico de AA en el global de nuestros pacientes; comparable al ABC que obtuvieron Tekeli *et al.* en un estudio realizado durante 3 años y con 281 pacientes (ABC=0,78) (35). Esta escala consta de 8 ítems objetivos y sencillos de obtener, facilitando por tanto su uso a pie de cama del paciente.

La escala de RIPASA: creada para mejorar la sensibilidad y la especificidad de la escala de Alvarado en la población de Europa del Este, Oriente medio y Asia (36). Esta escala fue seleccionada para nuestro estudio debido a las connotaciones socio-demográficas de nuestra zona; ya que, la comarca de Medina del Campo, presenta múltiples explotaciones agrícolas que dan trabajo, en las diferentes épocas del año, a una amplia población de Europa del Este. La puntuación de 7,5 es el punto de corte más útil para esta escala; el que mejor sensibilidad y especificidad proporciona, como establecieron Butt *et al.* en su estudio (37). Pese a que, en nuestro trabajo, la escala de RIPASA era estadísticamente significativa a la hora de diferenciar entre AA y AN, fue la que obtuvo un menor ABC. Esto entra dentro de lo esperado ya que en nuestra población el porcentaje de pacientes de Europa del Este, Asia u Oriente Medio fue de tan solo un 8,3%. Esta escala consta de 18 ítems e introduce el sexo y la edad como novedades, pero mantiene signos que no siempre quedan recogidos en las historias clínicas (Rovsing, sistemático de orina...).

La escala AIR: en algunos trabajos ha demostrado mejorar la sensibilidad, especificidad y ABC de la escala de Alvarado, aunque ha sido validada en muy pocos estudios (30). En el estudio de De Castro *et al.* (38), se obtuvo un ABC para esta escala de 0,96, consolidándose como una muy buena escala, capaz de reducir los ingresos hospitalarios, las pruebas de imagen, las cirugías innecesarias y, por tanto, los costes que se derivan del proceso diagnóstico – terapéutico de los pacientes con sospecha de AA (39). En nuestra cohorte, el ABC de esta escala fue 0,70, considerándose como una escala de probabilidad aceptable, pero sin mejorar los resultados de Alvarado, por lo que no podemos apoyar los resultados de De Castro *et al.* Esta escala está formada por 12 ítems pero, en nuestra opinión, es poco objetiva a la hora de recoger la exploración física del paciente.

La escala AAS: ideada para la evaluación de pacientes adultos; que son la población más numerosa en nuestra cohorte, y es de las últimas en ser desarrollada. Sammalkorpi *et. al.* (40) mostraron la utilidad de esta escala para disminuir las pruebas de imagen innecesarias, así como la tasa de apendicectomías negativas. Otros trabajos, como los de Bhangu (23) o Kollar (33), han encontrado superioridad de esta escala sobre Alvarado y AIR. En nuestro estudio, el ABC para esta escala fue 0,70, no superando ni a la escala de Alvarado ni al AIR Score. Esta escala está conformada por 20 ítems, también valora la edad y el sexo, pero a nuestro juicio es la más complicada de obtener y tampoco elimina la subjetividad de la anterior escala.

En nuestro estudio estas escalas han presentado una aceptable capacidad predictora para apendicitis aguda. La mediana de puntuación en todas ellas fue superior en los casos de AA frente a los casos de AN. A pesar de que las diferencias entre estas puntuaciones fueron escasas (1 punto para AIR y Alvarado y 0,5 puntos para RIPASA), los resultados fueron estadísticamente significativos. La explicación a esto podría deberse a que las escalas se mostraron como modelos predictores aceptables y no óptimos. No consiguen una discriminación óptima y la mayoría de pacientes se engloban en los grupos de probabilidad intermedia.

Finalmente, nuestro estudio cuenta con diversas limitaciones. La primera de ellas es que se trata de un estudio retrospectivo, basado en datos administrativos y en un solo centro. Otro inconveniente atribuible a este tipo de escalas, es que se han realizado sobre una base de datos de pacientes apendicectomizados, no teniendo en cuenta los que fueron desestimados para la intervención. Se trata, por lo tanto, de un sesgo de selección, ya que no están representados todos los pacientes con dolor

abdominal en FID en la base de datos a partir de la que se ha desarrollado el modelo.

Conclusiones:

Podemos afirmar que las escalas de probabilidad diagnóstica de AA son una herramienta útil para diferenciar el dolor abdominal de FID, que puede facilitar el trabajo en los Servicios de Urgencias y ahorrar tiempo y pruebas innecesarias. Su precisión diagnóstica puede incrementarse con la ayuda, en los casos identificados como probables o dudosos, de la implementación de la ecografía o la TC si la primera no es concluyente, siempre y cuando se tenga acceso a estas pruebas.

La escala de Alvarado, se muestra como la más sencilla, objetiva y fiable de las cuatro escalas a estudio en los pacientes de nuestra cohorte.

Declaraciones y conflicto de intereses: Parte del contenido de trabajo ha sido presentada como Comunicación tipo Póster en el XXI Congreso de la Asociación de Cirujanos de Castilla y León, celebrado los días 6 y 7 de junio de 2019 en Segovia. Miguel Toledano Trincado, coautor de este artículo, es Editor Jefe de la Revista ACIRCAL. El manuscrito ha sido evaluado por otros miembros de los Comités Editorial y Científico. No existen otros conflictos de interés.

Referencias bibliográficas:

1. P. Ansari. Appendicitis Symptoms, Causes & Diagnosis | Merck Manuals - Gastrointestinal Disorders. In: Merck Manuals Professional Edition [Internet]. 2014 [cited 2017 Jul 24]. Available from: <http://www.merckmanuals.com/professional/gastrointestinal-disorders/acute-abdomen-and-surgical-gastroenterology/appendicitis>
2. Kabir SA, Kabir SI, Sun R, Jafferbhoy S, Karim A. How to diagnose an acutely inflamed appendix; a systematic review of the latest evidence. *Int J Surg.* 2017;40:155–62.
3. Jones K, Peña AA, Dunn EL, Nadalo L, Mangram AJ. Are negative appendectomies still acceptable?. *Am J Surg.* 2004;188(6):748–54.
4. Shogilev D, Duus N, Odom S, Shapiro N. Diagnosing Appendicitis: Evidence-Based Review of the Diagnostic Approach in 2014. *West J Emerg Med.* 2014;15(7):859–71.

5. Kirkil C, Karabulut K, Aygen E, Ilhan YS, Yur M, Binnetoğlu K, et al. Appendicitis scores may be useful in reducing the costs of treatment for right lower quadrant pain. *Ulus Travma Ve Acil Cerrahi Derg Turk J Trauma Emerg Surg TJTES*. 2013 Jan;19(1):13–9.
6. Flum DR, Koepsell T. The clinical and economic correlates of misdiagnosed appendicitis: nationwide analysis. *Arch Surg Chic Ill 1960*. 2002 Jul;137(7):799–804; discussion 804.
7. Lahaye MJ, Lambregts DMJ, Mutsaers E, Essers B a. B, Breukink S, Cappendijk VC, et al. Mandatory imaging cuts costs and reduces the rate of unnecessary surgeries in the diagnostic work-up of patients suspected of having appendicitis. *Eur Radiol*. 2015;25(5):1464–70.
8. Wilson EB, Cole JC, Nipper ML, Cooney DR, Smith RW. Computed tomography and ultrasonography in the diagnosis of appendicitis: when are they indicated? *Arch Surg Chic Ill 1960*. 2001;136(6):670–5.
9. Pinto F, Pinto A, Russo A, Coppolino F, Bracale R, Fonio P, et al. Accuracy of ultrasonography in the diagnosis of acute appendicitis in adult patients: review of the literature. *Crit Ultrasound J*. 2013 Jul 15;5 Suppl 1:S2.
10. Liu W, Wei Qiang J, Xun Sun R. Comparison of multislice computed tomography and clinical scores for diagnosing acute appendicitis. *J Int Med Res*. 2015;43(3):341–9.
11. Kim K, Kim YH, Kim SY, Kim S, Lee YJ, Kim KP, et al. Low-dose abdominal CT for evaluating suspected appendicitis. *N Engl J Med*. 2012 26;366(17):1596–605.
12. Reddy SB, Kelleher M, Bokhari SAJ, Davis KA, Schuster KM. A Highly Sensitive and Specific Combined Clinical and Sonographic Score to Diagnose Appendicitis. *J Trauma Acute Care Surg*. 2017;83(4):643-649.
13. 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(8):1843–9.
14. Robert D. Ozde, MD, John R. Goldblum, MD. Odze and Goldblum Surgical Pathology of the GI Tract, Liver, Biliary Tract, and Pancreas: Chapter 18. Inflammatory Disorders of the Appendix. Third Edition. Elsevier; 512–532 p.
15. Mariadason JG, Wang WN, Wallack MK, Belmonte A, Matari H. Negative appendectomy rate as a quality metric in the management of appendicitis: impact of computed tomography, Alvarado score and the definition of negative appendectomy. *Ann R Coll Surg Engl*. 2012;94(6):395–401.
16. Grundmann RT, Petersen M, Lippert H, Meyer F. The acute (surgical) abdomen - epidemiology, diagnosis and general principles of management. *Z Gastroenterol*. 2010;48(6):696–706.
17. Addiss DG, Shaffer N, Fowler BS, Tauxe RV. The epidemiology of appendicitis and appendectomy in the United States. *Am J Epidemiol*. 1990;132(5):910–25.
18. Prystowsky JB, Pugh CM, Nagle AP. Current problems in surgery. Appendicitis. *Curr Probl Surg*. 2005;42(10):688–742.
19. Silberman VA. Appendectomy in a large metropolitan hospital. Retrospective analysis of 1,013 cases. *Am J Surg*. 1981;142(5):615–8.
20. Peltokallio P, Tykkä H. Evolution of the age distribution and mortality of acute appendicitis. *Arch Surg Chic Ill 1960*. 1981;116(2):153–6.
21. Primates P, Goldacre MJ. Appendectomy for acute appendicitis and for other conditions: an epidemiological study. *Int J Epidemiol*. 1994;23(1):155–60.
22. Ricci MA, Trevisani MF, Beck WC. Acute appendicitis. A 5-year review. *Am Surg*. 1991;57(5):301–5.
23. Bhangu A, Søreide K, Di Saverio S, Assarsson JH, Drake FT. Acute appendicitis: modern understanding of pathogenesis, diagnosis, and management. *Lancet*. 2015;386(10000):1278–87.
24. Old JL, Dusing RW, Yap W, Dirks J. Imaging for suspected appendicitis. *Am Fam Physician*. 2005;71(1):71-8.
25. Abu Foul S, Egozi E, Assalia A, Kluger Y, Mahajna A. Is early appendectomy in adults diagnosed with acute appendicitis mandatory? A prospective study. *World J Emerg Surg WJES*. 2019;14:2.
26. Lu Y, Friedlander S, Lee SL. Negative Appendectomy: Clinical and Economic Implications. *Am Surg*. 2016;82(10):1018–22.
27. Paulson EK, Kalady MF, Pappas TN. Suspected Appendicitis. *N Engl J Med*. 2003;348(3):236–42.
28. Kryzauskas M, Danys D, Poskus T, Mikalauskas S, Poskus E, Jotautas V, et al. Is acute appendicitis still misdiagnosed?. *Open Med*. 2016;11(1).
29. Kartal K, Yazıcı P, Ünlü TM, Uludağ M, Mihmanlı M. How to avoid negative appendectomies: Can US achieve this? *Ulus Travma Ve Acil Cerrahi Derg Turk J Trauma Emerg Surg TJTES*. 2017 Mar;23(2):134–8.
30. 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(7):1769–81.

31. Alvarado A. How to improve the clinical diagnosis of acute appendicitis in resource limited settings. *World J Emerg Surg.* 2016;11(1).
32. Alvarado A. A practical score for the early diagnosis of acute appendicitis. *Ann Emerg Med.* 1986 May;15(5):557–64.
33. Kollár D, McCartan DP, Bourke M, Cross KS, Dowdall J. Predicting Acute Appendicitis? A comparison of the Alvarado Score, the Appendicitis Inflammatory Response Score and Clinical Assessment. *World J Surg.* 2015;39(1):104–9.
34. Kariman H, Shojaee M, Sabzghabaei A, Khatamian R, Derakhshanfar H, Hatamabadi H. Evaluation of the Alvarado score in acute abdominal pain. *Ulus Travma Ve Acil Cerrahi Derg Turk J Trauma Emerg Surg TJTES.* 2014 Mar;20(2):86–90.
35. Tekeli MT, Ilhan E, Ureyen O, Senlikci A, Yeldan E, Ozturk M, et al. How much Reliable Is Alvarado Scoring System in Reducing Negative Appendectomy? *Indian J Surg.* 2017;79(2):106–10.
36. Chong CF, Adi MIW, Thien A, Suyoi A, Mackie AJ, Tin AS, et al. Development of the RIPASA score: a new appendicitis scoring system for the diagnosis of acute appendicitis. *Singapore Med J.* 2010;51(3):220–5.
37. Butt MQ, Chatha SS, Ghumman AQ, Farooq M. RIPASA score: a new diagnostic score for diagnosis of acute appendicitis. *J Coll Physicians Surg--Pak JCPSP.* 2014;24(12):894–7.
38. de Castro SMM, Ünlü ç., Steller EPh, van Wagensveld BA, Vrouenraets BC. Evaluation of the Appendicitis Inflammatory Response Score for Patients with Acute Appendicitis. *World J Surg.* 2012;36(7):1540–5.
39. Andersson M, Kolodziej B, Andersson RE, STRAPPSCORE Study Group. Randomized clinical trial of Appendicitis Inflammatory Response score-based management of patients with suspected appendicitis: Appendicitis Inflammatory Response score-based management of suspected appendicitis. *Br J Surg.* 2017;104(11):1451-1461
40. Sammalkorpi HE, Mentula P, Savolainen H, Leppäniemi A. The Introduction of Adult Appendicitis Score Reduced Negative Appendectomy Rate. *Scand J Surg.* 2017;106(3):196-201.