



ARTÍCULO ORIGINAL

Post-Ureteroscopic Lesion Scale: limitaciones en su aplicabilidad clínica

R. Polo^a, À. Canós-Nebot^b, J.P. Caballero-Romeu^b, P. Caballero^{c,*}, J.A. Galán-Llopis^b, F. Soria^d, J.E. de la Cruz-Conty^e y J. Tuells^c

^a Servicio de Urología, Hospital Marina Baixa, Denia, Valencia, España

^b Servicio de Urología, Hospital General de la Universidad de Alicante, Instituto de Investigación Sanitaria y Biomédica (ISABIAL – Fundación FISABIO), Alicante, España

^c Departamento de Enfermería Comunitaria, Medicina Preventiva y Salud Pública e Historia de la Ciencia, Universidad de Alicante, Alicante, España

^d Unidad de Cirugía Experimental, Hospital Universitario Ramón y Cajal, Madrid, España

^e Departamento de Fisiología-Laboratorio del Dr. Adebisi, Centro de Ciencias de la Salud de la Universidad de Tennessee (UTHSC), Memphis, Tennessee, EE. UU.

Recibido el 13 de junio de 2023; aceptado el 14 de agosto de 2023

Disponible en Internet el 20 de septiembre de 2023

PALABRAS CLAVE

Lesión ureteral;
Escala;
Concordancia;
Ureteroscopia;
Micro-ureteroscopia;
Post-Ureteroscopic
Lesion Scale

Resumen

Objetivo: Analizar el nivel de concordancia de la *Post-Ureteroscopic Lesion Scale* (PULS), y examinar las consecuencias de su aplicación en la práctica clínica con datos estadísticos más fiables que los utilizados en el trabajo original.

Métodos: Se realizaron 14 ureteroscopias (URS) y 14 micro-ureteroscopias (micro-URS) en 14 cerdos hembra. Todos los procedimientos se grabaron en vídeo y se realizó un análisis anatómopatológico en cada uréter. Dieciséis urólogos (9 endourólogos y 7 urólogos generales) y 4 médicos internos residentes evaluaron las lesiones ureterales según la escala PULS, con grados 0, 1 y ≥ 2 . La concordancia se calculó mediante porcentajes, el Coeficiente W de Kendall, el índice kappa de Fleiss y el alfa de Krippendorff. La concordancia entre evaluadores se calculó con la Correlación de Spearman y el Coeficiente kappa de Cohen.

Resultados: El porcentaje de concordancia fue del 11,1%. Los coeficientes se clasificaron como bajos o muy bajos, y encontramos una mayor concordancia entre los evaluadores más inexpertos. Por otro lado, no hubo acuerdo entre el 50% de los evaluadores.

Conclusiones: La baja concordancia entre evaluadores, la especificidad de la PULS y la correlación clínico/patológica sugieren dificultades del uso de esta escala y una curva de aprendizaje probablemente larga.

© 2023 Los Autores. Publicado por Elsevier España, S.L.U. en nombre de AEU. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

* Autor para correspondencia.

Correo electrónico: pablo.caballero@ua.es (P. Caballero).

KEYWORDS

Ureteral lesion;
Scale;
Concordance;
Ureteroscopy;
Micro-ureteroscopy;
Post-Ureteroscopic
Lesion Scale

Post-Ureteroscopic Lesion Scale to determine ureteral wall damage, not so easy to employ

Abstract

Objective: To analyze the level of agreement of the Post-Ureteroscopic Lesion Scale (PULS) and the consequences on its application in clinical practice with more reliable statistical data than the one used in the original work.

Methods: 14 URS and 14 micro-URS were performed in 14 female porcine model. All the procedures were video recorded and an anatomopathological analysis was performed in each ureter. Sixteen urologists (9 endourologists and 7 general urologists) and 4 residents evaluated the ureteral lesions according to the PULS, with degrees 0, 1 and ≥ 2 . The agreement was calculated with percentages, Kendall's W coefficient and the indicators Fleiss' Kappa and Krippendorff's Alpha, while the inter-rater agreement was calculated with Spearman's correlation and Cohen's Kappa.

Results: The percent of agreement was 11.1%. The coefficients were likewise classified as low or very low, with the greatest agreement found among the inexperienced. Also, 50% of the raters did not agree with the rest.

Conclusions: The low inter-rater agreement, the specificity of the PULS and the clinical-pathological correlation suggests that this scale is not simple, and probably has a long learning curve.

© 2023 The Authors. Published by Elsevier España, S.L.U. on behalf of AEU. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introducción

La ureteroscopia (URS) es una técnica endourológica esencial en el diagnóstico y tratamiento de patología unilateral. La URS presenta ventajas frente a la litotricia extracorpórea por ondas de choque (LEOC) en el manejo de la litiasis alojada en el uréter distal, y en la necesidad de un número menor de sesiones de tratamiento para conseguir su resolución¹. Sin embargo, la URS se asocia a una estancia hospitalaria más larga y a un número elevado de complicaciones².

La clasificación de las lesiones tiene 2 objetivos fundamentales. Por un lado, permite homogeneizar el tratamiento requerido según la gravedad de la lesión, y por otro, facilita la comparación de los resultados de los diferentes grupos al publicarlos.

Las lesiones ureterales, que pueden haberse producido como consecuencia de la aplicación de técnicas endourológicas, son clasificadas por la *American Association for the Surgery of Trauma* (AAST) según la escala *Organ Injury Severity Scale*³. Las lesiones ureterales también se pueden clasificar según el procedimiento realizado para su tratamiento, de acuerdo con la Clasificación de Clavien Dindo modificada^{4,5}. Esta, al igual que la escala de la AAST, constituye una clasificación válida para todas las lesiones ureterales producidas por traumatismos externos o iatrogénicos.

En 2012, Schoenthaler et al. desarrollaron la escala *Post-Ureteroscopic Lesion Scale* (PULS). Esta escala fue creada con el objetivo de clasificar las lesiones ureterales presentes tras una URS, independientemente de su existencia antes del procedimiento quirúrgico. La escala asociaba un grado de lesión con la necesidad o no de colocar un *stent* ureteral, así como con el tiempo de permanencia de este⁶⁻⁸. La

escala PULS ofrece 6 grados, de 0 a 5, como se muestra en la [figura suplementaria 1](#).

Desde la fecha de su publicación, el uso y presencia de la PULS han ido aumentando paulatinamente en la investigación clínica de las lesiones postureteroscopia, como herramienta para clasificar las lesiones producidas después del procedimiento. Según la *Web of Science*, el número de referencias relacionadas con esta escala ha ido aumentando progresivamente, pasando de 2 citas en 2015 a 61 en septiembre de 2020. Además, con el desarrollo de los endoscopios flexibles y las vainas de acceso ureteral (VAU), la escala PULS ha desempeñado un papel relevante en la evaluación de lesiones relacionadas con su uso^{6,9,10}.

En 2014, los autores de la escala mostraron una cierta concordancia entre evaluadores de la PULS con la prueba W de Kendall, trabajando con 37 evaluadores y 100 vídeos de procedimientos quirúrgicos⁷. En 2018, los mismos autores estudiaron el grado de concordancia alcanzado entre el personal asistencial y los cirujanos. Sin embargo, esta concordancia global ha sido puesta en duda, ya que cada paciente fue evaluado por un asistente y un cirujano, no por el conjunto de participantes de cada grupo¹¹.

En la actualidad, existen mejores indicadores para analizar la concordancia entre evaluadores que la prueba W de Kendall¹². Con base en el índice kappa de Cohen para evaluar la concordancia entre 2 evaluadores, puede utilizarse el coeficiente kappa de Fleiss¹³ o el coeficiente kappa de Light¹⁴, pudiendo aplicarse en caso de que todos los evaluadores examinen todos los casos. Otro indicador de concordancia entre evaluadores es el coeficiente alfa de Krippendorff¹⁵.

Por lo tanto, el objetivo de este estudio es analizar la concordancia de la escala PULS entre evaluadores, en un modelo experimental para evaluar su aplicación en la práctica clínica.

Material y métodos

Vídeos

Los vídeos corresponden a un estudio experimental realizado con 14 cerdos hembra de raza Large White, con un peso de entre 30 y 35 kg. Se excluyeron los machos porque suelen tener una conformación anatómica diferente, que los hace inadecuados para estudios endourológicos. El riñón porcino es multipapilar, al igual que el humano. Los modelos porcinos que pesan entre 30-35 kg son similares en tamaño a un humano de aproximadamente 70 kg. Además, el uréter del cerdo tiene una longitud aproximada de 22-26 cm y un diámetro interno de 4-6 Fr, muy similar al de un humano¹⁶. Se descartó cualquier anomalía bajo anestesia general. La fase experimental fue realizada por dos endourólogos con experiencia tanto en URS como en micro-ureteroscopia (m-URS). Cada cirujano realizó 7 ureteroscopias y 7 m-URS¹⁷.

Los vídeos mostraban 14 URS y 14 m-URS. Cada vídeo tenía una duración de aproximadamente 30 s, e incluía el paso final de la URS, mostrando 17-18 cm de uréter distal. De los primeros 28 vídeos, se excluyó el estudio de uno de ellos, correspondiente a una m-URS, debido a problemas de visualización.

El estudio fue aprobado por el Comité Ético de Experimentación Animal. La fase experimental se llevó cabo siguiendo las directrices relativas al cuidado de los animales utilizados para experimentación, así como las directrices de la Unión Europea relativas a la protección de animales para fines científicos.

Evaluadores

Se seleccionaron 16 urólogos (9 de ellos endourólogos) y 4 médicos internos residentes. Este fue el tamaño de muestra del estudio, ya que nuestro objetivo era analizar la concordancia entre sus respuestas. Todos los participantes recibieron las mismas instrucciones para la clasificación de las lesiones (figura suplementaria 1), que fueron extraídas de los estudios publicados por Schoenthaler et al.^{6,7}. Algunos de los endourólogos estaban familiarizados con la escala antes del estudio.

Evaluación

Los evaluadores recibieron un cuestionario para determinar el grado de las lesiones ureterales. En el cuestionario se consignaba una sola respuesta para cada vídeo. Había 6 opciones disponibles, dado que la puntuación PULS iba de 0 a 5⁷.

Para categorizar las lesiones por perforación (grado > 2) se requiere una pielografía anterógrada con verificación fluoroscópica. Sin embargo, en nuestro estudio esto no pudo realizarse, por lo que todas las lesiones de grado superior a 2, se incluyeron en un mismo grupo denominado «grado > 0 = 2».

Otras variables. Evaluación patológica

Además de los vídeos, se realizó un análisis anatomopatológico. El estudio experimental se completó con la extirpación

del tracto urinario en bloque para su estudio anatomopatológico. El patólogo, que desconocía el grupo del animal, realizó el estudio anatomopatológico del uréter. Se utilizó una tinción de hematoxilina y eosina en las muestras, y se empleó una escala de cicatrización validada en la que se evaluaron 5 parámetros, con puntuaciones de 0 a 3, donde 0 era equivalente a la ausencia de cambios, y 3 a la presencia de alteraciones graves¹⁷. Estos parámetros fueron: inflamación, fibrosis de lámina propia, fibrosis de capa muscular, integridad muscular y alteraciones de la membrana serosa. Los vídeos se clasificaron según la escala PULS, basándose en los resultados de la anatomía patológica. Se utilizaron puntuaciones histológicas como referencia para establecer una correspondencia entre las puntuaciones PULS y la lesión ureteral microscópica.

Análisis de datos

Para analizar la concordancia, se calcularon el porcentaje de concordancia, la W de Kendall y su nivel de significación, y los siguientes índices de concordancia: kappa de Fleiss, kappa de Light, y alfa de Krippendorff. Para evaluar la concordancia con los índices kappa de Fleiss y Light, los valores se consideraron bajos entre 0,21 y 0,40, normales entre 0,41 y 0,60, buenos entre 0,61 y 0,80, y excelentes por encima de 0,81¹⁸. Para el alfa de Krippendorff, los valores inferiores a 0,67 se consideraron como desacuerdo, entre 0,67 y 0,80 concordancia considerable, y superiores a 0,80 acuerdo perfecto¹⁵.

Para determinar la concordancia entre dos observadores, se calcularon el coeficiente de Correlación de Spearman y el kappa de Cohen, que se interpretaron en la misma escala que el kappa de Fleiss y Light. Así mismo, se calculó para ambos indicadores, el porcentaje de compañeros que estaba significativamente de acuerdo.

Resultados

La tabla 1 muestra los indicadores de concordancia calculados el porcentaje de acuerdo calculado fue del 11,1%, aunque este valor sobreestima el acuerdo ya que no tiene en cuenta el acuerdo debido al azar únicamente. Además, el kappa de Fleiss, así como la corrección de Light para modelos completos, no supera los 0,40 para todos los evaluadores, y solo se aproxima a 0,40 entre los evaluadores más novatos, que mostraron un mayor acuerdo entre ellos. Los valores alfa de Krippendorff se clasificaron como «muy bajos» para todos los evaluadores de los diferentes grupos, y el acuerdo volvió a ser mayor entre los más inexpertos. El clásico coeficiente W de Kendall fue significativo, como cabía esperar, aunque sus valores no fueron excesivamente elevados.

En cuanto a la concordancia por pares, se observa que tanto para el kappa de Cohen como para la Correlación de Spearman, el porcentaje de parejas de evaluadores con estos indicadores en valores distintos de cero, giran en torno al 57,2 y 54,4% para la totalidad de los evaluadores (tabla 2). Sin embargo, en el caso del kappa de Cohen, y siguiendo la clasificación propuesta, solo 8 parejas de evaluadores superan valores de 0,60, considerados como valores de concordancia buena (tabla 2). Siguiendo este criterio, solo habría cierta concordancia entre los evaluadores 1, 2 y

Tabla 1 Concordancia

Indicadores	Total	Principiantes	Endoscopistas expertos	Otros
Número de evaluadores	20	4	9	7
Porcentaje de concordancia	11,1	51,9	37	22,2
Kappa de Fleiss	0,24*	0,37*	0,25*	0,17*
Kappa de Light	0,26	0,4	0,26	0,20
Alfa de Krippendorff	0,24	0,37	0,25	0,17
Coefficiente W de Kendall por parejas	0,40*	0,67*	0,43*	0,45*
Porcentaje significativo Kappa de Cohen	57,2	50,0	50,0	52,4
Porcentaje significativo Correlación de Spearman	54,4	100	58,3	61,9

* Significativo $p < 0,01$.

16, el evaluador 20 con los evaluadores 4 y 12, el evaluador 11 con los evaluadores 17 y 1, y el evaluador 6 con el 19. Por lo tanto, según el kappa de Cohen, 10 de los 20 evaluadores del sondeo no estaban particularmente de acuerdo con el resto de sus compañeros.

Según el estudio patológico (tabla 3), 12 de los 27 vídeos (44,4%) tenían grado 1 en la escala PULS, y el resto tenía un grado 0. Al comparar estos resultados con las observaciones de los evaluadores, se obtuvieron discrepancias sustanciales, especialmente cuando la puntuación patológica se clasificó como grado 1. En este caso, solo el 23,5% de las clasificaciones realizadas por los evaluadores coincidía; la mayoría de las discrepancias se debían a infraestimación (67,5%), y solo el 8,7% a sobreestimación. Esta concordancia aumentó al 80,7% en el caso del grado 0 para la clasificación realizada a través de la puntuación patológica, resultando en una sobreestimación del 19,3% (tabla 3).

En cuanto a la sensibilidad y a la especificidad registradas por el conjunto de los evaluadores para diagnosticar un grado 0 con la PULS, se obtuvieron 242 resultados positivos (clasificados como grado 0) de 300 visualizaciones de vídeos clasificados como grado 0 según el análisis histológico, y 162 resultados positivos (clasificados como grado 0/1) de 240 vídeos clasificados como grado 1 según la anatomía patológica. Por lo tanto, se obtuvo una sensibilidad de 0,807 y una especificidad de 0,325. Los valores predictivos se muestran en la figura 1. Se observa un comportamiento sesgado para ambos valores predictivos, casi sin curva.

Discusión

La escala PULS pretende ser un sistema de clasificación sencillo y accesible, no solo para especialistas sino también para médicos residentes. Bajo la premisa de ser una escala intuitiva y sencilla, se busca un alto nivel de concordancia, para hacerla altamente reproducible.

Los autores de la escala limitan su utilidad a tres aspectos: establecer la necesidad o no de stent ureteral, la duración recomendada del uso del stent, o la necesidad de reparación inmediata del uréter a causa de una lesión grave. Así, la PULS se considera una escala que contribuye a complementar otras existentes, como la AAST, el sistema Clavien-Dindo modificado o el sistema de gradación Satava^{19,20}. El estudio de la utilidad y fiabilidad de esta escala es de gran interés, ya que los *stents* ureterales

disminuyen la calidad de vida hasta en el 80% de los pacientes portadores de estos dispositivos²¹.

Los resultados de concordancia entre evaluadores comunicados por Schoenhaler et al.⁶ se limitan a la W de Kendall y su significación. Dado que la prueba W de Kendall puede proporcionar un contraste si hay al menos un evaluador que está de acuerdo con otro del grupo, un aumento del número de evaluadores aumentaría la probabilidad de obtener resultados significativos, por lo que el valor de W de Kendall sería más interesante que su significación. Los valores obtenidos en este estudio también son significativos, pero ligeramente inferiores. No obstante, no se ha desarrollado una escala para evaluar la concordancia en función del valor de la W de Kendall, por lo que se han utilizado otros indicadores, como los que aplicamos en nuestro estudio: el kappa de Fleiss, el kappa de Light o el alfa de Krippendorff. Todos ellos mostraron una concordancia baja o moderada tanto para el conjunto de evaluadores como para los 3 grupos clasificados.

Otros indicadores calculados a partir del análisis por parejas de evaluadores, como el kappa de Cohen y la Correlación de Spearman, también presentaron una baja concordancia. Que el 50% de los evaluadores no obtuvieran valores de kappa de Cohen superiores a 0,6 con respecto a los demás evaluadores corroboró la falta de concordancia en la aplicación de la escala PULS.

El primer estudio en el que se compararon los resultados de la PULS con los hallazgos histopatológicos fue el realizado por Lildal et al.²². Los resultados de este estudio experimental en un modelo porcino para cirugía retrógrada intrarrenal (RIRS) con utilización de vaina de acceso ureteral (VAU) demostraron que la PULS subestimaba las lesiones ureterales en al menos 1 grado con respecto a los hallazgos histopatológicos. Bozzini et al.⁹ revelaron en su estudio que no había un mayor número de lesiones ureterales en las RIRS que utilizaban VAU con respecto a las que no la utilizaban, habiéndose usado en este estudio la PULS para la evaluación de las lesiones ureterales. Según nuestro estudio, los resultados de Bozzini et al. podrían tener un significado diferente dada la escasa concordancia entre los distintos índices.

Al comparar los resultados de los evaluadores con el análisis histopatológico, la infraestimación de grado 1 a grado 0 daría lugar a una falta de atención médica para el 67,5% de los pacientes que realmente la necesitaban. Por otro lado, la sobrevaloración de los casos de grado 0 en grado 1 resultaría en que el 19,3% de los casos recibiría un sobretratamiento, con posibles molestias secundarias provocadas por el uso

Tabla 2 Kappa de Cohen y Correlación de Spearman

Correlación de Spearman	Kappa de Cohen																			
	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15	R16	R17	R18	R19	R20
R1		0,763*	0,252	0,287	0,329*	0,140*	0,232	0,330*	-0,024	0,428*	0,618*	0,301*	0,369*	0,028	0,338*	0,614*	0,563*	0,217*	0,140*	0,344*
R2	0,774*		0,327*	0,491*	0,329*	0,140*	0,232	0,330*	-0,024	0,428*	0,428*	0,301*	0,369*	0,028	0,338*	0,614*	0,563*	0,217*	0,140*	0,563*
R3	0,667*	0,594*		0,278	0,391*	0,121*	-0,104	0,359*	-0,062	0,361*	0,289	0,151	0,256*	0,136	0,359*	0,163	0,340*	0,204	0,121*	0,193
R4	0,423*	0,623*	0,397*		0,285	0,221*	0,032	-0,055	0,244	0,382*	0,206	0,314*	0,156	0,049	0,133	0,139	0,320	0,137	0,221*	0,709*
R5	0,455*	0,455*	0,403*	0,293		0,150*	0,184	0,304	0,137	0,536*	0,536*	0,201	0,241*	0,281*	0,345*	0,224	0,184	0,197	0,150*	0,184
R6	0,485*	0,485*	0,368	0,429*	0,384*		0,115	0,081	0,377*	0,192*	0,192*	0,481*	0,377*	0,258*	0,029	0,176*	0,258*	0,077	1,000*	0,258*
R7	0,268	0,268	-0,025	0,165	0,425*	0,452*		0,265	0,182	-0,007	-0,007	0,247	-0,034	-0,112	0,002	0,296	0,076	0,147*	0,115	0,076
R8	0,376	0,376	0,392*	0,072	0,374	0,411*	0,303		-0,096	0,408*	0,408*	0,166	0,318*	0,013	0,530*	0,400*	0,359*	0,252	0,081	-0,007
R9	0,143	0,143*	-0,043	0,295*	0,189	0,530*	0,338	0,048		0,200	0,000	0,524*	0,161	0,073	-0,182	0,010	0,061	0,225*	0,377*	0,530*
R10	0,568*	0,568*	0,453*	0,411	0,520*	0,411*	0,124	0,513*	0,258		0,491*	0,477*	0,513*	0,013	0,330*	0,303*	0,449*	0,314*	0,192*	0,449*
R11	0,759*	0,568*	0,515*	0,241*	0,520*	0,411*	0,124	0,513*	0,048	0,513*		0,267*	0,513*	0,372*	0,464*	0,303*	0,633*	0,314*	0,192*	0,265
R12	0,485*	0,485*	0,237	0,390	0,290	0,600*	0,432*	0,356	0,554*	0,593*	0,356		0,215	0,125	0,019	0,368*	0,369*	0,175*	0,481*	0,621*
R13	0,646*	0,646	0,395*	0,311	0,384*	0,530*	0,106	0,451*	0,187	0,677*	0,677*	0,247		0,314*	0,172	0,301*	0,540*	0,191*	0,377*	0,196
R14	0,299	0,299*	0,218	0,193	0,378	0,451*	0,015	0,117	0,106	0,151	0,494*	0,170	0,392*		0,164	0,189	-0,011	0,114	0,258*	0,090
R15	0,501*	0,501*	0,563*	0,278	0,403*	0,372	0,056	0,628*	-0,110	0,481*	0,628*	0,207	0,460*	0,385*		0,276*	0,337*	0,186	0,029	0,071
R16	0,638*	0,638*	0,303	0,295	0,377	0,530*	0,338	0,468*	0,188	0,468*	0,468*	0,554*	0,447*	0,328	0,460*		0,416*	0,106	0,176*	0,182
R17	0,690*	0,690*	0,487*	0,352	0,206	0,452*	0,212	0,482*	0,107	0,482*	0,661*	0,432*	0,550*	0,015	0,542*	0,569*		0,330*	0,258*	0,382*
R18	0,419*	0,419*	0,287	0,183	0,144	0,371	0,292	0,367	0,393*	0,367	0,367	0,371	0,393*	0,392*	0,353	0,510*	0,447*		0,077*	0,330*
R19	0,485*	0,485*	0,368	0,429*	0,384*	1,000*	0,452*	0,411*	0,530*	0,411*	0,411*	0,600*	0,530*	0,451*	0,372	0,530*	0,452*	0,371		0,258*
R20	0,479*	0,690*	0,339	0,725*	0,206	0,452*	0,212	0,124	0,569*	0,482*	0,303	0,693*	0,355	0,242	0,218	0,338	0,409*	0,447*	0,452*	

R1-R4: evaluadores principiantes; R5-R13 endoscopistas expertos; R14-R20: otros urólogos.

* Significativo p < 0,05

Tabla 3 Clasificación de vídeos según evaluadores y anatomía patológica. Porcentaje de concordancia entre evaluadores y anatomía patológica

Video	0 % (n)	I % (n)	II o +II % (n)	AP			% PAA
				Epitelial	Muscular	PULS (AP)	
1	90 (18)	5 (1)	5 (1)	1	0	0	0,9
2	100 (20)			1	0	0	1
3	100 (20)	—	—	1	0	0	1
4	65 (13)	35 (7)	—	2	0	1	0,35
5	75 (15)	25 (5)	—	2	0	1	0,25
6	65 (13)	35 (7)	—	3	0	1	0,35
7	40 (8)	60 (12)	—	3	0	1	0,6
8	95 (19)	5 (1)	—	1	0	0	0,95
9	95 (19)	5 (1)	—	1	0	0	0,95
10	95 (19)	5 (1)	—	1	0	0	0,95
11	85 (17)	15 (3)	—	2	0	1	0,15
12	25 (5)	75 (15)	—	1	0	0	0,25
13	80 (16)	20 (4)	—	1	0	0	0,8
14	—	—	100 (20)	3	0	1	0
15	70 (14)	25 (5)	5 (1)	2	0	1	0,25
16	75 (15)	25 (5)	—	1	0	0	0,75
17	70 (14)	30 (6)	—	3	0	1	0,3
18	65 (13)	35 (7)	—	1	0	0	0,65
19	80 (16)	15 (3)	5 (1)	1	0	0	0,8
20	80 (16)	20 (4)	—	3	0	1	0,2
21	100 (20)	—	—	1	0	0	1
22	80 (16)	20 (4)	—	2	0	1	0,2
23	90 (18)	10 (2)	—	2	0	1	0,1
24	90 (18)	10 (2)	—	2	0	1	0,1
25	90 (18)	10 (2)	—	1	0	0	0,9
26	40 (8)	45 (9)	15 (3)	1	0	0	0,4
27	80 (16)	20 (4)	—	1	0	0	0,8
						Media	0,55

AP: anatomía patológica; PULS: post-ureteroscopy lesion scale; % PAA: porcentaje de concordancia en la anatomía patológica: Grado 0, I, II o +II.

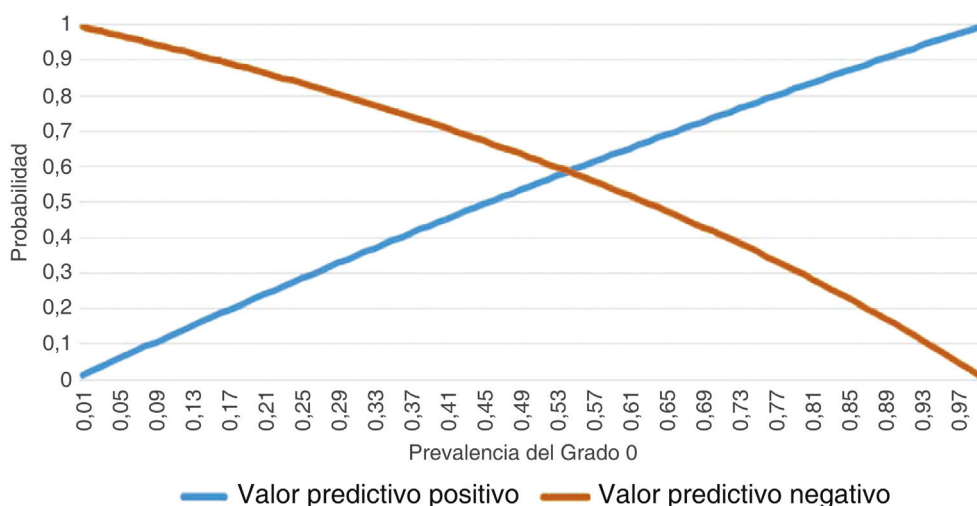


Figura 1 Valores predictivos positivos y negativos para la prevalencia del grado 0.

del stent ureteral (hematuria, síntomas del tracto urinario inferior, infecciones, etc.)⁶. Teniendo en cuenta:

La mencionada infravaloración de lesiones ureterales. La ineficacia de la escala PULS para diferenciar entre grados 0 y 1. Y el hecho de que Schoenthaler et al. proponen que los grados de 0 a 2 no sean declarados como casos de complicación postoperatoria. Consideramos que la escala PULS puede constituir una herramienta limitada para identificar las complicaciones relacionadas con procedimientos ureteroscópicos⁷.

Una de las principales desventajas de la escala PULS es que no tiene en cuenta el estado del uréter al inicio de la URS. En los pacientes que se han sometido a un procedimiento endourológico previo o a una LEOC en el mismo uréter, la categorización de la lesión posterior a la URS puede sobreestimarse al utilizar la PULS.

Por último, nos gustaría considerar algunos puntos que podrían mejorar la concordancia en las evaluaciones de lesiones ureterales mediante la escala PULS. En primer lugar, es necesario desarrollar un programa de formación y seguimiento para que los urólogos aprendan a utilizar la escala correctamente. Es importante generalizar el uso de la pielografía ante cualquier sospecha de lesión ureteral durante la URS, y saber evaluar la gravedad de la lesión en relación con la cantidad de contraste extravasado. En relación con esto, podría ser útil el desarrollo de herramientas tecnológicas que permitan automatizar el proceso de evaluación y reducir el riesgo de error en la categorización del grado de lesión. En este sentido, el desarrollo de herramientas software que permitan comparar imágenes de pielografía podría mejorar la evaluación imparcial de las lesiones ureterales.

Limitaciones

A pesar de los resultados relativos a la concordancia entre características histológicas y clínicas y su apoyo a determinados hallazgos clínicos, debemos considerar algunas limitaciones en este aspecto.

En primer lugar, no disponemos de información sobre las características histológicas previas de los tejidos ureterales. Dicha información nos permitiría analizar los cambios generados por la propia intervención.

Por otro lado, el análisis histopatológico fue exhaustivo, pero también focal, por lo que se trata de una técnica transversal o estática, mientras que la visualización del uréter mediante vídeos se considera una técnica longitudinal o dinámica. Por este motivo, en los casos en los que la PULS podría sobreestimar los resultados o la sensibilidad y especificidad de la escala, no podemos afirmar que difieran, debido a la ausencia de un análisis de la porción ureteral que podría estar dañada.

Por último, el análisis histopatológico no permite distinguir entre puntuaciones de 3, 4 o 5. La diferencia entre grados es relevante en la práctica clínica, debido a que cada uno implica necesidades terapéuticas diferentes.

Conclusión

En nuestro estudio, la escala PULS presentó una baja concordancia entre evaluadores, una escasa correlación con los hallazgos histopatológicos y una baja especificidad. Además,

su uso adecuado puede implicar una larga curva de aprendizaje, y puede no ser tan fácil de utilizar como los autores podían prever inicialmente.

Aprobación ética

Los datos utilizados proceden de fuentes secundarias, por lo que no requieren la aprobación de un comité de ética.

Financiación

El presente estudio no recibió ningún tipo de financiación.

Autorías

Rebeca Polo: Desarrollo del protocolo/proyecto y recolección de datos.

Àngela Canós-Nebot: Desarrollo del proyecto y redacción del manuscrito.

Juan Pablo Caballero-Romeu: Desarrollo del proyecto, redacción del manuscrito y coordinación.

Juan Antonio Galán-Llopis: Desarrollo del proyecto y edición del manuscrito.

Pablo Caballero: Análisis de datos, desarrollo del proyecto y redacción del manuscrito.

Federico Soria: Recolección de datos y edición del manuscrito.

Julia E. de la Cruz-Conty: Recolección de datos.

José Tuells: Desarrollo del proyecto y edición del manuscrito.

Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

Anexo. Material adicional

Se puede consultar material adicional a este artículo en su versión electrónica disponible en [doi:10.1016/j.acuro.2023.08.004](https://doi.org/10.1016/j.acuro.2023.08.004).

Bibliografía

1. Drake T, Grivas N, Dabestani S, Knoll T, Lam T, MacLennan S, et al. What are the Benefits and Harms of Ureteroscopy Compared with Shock-wave Lithotripsy in the Treatment of Upper Ureteral Stones? A Systematic Review. *Eur Urol*. 2017;72:772–86.
2. Aboumarzouk OM, Kata GS, Keeley FX, McClinton SNG. Extracorporeal shock wave lithotripsy (ESWL) versus ureteroscopy management for ureteric calculi. *Cochrane Database Syst Rev*. 2012;24:CD006029.
3. Moore EE, Cogbill TH, Jurkovich GJ, McAninch JW, Champion HR, Gennarelli TA, et al. Organ injury scaling. III: Chest wall, abdominal vascular, ureter, bladder, and urethra. *J Trauma*. 1992;33:337–9.
4. Dindo D, Demartines N, Clavien PA. Classification of surgical complications: A new proposal with evaluation in a cohort of 6336 patients and results of a survey. *Ann Surg*. 2004;240:205–13.

5. Ataulaevich Akilov F, Iskandarovich Giyasov S, Tursunovich Mukhtarov S, Raufovich Nasirov F, Fatikhovich Alidjanov J. Applicability of the Clavien-Dindo grading system for assessing the postoperative complications of endoscopic surgery for nephrolithiasis: A critical review. *Turkish J Urol*. 2014;39:153–60.
6. Schoenthaler M, Wilhelm K, Kuehhas FE, Farin E, Bach C, Buchholz N, et al. Postureteroscopic Lesion Scale: A New Management Modified Organ Injury Scale—Evaluation in 435 Ureteroscopic Patients. *J Endourol*. 2012;26:1425–30.
7. Schoenthaler M, Buchholz N, Farin E, Ather H, Bach C, Bach T, et al. The Post-Ureteroscopic Lesion Scale (PULS): A multicenter video-based evaluation of inter-rater reliability. *World J Urol*. 2014;32:1033–40.
8. Traxer O, Thomas A. Prospective evaluation and classification of ureteral wall injuries resulting from insertion of a ureteral access sheath during retrograde intrarenal surgery. *J Urol*. 2013;189:580–4.
9. Bozzini G, Bevilacqua L, Besana U, Calori A, Pastore A, Romero Otero J, et al. Ureteral access sheath-related injuries vs. post-operative infections Is sheath insertion always needed? A prospective randomized study to understand the lights and shadows of this practice. *Actas Urol Esp (Engl Ed)*. 2021;45:576–81.
10. Cristallo C, Santillán D, Tobia I, Tirapegui FI, Daels FPGM. Flexible ureteroscopy without ureteral access sheath. *Actas Urol Esp (Engl Ed)*. 2022;46:354–60.
11. May M, Schönthaler M, Gilfrich C, Wolff I, Peter J, Miernik A, et al. Interrater-Übereinstimmung und klinischer Nutzen der ‘‘Post-Ureteroscopic Lesion Scale’’ (PULS) zur Graduierung von intraoperativen Harnleiterverletzungen einer Ureterorenoskopie: Ergebnisse der deutschen prospektiv-multizentrischen BUSTER-Studie. *Urologe*. 2018;57:172–80.
12. Matthias Gammer, Jim Lemon. Ian Fellows PS. Package ‘‘irr’’ Various coefficients of interrater reliability and agreement. 2012. Report No.: 0.84.
13. McGraw KO, Wong SP. Forming inferences about some intraclass correlations coefficients: Correction. *Psychol Methods*. 1996;1:390, 390.
14. Conger AJ. Integration and generalization of kappas for multiple raters. *Psychol Bull*. 1980;88:322–8.
15. Krippendorff K. Content analysis: An introduction to its methodology. 2.nd Edit SAGE Publications; 2004. p. 412.
16. Schwalb D, Eshghi M, Cord J, Evans R, Braga E, Franco I, et al. The Minipig as a Practical Endourologic Model. *J Endourol*. 1989;3:85–90.
17. Caballero-Romeu JP, Galán-Llopis JA, Soria F, Morcillo-Martín E, Caballero-Pérez P, de la Cruz-Conty JE, et al. Outcomes of ureteroscopy miniaturization on tissue damage and tissue hypoxia in a pig model. *Sci Rep*. 2018;8:1–9.
18. Gwet K. Benchmarking Inter-Rater Reliability Coefficients. *Handb Inter-rater Reliab*. 2012:121–8.
19. Tepeler A, Resorlu B, Sahin T, Sarikaya S, Bayindir M, Oguz U, et al. Categorization of intraoperative ureteroscopy complications using modified Satava classification system. *World J Urol*. 2014;32:131–6.
20. Karakan T, Kilinc M, Demirbas A, Hascicek A, Doluoglu O, Yucel MRB. Evaluating of Ureteral Wall Injuries with Endoscopic Grading System and Analysis of the Predisposing Factors. *J Endourol*. 2016;30:375–8.
21. Joshi HB, Stainthorpe A, MacDonagh RP, Keeley FX, Timoney AG. Indwelling ureteral stents: Evaluation of symptoms, quality of life and utility. *J Urol*. 2003;169:1065–9.
22. Lildal SK, Sørensen FB, Andreassen KH, Christiansen FE, Jung H, Pedersen MR, et al. Histopathological correlations to ureteral lesions visualized during ureteroscopy. *World J Urol*. 2017;35:1489–96.