

# Valor diagnóstico de las clasificaciones outerbridge e ICRS para la clasificación artroscópica de lesiones condrales de rodilla

## *Diagnostic Value of Outerbridge and ICRS Classifications for Arthroscopic Grading of Chondral Lesions in the Knee*

María Tuca<sup>1</sup>  Tomás Pineda<sup>2</sup>  Javier Oyarce<sup>3</sup>  Felix Etchegaray<sup>4</sup> David Figueroa<sup>1</sup>

<sup>1</sup>Departamento de Ortopedia y Traumatología, Clínica Alemana de Santiago, Universidad del Desarrollo, Santiago de Chile

<sup>2</sup>Departamento de Traumatología y Ortopedia, Universidad de Chile, Santiago, Chile

<sup>3</sup>Servicio de Ortopedia y Traumatología, Hospital Padre Hurtado, Santiago, Chile

<sup>4</sup>Departamento Ortopedia y Traumatología, Clínica Santa María, Santiago, Chile

Address for correspondence: María Tuca, MD, Departamento de Ortopedia y Traumatología, Clínica Alemana de Santiago, Universidad del Desarrollo, Santiago de Chile (e-mail: mjtuca@gmail.com).

Rev Chil Ortop Traumatol 2020;61:53–59.

### Resumen

**Objetivos** Determinar el rendimiento de las clasificaciones de *Outerbridge* (OB) e *International Cartilage Repair Society* (ICRS).

**Método** Estudio de test diagnóstico, diseño y recolección de datos prospectivo. Siete traumatólogos sub-especialistas observaron un mismo video donde se exponían 30 lesiones condrales bajo visión artroscópica, para luego clasificarlas según OB e ICRS y consignar el tratamiento de elección para cada una de las lesiones, eligiendo entre 6 alternativas: observación, debridamiento mecánico o térmico, microfractura, OATS o terapias biológicas. Tras 7 días, los evaluadores debían repetir el mismo procedimiento.

**Resultados** La concordancia entre los observadores fue débil para clasificación de ICRS ( $k$  0,25  $p < 0,001$ ) y moderada para la clasificación de OB ( $k$  0,45  $p < 0,001$ ). La concordancia intra-observador para ICRS oscilaba entre moderada y excelente ( $k$  promedio de 0,67), y para la clasificación de OB entre buena y excelente ( $k$  promedio 0,83). Ninguna de las dos clasificaciones mostró correlación con la experiencia del cirujano. En la elección de tratamiento, la concordancia entre observadores fue débil ( $k$  0,33  $p < 0,001$ ), sin embargo, la concordancia intra-observador fue en todos los casos buena o excelente ( $k$  0,82), mostrando una correlación directamente proporcional a la experiencia del cirujano. La capacidad de discriminación terapéutica, evaluada mediante una regresión logística, mostró un área bajo la curva roc en el rango del no-efecto.

**Conclusión** Ambas clasificaciones mostraron una baja correlación inter-observador y una elevada concordancia intra-observador. En ambas categorías, *Outerbridge* fue más concordante que ICRS. En cuanto al tratamiento, ninguna de las dos clasificaciones logra unificar criterios quirúrgicos.

**Nivel de evidencia** Nivel I (test diagnóstico).

### Palabras clave

- ▶ cartílago
- ▶ rodilla
- ▶ clasificaciones
- ▶ correlación interobservador
- ▶ concordancia intraobservador

received  
January 28, 2020  
accepted  
May 5, 2020

DOI <https://doi.org/10.1055/s-0040-1713381>.  
ISSN 0716-4548.

Copyright © 2020 by Thieme Revinter Publicações Ltda, Rio de Janeiro, Brazil

License terms



**Abstract**

**Objectives** Assess de diagnostic accuracy of Outerbridge (OB) and ICRS (International Cartilage Repair Society) classifications.

**Methods** We performed a diagnostic test study, with a prospective design and data collection. Seven knee surgeons were asked to observe a video were the 30 chondral lesions were shown through arthroscopic view. Simultaneously they were asked to classify them according to OB and ICRS. Besides, they had to define how they would manage the chondral lesion, choosing among six treatment options (observation, mechanical or thermic chondroplasty, microfracture, osteochondral autologous transfer system (OATS) or biological therapies). A week later, they repeated the same procedure. Intra and interobserver agreement were characterized by  $\kappa$  statistical analysis, and a logistic regression was used to assess the ability of both classifications to discriminate among treatment options. P values  $< 0,05$  were considered significant.

**Results** Interobserver agreement was weak ( $\kappa 0.25$   $p < 0.001$ ) for ICRS classification and moderate for OB classification ( $\kappa 0.45$   $p < 0.001$ ). Intraobserver agreement for ICRS ranged from moderate to excellent (average  $\kappa$  of 0.67), and for the OB classification ranged from good to excellent (average kappa 0.83). Neither classification correlated with the surgeon's experience. Interobserver agreement for therapeutic choice was poor ( $\kappa 0.33$   $p < 0.001$ ). However, intraobserver agreement was good to excellent ( $\kappa 0.82$ ) in all cases, showing a direct correlation with the surgeon's experience. Logistic regression used to assess the ability of both classifications to discriminate among treatment options, showed in both cases an area under the roc curve in the no-effect range.

**Conclusion** Both classifications showed low interobserver and high intraobserver agreements for arthroscopic grading of chondral lesions. In both, Outerbridge was more reliable than ICRS. As for guiding therapeutic management, none of the classifications could unify surgical criteria.

**Keywords**

- ▶ cartilage
- ▶ knee
- ▶ classifications
- ▶ interobserver concordance
- ▶ intraobserver concordance

**Introducción**

Las lesiones condrales son un hallazgo frecuente en cirugía artroscópica de rodilla, tanto en cohortes nacionales como internacionales, donde se ha descrito una prevalencia alrededor del 50%.<sup>1,2</sup> Sin embargo, el foco de investigación se ha centrado en el tratamiento de las lesiones condrales, existiendo escasas publicaciones que reporten la validez y el rendimiento de las herramientas diagnósticas para esas lesiones. Actualmente, el *Gold Standard* para su diagnóstico es la visión y exploración directa bajo artroscopia.

Una clasificación ideal para las lesiones condrales, es aquella que permita determinar pronóstico, ayude a orientar el tratamiento, y logre estandarizar un lenguaje común, unificando criterios entre los observadores. Existen múltiples clasificaciones diagnósticas descritas y publicadas sin embargo las dos más ampliamente utilizadas, son la clasificación de *Outerbridge* (OB) y la de *International Cartilage Repair Society* (ICRS).

En un estudio realizado en Alemania, Spahn G. y col., muestran una encuesta realizada a 105 cirujanos con gran experiencia en artroscopia, confirmando que la gran mayoría utiliza la clasificación de OB, y en segundo lugar la

clasificación de ICRS, conformando entre ambas el 95% de las preferencias. Además, destaca que el 50% de los cirujanos encuestados consideraba que los sistemas de clasificación disponibles requerían una mejoría, y que desconfiaban de su validez.<sup>3</sup>

El objetivo de este estudio, es evaluar el rendimiento diagnóstico de las clasificaciones de OB e ICRS para lesiones condrales de rodilla, y su utilidad para orientar el tratamiento. La hipótesis de nuestro estudio es que la clasificación de *Outerbridge* es más concordante que la clasificación de ICRS, siendo ambas útiles para determinar la elección terapéutica.

**Métodos**

Se realizó un estudio de test diagnóstico, con diseño y recolección de datos prospectivo. Se grabó un video DVD-HD que mostraba 30 lesiones condrales focales sucesivas, recopiladas en artroscopias de rodilla realizadas entre mayo y junio del 2013 por el mismo equipo quirúrgico en nuestro centro. Este video, mostraba cada una de las lesiones condrales desde su identificación inicial, posterior exploración, debridamiento y palpación.

Se reclutó a 7 cirujanos sub-especialistas en rodilla, que se graduaron del 1 al 7 acorde a su nivel de experiencia, el cual fue definido según años de sub-especialista y número de artroscopias año (n° artroscopias/año). Se les solicitó que tras revisar las imágenes clasificaran cada una de las lesiones según OB e ICRS tras la exploración, debridamiento y palpación de la lesión observada en el video. Además, debían indicar para cada una de las lesiones, una conducta terapéutica entre 6 alternativas: Observación, Debridamiento mecánico, Debridamiento térmico, OATS, Microfracturas o Terapias Biológicas, todo ello considerando el contexto de un paciente joven, activo, sintomático y sin limitación de recursos. Ese proceso se realizó en un primer tiempo, y fue repetido exactamente igual a los 7 días. Cada cirujano contaba con las clasificaciones de OB e ICRS, en el momento de ver el video. [► **Tabla 1**]

Para el análisis estadístico se utilizó el programa STATA v12.0. Se analizó la concordancia inter-observador e intra-observador utilizando el índice de kappa y se crearon variables que explicitaban la concordancia global intra-observador para las escalas OB e ICRS con la concordancia del tratamiento.<sup>4</sup> Para determinar la capacidad de discriminación terapéutica de ambas clasificaciones, se utilizó un modelo de regresión logística, mediante el cálculo del área bajo la curva ROC.<sup>5</sup>

Se crearon variables que explicitaban la concordancia global intra-observador para las escalas ICRS y OB con la concordancia del tratamiento. Mediante un análisis de regresión logística para la respuesta concordancia del tratamiento con cada una de las concordancias de las escalas, se estimó la capacidad de discriminación mediante el área bajo la curva ROC.

Para ver la asociación entre diagnóstico y tratamiento, se realizó una regresión logística a partir de la concordancia global, con lo cual se estimó el tratamiento esperado. Utilizando el análisis de curva ROC, se contrasta el tratamiento esperado con el tratamiento observado.

**Tabla 1** Clasificaciones

	OUTERBRIDGE	ICRS
<b>Grado 0</b>		Cartílago normal
<b>Grado 1</b>	Reblandecimiento y edema de cartílago	1. Cercano a lo normal
		1.a Fibrilación y/o leve reblandecimiento
		1.b Laceración o fisuras superficiales
<b>Grado 2</b>	Fragmentación y fisuración menor a 0,5 pulgadas de diámetro	2. Cartílago anormal, lesión que compromete menos del 50% del cartílago articular
<b>Grado 3</b>	Fragmentación y fisuración mayor a 0,5 pulgadas de diámetro	3. Severamente anormal, defecto en cartílago mayor a 50% del espesor
		3.a Hasta capa calcificada
		3.b Profundo a capa calcificada
		3.c Hasta hueso subcondral
		3.d Ampollas
<b>Grado 4</b>	Erosión del cartílago que expone hueso subcondral	4. Severamente anormal
		4.a Lesiones que abarcan el hueso subcondral
		4.b Lesiones que se extienden más profundo que el hueso subcondral

**Tabla 2** Concordancia inter-observador

	Outerbridge	ICRS
Día 0	0,43	0,24
Día 7	0,47	0,26
Promedio	0,45	0,25
Valoración índice Kappa	Moderado	Débil

## Resultados

### Concordancia Inter-observador

Los resultados del estudio mostraron que para la clasificación de OB, el promedio del índice kappa fue de 0,45 ( $p < 0,001$ ), evidenciando una concordancia moderada para esa clasificación. [► **Tabla 2**]

Con respecto a la clasificación de ICRS, el promedio del índice kappa fue de 0,25 ( $p < 0,001$ ), mostrando una débil concordancia para esa clasificación. [► **Tabla 2**]

### Concordancia Intra-observador

Para la clasificación de OB, la concordancia intra-observador fue en todos los casos, buena a excelente, con un promedio de índice kappa de 0,83 ( $p < 0,001$ ), representando una excelente concordancia intra-observador para esa clasificación. Podemos constatar que no existe correlación lineal entre la pericia del cirujano y la concordancia intra-observador [► **Tabla 3**].

Para la clasificación de ICRS, la concordancia intra-observador varió de moderada a excelente, con un promedio de índice kappa de 0,67 ( $p < 0,001$ ), resultando en una buena concordancia intra-observador para esa clasificación. Al igual que para la clasificación de OB, se puede observar que no existe correlación lineal entre la pericia de los cirujanos y la concordancia intra-observador [► **Tabla 3**].

**Tabla 3** Concordancia intra-observador

Cirujano	Outerbridge	ICRS
1	0,66	0,54
2	0,85	0,82
3	0,89	0,67
4	0,90	0,83
5	0,75	0,59
6	0,90	0,75
7	0,90	0,54
Promedio	0,83	0,67
Valoración índice Kappa	Excelente	Buena

**Concordancia de Tratamiento Intra-observador**

Analizando la concordancia de tratamiento intra-observador, pudimos apreciar que en todos los casos fue de buena (rango 0,62- 0,8) a excelente (rango de 0,81- 1), con un índice kappa promedio excelente (0,82), revelando que los cirujanos eran consistentes con su opción terapéutica a los 0 y 7 días. Si observamos los resultados de la [►Figura 1], podemos ver que existe una correlación lineal entre la pericia del cirujano y la concordancia del tratamiento.

**Discriminación Terapéutica**

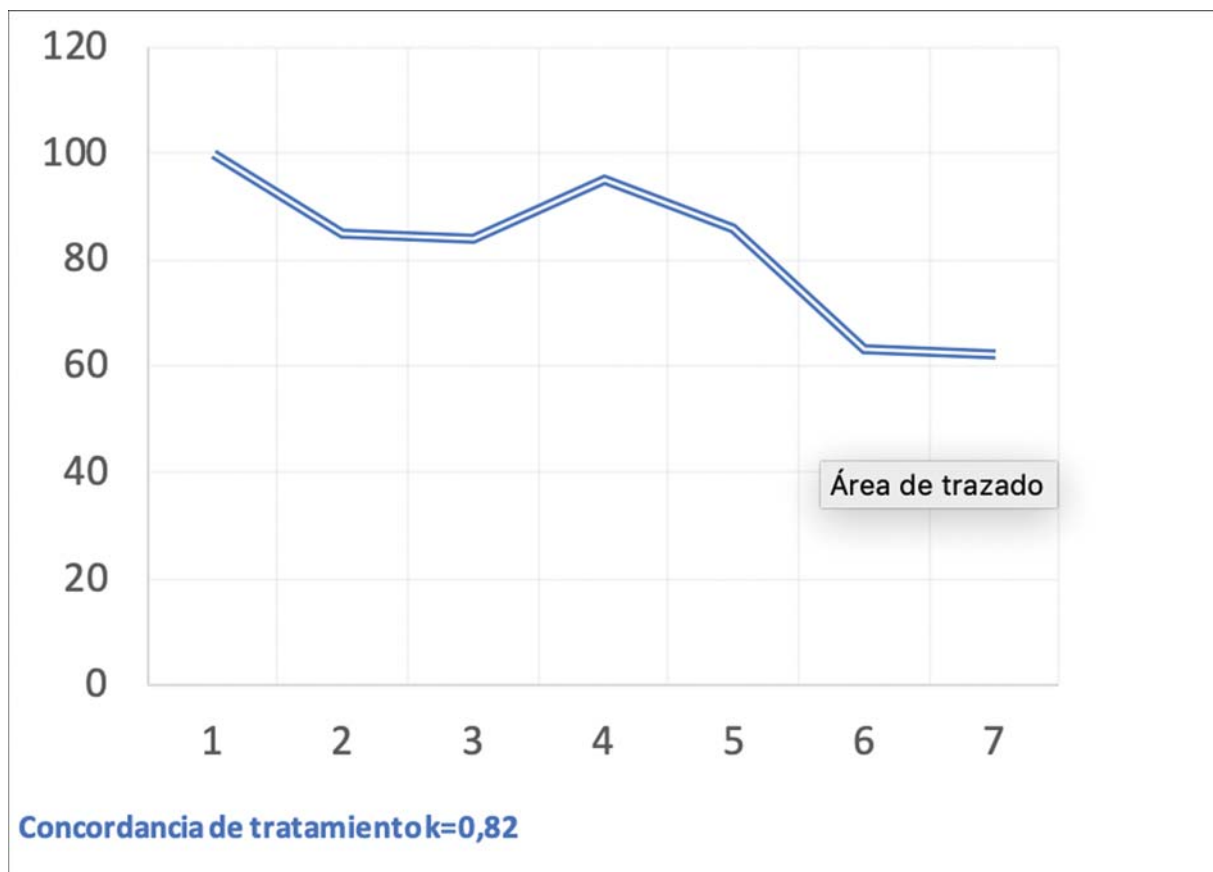
Finalmente, el modelo de regresión logística utilizado para evaluar la habilidad de ambas clasificaciones para discriminar

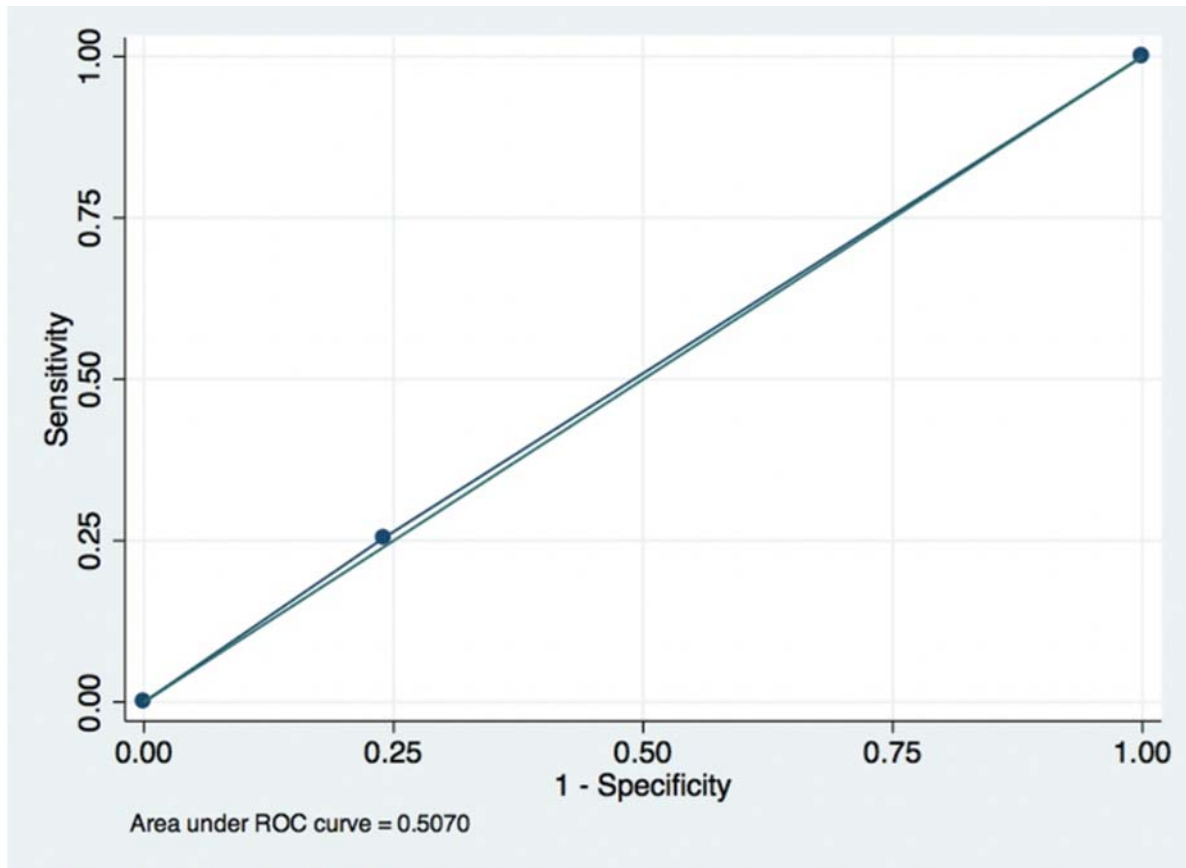
entre las distintas opciones de tratamiento, mostró en ambos casos el área bajo la curva ROC en el rango de no-efecto, evidenciando un ROC de 0,50 para OB y un 0,55 para ICRS. [►Figura 2] [►Figura 3]

**Discusión**

La adecuada valoración de las lesiones condrales es de gran importancia en la práctica clínica. Permite la correcta comunicación entre los distintos especialistas, determina el algoritmo terapéutico y guía el pronóstico.<sup>6</sup> Se han reportado múltiples clasificaciones para las lesiones del cartílago articular, con criterios radiológicos, histológicos y macroscópicos, cada una con sus ventajas y desventajas, siendo la artroscopia el método más validado para diagnosticar y clasificar.<sup>7</sup> La literatura actual sobre la validez de las clasificaciones para lesiones condrales es escasa, y en general solo evalúa cada clasificación por separado.

El método de *Outerbridge* fue de los primeros sistemas de graduación de lesiones condrales descrito en el año 1961, que fue inicialmente creado para la evaluación de condromalacia patelar. Al ser un método relativamente fácil de aplicar, fue ampliado para el resto de la rodilla.<sup>8</sup> Existen estudios previos que han evaluado la concordancia inter-observador utilizando esa clasificación, mostrando resultados que oscilan entre débiles a moderados, coincidiendo con los resultados encontrados en nuestra serie. Destacando que este reporte supera en potencia estadística a las publicaciones previas,

**Fig. 1** Concordancia de tratamiento intra-observador (índice kappa promedio: 0.82).



### Opciones de tratamiento (OB)

**Fig. 2** Curva ROC para Clasificación de *outerbridge*.

tanto en número de observadores como de casos analizados. Respecto a la concordancia intra-observador, solo el trabajo de Cameron y colaboradores, incorpora dos mediciones en el tiempo, permitiéndole así calcular la concordancia intra-observador, que al igual que nuestra serie, muestra valores kappa excelentes.<sup>9-11</sup>

El otro método evaluado en nuestro estudio fue el descrito en 1998 por la *International Cartilage Repair Society* (ICRS), el cual es más detallado que el de *Outerbridge*,<sup>12</sup> y es el más popular en la literatura actual.<sup>3,13</sup> La desventaja es que al contar con más subclasificaciones, su descripción y memorización tiende a ser más compleja.

En cuanto a esa clasificación, sólo se ha publicado un trabajo que evalúa su rendimiento, mostrando una concordancia inter-observador débil, que concuerda con los resultados de nuestro trabajo. En él sólo se realiza una evaluación en el tiempo, por tanto, no describe la concordancia intra-observador de ICRS.<sup>13</sup>

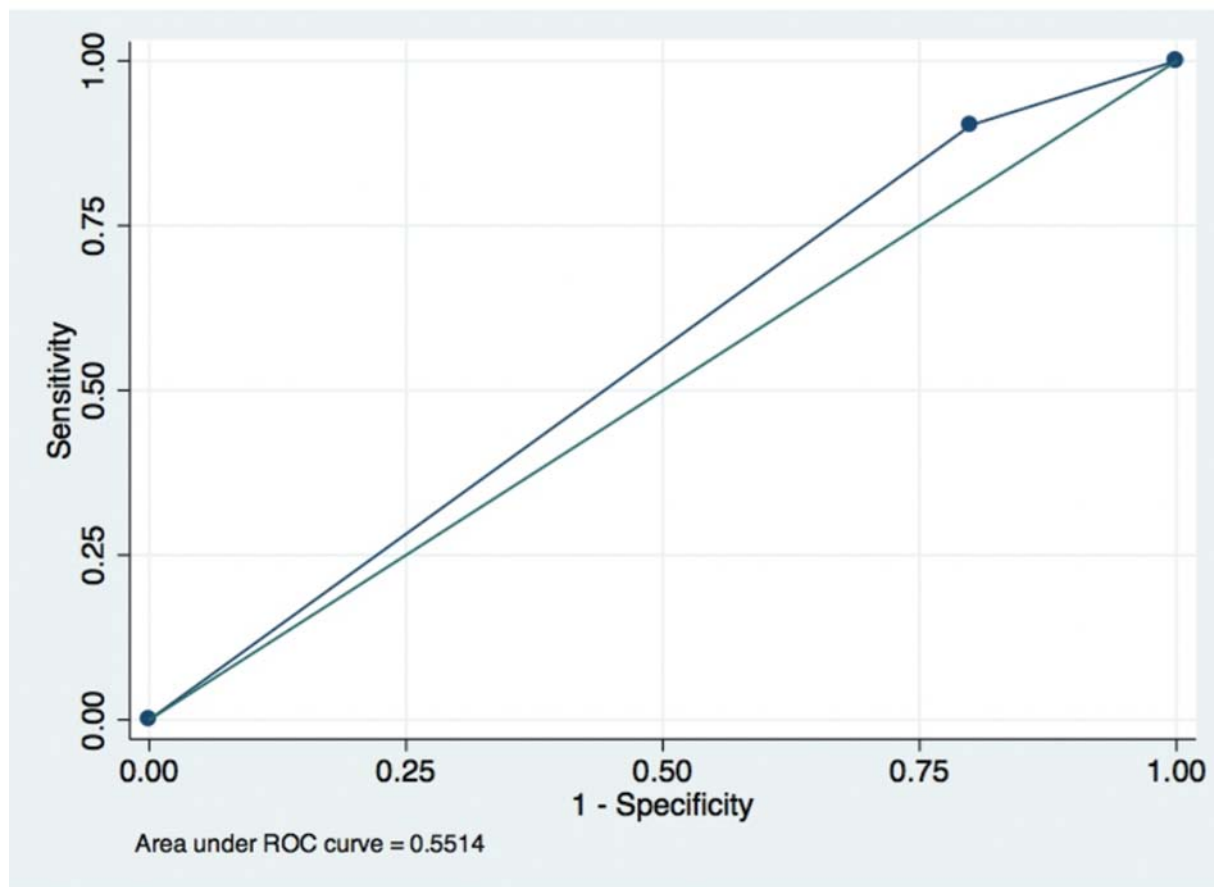
Ambas clasificaciones tuvieron concordancia inter-observador de rango débil a moderado, lo que explica la sensación de desconfianza respecto a esas clasificaciones, confirmando que no logra unificar criterios entre los distintos observadores. Eso sugiere que en la práctica, pudiese ser preferible una detallada descripción de las

lesiones observadas, más que registrarlas mediante alguna de esas clasificaciones.

Por otra parte, la concordancia intra-observador en nuestro estudio, fue en rango excelente para ambas clasificaciones, mostrando una correlación lineal con la pericia del cirujano, lo que significa que ellas pueden ser una herramienta útil y confiable para el registro individual.

Además, debemos destacar que nuestro estudio sugiere que ninguna de las dos clasificaciones fue consistente en guiar la decisión terapéutica, lo que significa que ante una misma categorización los cirujanos elegían tratamientos dispares, dejando en evidencia que ellas no eran necesariamente utilizadas como una guía a la hora del tratamiento.

En cuanto a las debilidades de nuestro estudio, podemos mencionar que se clasificó a las lesiones mediante una grabación en video, lo que no permite al observador realizar pruebas de estabilidad y palpación in vivo. Sin embargo, varios de los estudios publicados que se han realizado bajo visualización directa, tienen resultados comparables a los publicados por nuestro grupo. Creemos que podría ser un aporte realizar un análisis que incluyera la localización de las lesiones, ya que el sitio de lesión pudiera influir en la elección de la clasificación y las distintas opciones terapéuticas.



**Opciones de tratamiento (ICRS) 1**

**Fig. 3** Curva ROC para Clasificación de ICRS.

## Conclusión

Los hallazgos de este estudio nos permiten concluir que ambas clasificaciones mostraron una baja concordancia inter-observador, lo que significa que al evaluar una lesión condral mediante artroscopia, la clasificación varía entre cirujanos de manera importante, produciendo diferencias en la categorización y por lo tanto, en el manejo de dicha lesión. Además, ninguna de las dos clasificaciones mostró concordancia con el tratamiento elegido por el cirujano, sugiriendo que no logran unificar criterios terapéuticos.

Al evaluar cada cirujano individualmente, mediante concordancia intra-observador, el estudio demostró que existe una buena a excelente concordancia en ambas clasificaciones, lo que significa que ella puede ser útil y confiable para el registro individual de cada cirujano.

### Conflictos de Intereses

Ninguno

## Bibliografía

- 1 Figueroa D, Calvo R, Vaisman A, Carrasco MA, Moraga C, Delgado I. Knee chondral lesions: incidence and correlation between

arthroscopic and magnetic resonance findings. *Arthroscopy* 2007;23(03):312-315

- 2 Curl WW, Krome J, Gordon ES, Rushing J, Smith BP, Poehling GGC. Cartilage injuries: a review of 31,516 knee arthroscopies. *Arthroscopy* 1997;13(04):456-460
- 3 Spahn G, Klinger HM, Hofmann GO. How valid is the arthroscopic diagnosis of cartilage lesions? Results of an opinion survey among highly experienced arthroscopic surgeons. *Arch Orthop Trauma Surg* 2009;129(08):1117-1121
- 4 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33(01):159-174
- 5 Franco N, Vivo Molina JM. 'Análisis de curvas ROC: Principios básicos y aplicaciones'. Cuadernos de Estadística. EdLa Muralla, S. A. ISBN: 978- 84-7133-772-6.
- 6 Figueroa D, Calvo R, Vaisman A. Actualización en lesiones del cartílago articular. 2 ed. Santiago de Chile 2015
- 7 Beguin JA, Héron JF, Sabatier JP, Locker B, Souquières G. [Arthroscopy of the knee. Diagnostic value. 1005 cases]. *Nouv Presse Med* 1982;11(49):3619-3621
- 8 Outerbridge RE. The etiology of chondromalacia patellae. *J Bone Joint Surg Br* 1961;43-B:752-757
- 9 Cameron ML, Briggs KK, Steadman JR. Reproducibility and reliability of the outerbridge classification for grading chondral lesions of the knee arthroscopically. *Am J Sports Med* 2003;31(01):83-86
- 10 Marx RG, Connor J, Lyman S, et al; Multicenter Orthopaedic Outcomes Network. Multirater agreement of arthroscopic grading of knee articular cartilage. *Am J Sports Med* 2005;33(11):1654-1657

- 11 Javed A, Siddique M, Vaghela M, Hui AC. Interobserver variations in intra-articular evaluation during arthroscopy of the knee. *J Bone Joint Surg Br* 2002;84(01):48-49
- 12 International Cartilage Repair Society Cartilage Injury Evaluation Package. Develop during ICRS 2000 Standarts Workshop. Schloss Mucnhenwiler, Switzerland January 27-30-2000
- 13 Spahn G, Klinger HM, Baums M, et al. Near-infrared spectroscopy for arthroscopic evaluation of cartilage lesions: results of a blinded, prospective, interobserver study. *Am J Sports Med* 2010;38(12):2516-2521