

Nivel de evidencia de la RICMA comparada con el JHSE y JHSA

R. SÁNCHEZ ROSALES¹, V. J. HERNÁNDEZ GONZÁLEZ², M. E. POLLWEIN², L. REBOSO-MORALES³, Y. MARTÍN HIDALGO³.

¹Unidad de Cirugía de la Mano y Microcirugía, GECOT, San Cristóbal de La Laguna, Tenerife, España. ²Clínica de Fisioterapia NOVA, SIC de Tenerife, España. ³Servicio de Cirugía Ortopédica y Traumatología, Hospital Universitario de Nuestra Señora de Candelaria HUNSC. Universidad de La Laguna. San Cristóbal de La Laguna, Tenerife, España.

Resumen: Objetivo: El propósito del presente trabajo fue evaluar el nivel de evidencia científica de artículos de investigación clínica publicados en las Revista Iberoamericana de Cirugía de la Mano (RICMA) y compararlo con las revistas Europea (JHSE) y Americana (JHSA). Material y método: Un total de 932 trabajos de investigación clínicos publicados entre 2005 y 2009 (RICMA 60, JHSE, JHSA 461 y 411) fueron revisados. Dos observadores independientes clasificaron el nivel de evidencia basado en la Clasificación Internacional de Oxford, siendo 5 el nivel más bajo y I el nivel más alto de evidencia científica. Las frecuencias observadas de nivel de evidencia para cada revista científica fueron comparadas con las frecuencias esperadas mediante un test de chi-cuadrado (X^2) para variables categóricas con un nivel de significancia del 0.05.

Resultados: Más del 80% de los trabajos publicados en la RICMA y el JHSE, y el 67.6% en el JHSA presentaron un nivel IV de evidencia. Trabajos de nivel I o II no fueron publicados en la RICMA, comparado con el JHSE (0.9% nivel I, 5% nivel II), y con el JHSA (8.3% nivel I, 10% nivel II). El porcentaje de artículos con nivel III publicados en la RICMA (16.7%) fue mayor al observado en el JHSE (11.1%) y en el JHSA (14.1%). Todos los resultados fueron estadísticamente significativos ($X^2=63.95$; $p<0.001$).

Conclusiones: El nivel de evidencia de las publicaciones en investigación clínica en cirugía de la mano depende del tipo de revista, siendo el mayor nivel de evidencia en aquellos trabajos publicados en el JHSA, seguido del JHSE y finalmente la RICMA.

Palabras claves: Nivel de evidencia científica, Cirugía de mano, Investigación clínica.

Level of evidence of the clinical research published in the Ibero-american Journal of Hand Surgery (RICMA) compared to the JHSE and JHSA

Abstract: Objective: The purpose of this study was to assess the level of evidence of the clinical research papers published in the Ibero-american (RICMA), the European (JHSE) and American (JHSA) Journals of Hand Surgery.

Methods: A total of 932 clinical research papers published between 2005 and 2009 (RICMA 60, JHSE 461, and JHSA 411) were reviewed. Two independent observers classified the level of evidence based on the Oxford International Classification, 5 being the lowest level and I the highest level. The observed frequencies of the level of evidence for each journal were compared with the expected frequencies by a chi-square (X^2) test for categorical variables with a significance level of 0.05.

Results: More than 80% of the papers in RICMA and JHSE and a 67.6% in the JHSA presented a level of 4. No level I or 2 studies were published in RICMA, compared to JHSE (0.9% level I and 5.0% level 2) and JHSA (8.3% level I and 10% level 2). The percentage of papers with level 3 published in RICMA (16.7%) was higher compared to the JHSE (11.1%) and the JHSA (14.1%). All the results were statistically significant ($X^2=63.945$; $p<0.001$).

Conclusion: The level of evidence in clinical research related to hand surgery depend on the type of journal, being the highest level of evidence in papers published in the JHSA, followed by the JHSE and finally, the RICMA.

Key words: Level of evidence, Hand surgery, Clinical research.

Desde que fue descrito el primer sistema de clasificación del nivel de evidencia científica de los artículos de investigación clínica¹, la Medicina Basada en la Evidencia ha llegado a ser una importante parte de la práctica clínica. Los cirujanos de mano deberían entender el nivel de evidencia para llegar a ser conocedores de la veracidad y la utilidad de los datos contenidos en los artículos consultados. Pocas investigaciones se han hecho para analizar el nivel de evidencia científica en revistas relacionadas con cirugía de mano en relación a otros campos de investigación clínica afines, como

en revistas de cirugía ortopédica^{2,3} y cirugía plástica⁴. Sólo una revista específica de cirugía de mano ha sido analizada para ver su nivel de evidencia en un



Roberto Sánchez Rosales
M^o del Cristo Ossuna, 20, Local A
38204 San Cristóbal de La Laguna, SIC de Tenerife, España
titorosales@telefonica.net

período de seis meses, y ha sido comparada con otras de publicaciones ortopédicas³. Hasta nuestro conocimiento, no existen trabajos previos publicados en lengua castellana que comparen el nivel de evidencia científica de la RICMA con otras revistas científicas del campo de la cirugía de la mano. El propósito de este trabajo fue evaluar el nivel de evidencia científica de los artículos de investigación clínica publicados entre 2005 y 2009 en la Revista Ibero-Americana de Cirugía de Mano (RICMA), considerada como la revista y órgano de expresión de la Sociedad Española de Cirugía de Mano y de las que son Sociedades Científicas Afiliadas, la Sociedad Portuguesa de Cirugía de Mano y la mayor parte de las sociedades de Cirugía de Mano de Latino América y compararlo con la Revista Europea de Cirugía de Mano (JHSE) y la Americana (JHSA), como las revistas oficiales de la Federación de las Sociedades Europeas de Cirugía de Mano (FESSH) y la Sociedad Americana de Cirugía de Mano (ASSH).

Los investigadores establecieron como hipótesis nula (Ho) que la variable “nivel de evidencia científica” era independiente de la variable “tipo de revista”

MATERIAL Y MÉTODO

Población de estudio

Criterios de inclusión: Todos los artículos de investigación clínica que fueron publicados entre Enero de 2005 y Diciembre de 2009 en la Revista Ibero-Americana de Cirugía de Mano (RICMA), la Revista Europea de Cirugía de la Mano (JHSE) y la Revista Americana de Cirugía de la Mano (JHSA).

Criterios de exclusión: Estudios animales, anatómicos y de cadáver; estudios básicos de ciencias, cursos de instrucción, suplementos de resúmenes, publicaciones cortas, cartas al editor y artículos de revisión no fueron considerados factibles para el estudio.

Por esta razón, un total de 932 artículos de investigación clínica siguieron los criterios de inclusión y exclusión (RICMA 60, JHSE 461 y JHSA 411).

Evaluación del Nivel de Evidencia

Los artículos que siguieron los criterios de inclusión y exclusión fueron asignados de forma aleatoria a dos observadores independientes (YMH, LRM), con una larga experiencia clínica en cirugía de la mano y bien familiarizados con la Medicina Basada en la Evidencia. Un número igual aproximado de artículos de cada una de las revistas fue evaluado por cada observador:

El nivel de evidencia científica para cada artículo fue evaluado en base a “Niveles de Evidencia científica – Centro de Oxford para la Medicina Basada en la Evidencia”⁵. La valoración fue a ciegas en relación a cualquier información previa del nivel de evidencia de los artículos analizados (e.g.: “nivel de evidencia y tipo de estudio”, el cual es incluido en la JHSA con el resumen de cada artículo desde 2006). Los artículos fueron catalogados de acuerdo a su nivel de evidencia siendo el nivel 1 (mayor nivel de evidencia, e.g.: revisiones sistemáticas (SR), meta-análisis (MA), y estudios individuales controlados y aleatorizados con estrechos intervalos de confianza), y el nivel 5 (menor nivel de evidencia, e.g.: opinión de experto). La evaluación de los sub-grupos de nivel de evidencia, así como los tipos de artículos científicos no fueron realizados por los observadores.  [Tabla I.](#)

Análisis de la fiabilidad

Antes de empezar el estudio, la fiabilidad de la evaluación fue hecha en base al análisis del error intra-observador e inter-observador. Una muestra aleatoria de 30 artículos de investigación clínica, de un total de 872 artículos, publicados en Inglés (461 de JHSE, y 411 de JHSA), fueron evaluados por dos observadores independientes asignados para el estudio. Después de 15 días, se hizo una segunda evaluación con el orden de los artículos cambiado. Ningún artículo de la RICMA fue incluido en la muestra de estudio para el análisis de la fiabilidad. Esto se hizo para evitar un sesgo de información⁶, porque las diferentes lenguas presentes en la RICMA (Español y Portugués), podría aumentar la fiabilidad intra-observador. La fiabilidad intra e inter-observador fue estudiada usando el test Coeficiente Kappa con un nivel de significancia de 0.05.

Análisis estadístico

Para la evaluación de los resultados, el número de artículos clasificados para cada nivel de evidencia fue expresado como un porcentaje del total de artículos que cumplieron con los criterios de inclusión y exclusión durante el período de tiempo estudiado. Las frecuencias observadas para cada nivel de evidencia de cada revista fueron comparadas con las frecuencias esperadas usando un test chi-cuadrado (χ^2) para variables categóricas con un nivel de significancia de 0.05.

RESULTADOS

El análisis inter-observador mostró un Kappa de 0.617, con un error asintótico estándar (SE) de 0.117. El análisis intra-observador presentó un Kappa de 0.66 (SE

TABLA I – NIVEL DE EVIDENCIA Y TIPO DE ESTUDIO.

	Tratamiento, prevención, etiología, daño	Pronóstico Investigar el efecto de la característica del paciente en el resultado de la enfermedad	Diagnóstico Investigar un test diagnóstico. ¿Es este test diagnóstico exacto?	Diagnóstico diferencial, estudios de prevalencia	Estudios económicos y Análisis de decisión
Nivel I	Revisión sistemática de ensayos aleatorizados (ECA) ECA de alta calidad (e.g.: seguimiento superior al 80%, con un intervalo de confianza estrecho)	Revisión sistemática de estudios de cohortes Estudio de cohortes individual con un seguimiento superior al 80%, con todos los pacientes listados al mismo tiempo	Revisión sistemática de estudios diagnósticos de nivel I Estudios diagnósticos de nivel I o estudios de validación que prueben la calidad de un test diagnóstico específico, previamente desarrollado, en series de pacientes consecutivos con referencia a "gold standard"	Revisión sistemática de estudios de cohortes clásico o prospectivos Estudios prospectivos o cohortes clásico con buen seguimiento (>80%)	Revisión sistemática de estudios económicos de nivel I Estudios de nivel I (Análisis basado en costes clínicamente sensibles o alternativos, valores obtenidos de muchos estudios , e incluyendo análisis de sensibilidad de diferentes maneras)
Nivel II	Revisión sistemática de estudios de cohortes ECA de menor calidad (e.g.: seguimiento inferior al 80%, con un intervalo de confianza amplio) Estudio de cohortes individual, incluye estudios de cohortes apareados (Estudios prospectivos comparativos) Estudios ecológicos	Revisión sistemática de cualquier estudio histórico de cohorte o de grupos controles no tratados (brazo control) en RCTs Estudio histórico de cohortes (retrospectivo) o brazo de control de un RCT	Revisión sistemática de estudios diagnósticos de nivel 2 Estudios de diagnóstico de nivel 2 o estudios exploratorios que, a través de una regresión logística, determinan factores significativos (e.g.: usando regresión de análisis)	Revisión sistemática de estudios de nivel 2 Estudios de nivel 2 (Estudios retrospectivos o de cohortes histórico , o con un seguimiento inferior al <80%) Estudios ecológicos	Revisión sistemática de estudios de nivel 2 Estudios de nivel 2 (Análisis basados en costes clínicamente sensibles o alternativas de estudios limitados , e incluyendo análisis de sensibilidad de diferentes maneras)
Nivel III	Revisión sistemática de estudios casos-control Estudio caso-control individual		Revisión sistemática de estudios de nivel 3 Estudios diagnósticos de nivel 3 o estudios en pacientes no consecutivos y sin referencias consistentes de "gold standards".	Revisión sistemática de estudios de nivel 3 Estudios de nivel 3 (cohortes no consecutivos o población muy limitada)	Revisión sistemática de estudios de nivel 3 Estudios de nivel 3 (análisis basado en pocas alternativas o costes, poca calidad de estimación de datos, pero incluyendo análisis de la sensibilidad)
Nivel IV	Serie de casos Estudios Cohortes y caso-control de poca calidad*	Serie de casos Estudios Cohortes y caso-control de poca calidad*	Estudio de caso - control Estándares de referencia pobres o no independientes	Serie de casos	No análisis de sensibilidad
Nivel V	Opinión de experto	Opinión de experto	Opinión de experto	Opinión de experto	Opinión de experto

Generalmente, una revisión sistemática (RS) es mejor que cualquier estudio individual. Un estudio experimental (e.g.: ensayo controlado aleatorizado de buena calidad) es normalmente mejor que cualquier estudio observacional. Para estudios observacionales: un estudio cohorte es mejor que cualquier estudio casos-control. Un estudio de casos-control es generalmente mejor que cualquier estudio de

serie de casos. *Como "estudios cohortes de baja calidad" nos referimos a aquellos cohortes que son deficientes a la hora de definir claramente los grupos comparados y/o deficientes a la hora de medir exposiciones y resultados (preferiblemente a ciegas) de la misma objetiva manera en ambos grupos (expuestos y no expuestos) y/o deficientes a la hora de identificar las variables de confusión y/o con un seguimiento muy pobre. Lo mismo para "estudios casos-control de baja calidad" a excepción de que en los diseños clínicos de casos-control los pacientes entran en el estudio en base a la variable "resultado". Así los "casos" (e.g.: fracaso de reimplante) son comparados con aquellos que no presentan la variable resultado (e.g.: reimplante exitoso) llamados "control" y por tanto no tenemos "expuestos y no expuestos" ni tampoco tenemos seguimiento longitudinal. Estudios ecológicos y Económicos/ análisis de desición son muy poco frecuentes en cirugía de la mano. Esta tabla ha sido adaptada de material publicado por el Centro de medicina Basada en la Evidencia, Oxford, Inglaterra. Marzo de 2009.

0.114) para el observador 1, y un Kappa de 0.751 (SE 0.103) para el observador 2. Todos los valores Kappa fueron significativos ($p < 0.001$). Más del 80% de los artículos publicados en RICMA y la JHSE; y un 67.6% en la JHSA presentaron un nivel 4 de evidencia. Ningún artículo con nivel I y II de evidencia fue publicado en la RICMA comparado con la JHSE (0.9% de nivel I y 5.0% de nivel II) y la JHSA (8.3% de nivel I y 10% de nivel II). El porcentaje de artículos con nivel III, publicados en la RICMA (16.7%) fue mayor comparado con la JHSE (11.1%) y la JHSA (14.1%)  **Tabla II**. Todos los resultados fueron estadísticamente significativos ($\chi^2 = 63.945$; $p < 0.001$) y la Hipótesis nula (H_0) fue rechazada.

DISCUSIÓN

Los resultados de este artículo han demostrado con un nivel de fiabilidad bueno-excelente que la variable "nivel de evidencia científica" es dependiente de la variable "tipo de revista".

El uso de Kappa es importante, ya que nos permite conocer el grado de acuerdo intra-observador e inter-observadores teniendo en cuenta el azar. Un coeficiente Kappa estadísticamente significativo indica que el acuerdo es diferente a cero (acuerdo nulo). Sin embargo, la interpretación de valores Kappa obtenidos es subjetivo, y han sido propuestas diferentes clasificaciones o guías de interpretación del coeficiente Kappa para el análisis de fiabilidad. En este artículo, el nivel de acuerdo de análisis del inter-observador y el intra-observador ha mostrado un valor Kappa comprendido desde 0.617 a 0.751, que puede ser considerado que tiene un nivel de fiabilidad de buena a excelente^{7,8}, en la evaluación del nivel de evidencia y el tipo de revista.

Resultados similares han sido publicados antes. Obrem-skey y cols.³, en el estudio del nivel de evidencia de revistas de ortopedia, han publicado valores Kappa de 0.62 para acuerdos inter-observadores entre revisores con poca experiencia, y valores Kappa de 0.75

TABLA II – TABULACIÓN CRUZADA DE “TIPO DE REVISTA” Y “NIVEL DE EVIDENCIA”

REVISTA	NIVEL DE EVIDENCIA					NUMERO COMPLETO DE TRABAJOS
	I	II	III	IV	V	
RICMA	0.0%	0.0%	16.7%	80.0%	3.3%	60
(IC 95%)	(N.A.)	(N.A.)	(7.3; 26.1)	(69.9; 90.1)	(N.A.)	
	0	0	10	48	2	
JHSE	0.9%	5.0%	11.1%	82.4%	0.7%	461
(IC 95%)	(N.A.)	(3.02; 6.9)			(N.A.)	
	4	23	51	380	3	
JHSA	8.3%	10%	14.1%	67.6%	0%	411
(IC 95%)	(5.6; 10.9)	(7.1; 12.9)	(10.7; 17.4)	(63.1; 72.1)	(N.A.)	
	34	41	58	278	0	

RICMA: Revista Iberoamericana de Cirugía de la Mano / **JHSE:** Journal of Hand Surgery European Volume / **JHSA:** Journal of Hand Surgery American Volume / **IC:** intervalo de confianza / **N.A.** no aplicable: (cuando el numero de proporción observado fue menor de 5/n)

para inter-observadores experimentados. Ningún análisis intra-observador fue publicado por estos autores.

No muchos artículos han estudiado el nivel de evidencia científica en revistas de cirugía de la mano o en campos similares de investigación, como revistas de cirugía ortopédica y plástica. Sinno y cols.⁴, revisaron 726 artículos de 6 revistas diferentes de cirugía plástica y el nivel de evidencia fue evaluado usando una clasificación basada en el Centro para el Nivel de Evidencia de Oxford (CEBM). En Hanzlik y cols.² fueron evaluados 551 artículos del Volumen Americano del JBS de los años 1975 (134 artículos), 1985 (123 artículos), 1995 (120 artículos) y 2005 (174 artículos). El nivel de evidencia fue evaluado siguiendo una clasificación incluida en la guía de autores (sistema de clasificación JBSA), la cual es muy similar a la desarrollada por la CEBM, con el fin de demostrar la tendencia del nivel de evidencia durante 30 años. Obremskey y cols.³ revisaron 382 artículos de investigación clínica de nueve revistas diferentes con el fin de evaluar el nivel de evidencia científica en revistas de cirugía ortopédica. En el presente trabajo fueron revisados 932 artículos de tres revistas específicas de cirugía de la mano, lo cual supone la población más amplia estudiada hasta ahora, para evaluar el nivel de evidencia científica.

Los resultados de este trabajo demuestran que la mayoría de los artículos clínicos publicados en cirugía de la mano tienen muy poco nivel de evidencia científica (80% nivel IV en la JHSE o RICMA y 67.6% en la JHSA). La mayoría de los artículos fueron serie de casos y menos frecuentemente estudios cohortes y casos-control de pobre calidad comparado con revistas científicas de cirugía ortopédica (48% estudios de nivel IV)², plástica (40% estudios de nivel IV)⁴ y de oftalmología (58% estudios de nivel IV)⁹. Sin embargo, otras revistas de cirugía, como las relativas a otorrinolaringología, presentan un porcentaje similar a la JHSE y RICMA (80% estudios de nivel IV)¹⁰. El porcentaje de artículos de nivel IV en la JHSA fue menor, comparado con el resto de revistas de cirugía de mano investigadas, y fue muy similar al del trabajo publicado por Obremskey y cols.³, donde se notifica la existencia de un 68.8% de artículos de nivel IV, en una revisión de 32 artículos publicados en la JHSA de Enero a Junio de 2003.

El porcentaje de artículos con un alto nivel de evidencia científica (nivel I y II) fue mayor en la JHSA (8.3% nivel I y 10% nivel II), comparado con la RICMA (0%) y la JHSE (0.9% nivel I y 5% nivel II). Mientras que

si se compara con otras revistas, hubo un 21% de nivel I y un 15% de nivel II de evidencia en revistas de cirugía ortopédica², 3% de nivel I y 16% de nivel II en cirugía plástica⁴, 18% de nivel I y 8% de nivel II en revistas de oftalmología⁹, y 7% de niveles I y II en otorrinolaringología¹⁰.

El porcentaje de artículos con nivel III (mayoritariamente estudios de casos-control y cohortes no consecutivos o con poblaciones muy limitadas) publicados en la RICMA (16.7%) fue mayor comparado con la JHSE (11%) y la JHSA (14%); y similar a otras revistas: 16% en revistas de cirugía ortopédica², 16% en otorrinolaringología¹⁰, y 16% en oftalmología⁹.

Algunos autores han criticado el escaso número de artículos de alto nivel de evidencia en cirugía¹¹. Aun así, las críticas puede que parezcan exageradas si se tiene en cuenta que un ensayo clínico aleatorizado quirúrgico es diferente de un ensayo que compara un medicamento con un placebo. Los procedimientos quirúrgicos son invasivos; es difícil aleatorizar pacientes, el enmascaramiento es un problema en un ensayo quirúrgico, y ambos son muy caros. Si no se tiene estudios aleatorios de alta calidad, no se puede tener una revisión sistemática que sintetice la evidencia anteriormente descrita.

Ningún análisis de tendencia es una limitación de este estudio, y dicha información debería ser el propósito de futuros estudios, en el sentido de entender cómo la evidencia científica publicada en revistas de cirugía de la mano ha cambiado, y cómo la relación entre cambios en el nivel de evidencia y cambios en el índice de impacto ha variado también con el tiempo.

Después de revisar una enorme cantidad de artículos publicados en revistas de diferentes partes del mundo, otras cuestiones afloran, tales como si las diferencias encontradas son un reflejo de diferentes prioridades regionales, o de qué manera las inversiones utilizadas para la investigación tienen un impacto en los resultados, e incluso si diferentes países son los principales contribuidores en estudios de alto nivel.

CONCLUSIONES

El nivel de evidencia científica en cirugía de la mano es dependiente del tipo de revista; siendo los artículos de mayor nivel de evidencia los publicados en la JHSA, seguido por la JHSE y finalmente la RICMA. Conocer el estado del nivel de evidencia publicado en cirugía

de la mano, es el punto de partida para enfocar los cambios de mejora de la calidad de nuestros estudios clínicos. Así, la existencia de un 16% de artículos de nivel 3 de evidencia en la RICMA, superior a las otras revistas, puede ser considerado ese "punto de partida" hacia la mejora de la calidad de los artículos que se publican en nuestra revista oficial de la Sociedad Española de Cirugía de la Mano.

CONFLICTOS DE INTERESES

Los autores declaran no tener conflictos de intereses.

BIBLIOGRAFÍA

1. Sackett DL. Rules of evidence and clinical recommendations on the use of antithrombotic agents. *Chest*. 1986;89(2 Suppl):25-35.
2. Hanzlik S, Mahabir RC, Baynosa RC, Khiabani KT. Levels of evidence in research published in the journal of bone and joint surgery (American volume) over the last thirty years. *J Bone Joint Surg*. 2009;91A:425-8.
3. Obrebsky WT, Pappas N, Attallah-Wasif E, Tornetta P, Bhandari M. Level of evidence in orthopedic journals. *J Bone Joint Surg*. 2005;87A:2632-8.
4. Sinno H, Neel OF, Lutfy J, Bartlett G, Gilardino M. Level of evidence in plastic surgery research. *Plast Reconstr Surg*. 2011;127:974-80.
5. Oxford Center for Evidence-based Medicine: Levels of Evidence. Disponible en: <http://www.cebm.net/oxford-centre-evidence-based-medicine-levels-evidence-march-2009/>.
6. Page RM, Cole GE, Timmreck TC. *Basic Epidemiological Methods and Biostatistics. A practical guidebook*. Boston: Jones and Bartlett publishers; 1995.
7. Landis JR, Koch GG. The measurements of observer agreement for categorical data. *Biometrics*. 1977;33:159-74.
8. Silman AJ. *Epidemiological Studies: a practical guide*. New York: Cambridge University Press; 1995.
9. Lai TY, Leung GM, Wong VW, Lam RF, Cheng AC, Lam DS. How evidence-based are publications in clinical ophthalmic journals? *Invest Ophthalmol Vis Sci*. 2006;47:1831-8.
10. Bentsianov BL, Boruk M, Rosendfield RM. Evidence-based medicine in otolaryngology journals. *Otolaryngol Head Neck Surg*.
11. Horton R. Surgical research or comic opera: questions, but few answers. *Lancet*. 1996;347:984-5.