

REVISIÓN

Calidad de los estudios de fiabilidad de la exploración ecográfica en el tendón de Aquiles: Revisión sistemática

Elena Portaceli Oñate^a, Federico D. Puglisi^a, Jacinto J. Martínez Payá^a,
M^a Elena del Baño Aledo^a y José Ríos Díaz^{a,b,*}

^aG.I ECOFISTEM. Facultad de Ciencias de la Salud. Universidad Católica de Murcia, España

^bCentro Universitario de Ciencias de la Salud San Rafael-Nebrija. Madrid, España

PALABRAS CLAVE

Fiabilidad;
Ultrasonografía;
Tendón de Aquiles;
QAREL

Resumen

Introducción: El tendón de Aquiles es uno de los tendones que más frecuentemente se lesiona. La ecografía, una técnica no invasiva, indolora y de bajo coste, tiene cada vez más importancia en la práctica clínica para valorar el estado del tendón; sin embargo, su fiabilidad depende de la experiencia y habilidad técnica del explorador. El objetivo del estudio fue revisar y evaluar la calidad metodológica de los trabajos de fiabilidad de la exploración ecográfica sobre el tendón de Aquiles.

Material y método: Se realizó una revisión sistemática, siguiendo las recomendaciones PRISMA, buscando en las bases de datos MEDLINE, SPORTDiscus, Academic Search, CINAHL, E-Journals, IBECs, LILACS OVID, DIALNET, COCHRANE y PEDro. Se generó una herramienta formada por 9 ítems para valorar la calidad metodológica de la descripción de las exploraciones ecográficas. Para evaluar la calidad metodológica de los estudios de fiabilidad seleccionados se utilizó la escala QAREL.

Resultados: Se localizaron 227 artículos y tras un proceso de selección exhaustivo se incluyeron 13 trabajos. La información proporcionada relativa al cegamiento de los exploradores y al análisis estadístico es insuficiente en un 70% y en un 69% de los estudios analizados respectivamente.

Conclusiones: La descripción del método ecográfico es insuficiente para asegurar la reproducibilidad de los estudios revisados. Además, los artículos presentan problemas metodológicos en aspectos relevantes tales como la representatividad de la muestra, el cegamiento de los exploradores y el adecuado análisis estadístico de los datos.

© 2016 MVclinic. Publicado por Elsevier España, S.L.U. Este es un artículo open access distribuido bajo los términos de la licencia CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

*Autor para correspondencia.

Correo electrónico: Jrios@nebrija.es (J. Ríos Díaz).

KEYWORDS

Reliability;
Ultrasonography;
Achilles tendon;
QAREL

Quality of studies on reliability of ultrasound examination in the Achilles tendon: systematic review

Abstract

Introduction: The Achilles tendon is one of the most frequently injured. Ultrasound is a non-invasive, painless and inexpensive technique that can assess the condition of the tendons, but the reliability depends on the experience and skills of the examiner. The objectives of this work are to review and assess the methodological quality of the ultrasound studies on the Achilles tendon.

Material and method: This is a systematic review in which we searched on MEDLINE, SPORTDiscus, Academic Search, CINAHL, E-Journals, IBECs, LILACS OVID, DIALNET, COCHRANE y PEDro. A tool consisting of 9 items to assess the methodological quality of the description of ultrasound examinations was created. The QAREL scale was used to assess the methodological quality of the reliability studies.

Results: 227 articles were found and after passing a selection process we have chosen 13 articles to realize this review. The information provided on blinding of explorers and statistical analysis is respectively insufficient in 70% and 69% of cases.

Conclusions: The sonographic method description is insufficient to ensure the reproducibility of the studies reviewed. Besides, the articles have methodological problems in relevant aspects such as the representativeness of the sample, blinding of explorers and appropriate statistical analysis of the data.

© 2016 MVclinic. Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introducción

El tendón de Aquiles es el tendón más grueso y fuerte del cuerpo y es capaz de soportar hasta 12,5 veces el peso del cuerpo durante la marcha. Debido a las elevadas cargas y tensiones que tiene que soportar es uno de los más comunes en sufrir lesiones¹. Hoy en día las alteraciones en el tendón y los motivos que las desencadenan siguen siendo motivo de estudio y de discusión.

Se sabe que las tendinopatías son el grupo de lesiones más frecuentes en personas que practican deporte, ya sea a nivel profesional o de forma recreativa². Como factores asociados se han identificado la edad, el sexo, una vascularización deficiente, aspectos genéticos y una flexibilidad inadecuada³. Clínicamente, la manifestación más clara es la aparición de dolor y en muchos casos impotencia funcional cuando ya se han producido cambios tisulares en la estructura del tendón. A nivel ecográfico, los cambios más frecuentes en presencia de tendinopatía son un engrosamiento del tendón, hipoecogenicidad focal o global, calcificaciones proximales a la inserción en el calcáneo y, en ocasiones, signos de neovascularización⁴.

Desde el punto de vista del diagnóstico y del control evolutivo de la lesión, la ecografía musculoesquelética presenta una excelente relación coste-beneficio, es inocua, no invasiva y permite evaluar al paciente de forma dinámica; no es casualidad que se haya convertido en la técnica de elección para los profesionales de la salud, entre ellos el fisioterapeuta^{5,6}.

Estas imágenes se consiguen mediante la representación gráfica de los ultrasonidos (US) que al reflejarse en los diferentes tejidos corporales son captados de nuevo por el transductor y procesados por el *software* del equipo ecográfico que proporciona la imagen ecográfica⁶. Sin embargo,

dadas las características de la exploración, la calidad de la imagen obtenida depende no sólo de la optimización adecuada de la imagen sino del conocimiento anatómico, la experiencia y la pericia del explorador. Es precisamente este factor el más criticado y discutido cuando se utiliza la ecografía para la realización de estudios cuantitativos.

Por ello hay que tener presente que, antes de utilizar cualquier protocolo ecográfico para la obtención de información en investigación, es necesario conocer y evaluar la fiabilidad de dicho método. Se dice que un instrumento de medida (o metodología de estudio) es fiable cuando permite obtener un resultado similar bajo las mismas condiciones de medición⁷⁻⁹. La fiabilidad del instrumento de medida es el fundamento para que sea válido para la valoración o el diagnóstico.

Aunque existen trabajos publicados en los que se ha incluido el correspondiente estudio de fiabilidad, no siempre se realiza o se informa de manera adecuada. Por tanto, se realizó una revisión sistematizada sobre los estudios de fiabilidad de la exploración ecográfica sobre el tendón de Aquiles en humanos, con el objetivo de evaluar el grado de descripción de los parámetros que afectan a la reproducibilidad de la exploración ecográfica y al grado de cumplimiento de los estándares para los estudios de fiabilidad.

Material y método

Diseño del estudio

Se llevó a cabo una revisión sistemática siguiendo las recomendaciones PRISMA¹⁰. El periodo en el que se realizaron las búsquedas fue desde diciembre de 2015 a mayo de 2016.

Fuentes documentales

Las búsquedas se realizaron en la base de datos MEDLINE a través de la plataforma PUBMED; en las bases de datos SPORTDiscus, Academic Search, CINAHL y E-Journals a través de la plataforma EBSCO; en las bases de datos IBECs y LILACS a través de la plataforma BVS; en las bases de datos OVID y DIALNET a través de la plataforma virtual de la universidad CEU San Pablo, y en las bases de datos COCHRANE y PEDro. Además, se llevó a cabo una búsqueda manual para localizar el máximo de artículos posible.

Estrategias de búsqueda

Las búsquedas fueron realizadas por dos revisores independientes (EPO y FDP), se consensuaron las discrepancias con otro miembro del equipo (JRD), que adaptó la sintaxis de las búsquedas según la base de datos correspondiente. No se utilizó ningún tipo del filtro en cuanto a fecha desde publicación, edad, idioma o tipología de estudio para poder obtener una selección más sensible.

MEDLINE

#1 («Ultrasonography» [Mesh] OR ultrasonography[TIAB] OR echography[TIAB] OR sonography[TIAB] OR ultrasound*[TIAB]) AND (reliability[TIAB] OR reproducibility[TIAB] OR «Reproducibility of Results» [Mesh] OR agreement[TIAB] OR Validity [TIAB] OR inter-observer[TIAB] OR interobserver[TIAB] OR intra-observer[TIAB] OR intraobserver[TIAB] OR inter-rater[TIAB] OR interrater[TIAB] OR intra-rater[TIAB] OR intrarater[TIAB] OR intra-examiner[TIAB] OR intraexaminer[TIAB] OR inter-examiner[TIAB] OR interexaminer[TIAB] OR interexplorer[TIAB] OR test-retest[TIAB] OR «Observer Variation»[Mesh]) AND («Achilles tendon»[TIAB] OR «Achilles Tendon»[Mesh]).

EBSCO

Búsqueda avanzada por palabras en el resumen en SPORTDiscus with Full Text, Academic Search Complete, CINAHL Complete, E-Journals:

#2 (ultrasonography OR echography OR sonography OR ultrasound*) AND (reliability OR reproducibility OR agreement OR Validity OR inter-observer OR interobserver OR intra-observer OR intraobserver OR inter-rater OR interrater OR intra-rater OR intrarater OR intra-examiner OR intraexaminer OR inter-examiner OR interexaminer OR interexplorer OR test-retest) AND (Achilles tendon).

OVID

Búsqueda multi-campo por palabras en el resumen:

#3 (ultrasonography OR echography OR sonography OR ultrasound) AND (reliability OR reproducibility OR agreement OR Validity OR inter-observer OR interobserver OR intra-observer OR intraobserver OR inter-rater OR interrater OR intra-rater OR intrarater OR intra-examiner OR intraexaminer OR inter-examiner OR interexaminer OR interexplorer OR test-retest) AND (Achilles tendon).

Biblioteca Virtual en Salud (BVS)

Búsqueda por palabras en IBECs Y LILACS:

#4 (ultrasonography OR echography OR sonography OR ultrasound) AND (reliability OR reproducibility OR agreement OR Validity OR inter-observer OR interobserver OR intra-

observer OR intraobserver OR inter-rater OR interrater OR intra-rater OR intrarater OR intra-examiner OR intraexaminer OR inter-examiner OR interexaminer OR interexplorer OR test-retest) AND (Achilles tendon).

COCHRANE

Búsqueda avanzada por palabras en el título/resumen/palabras clave:

#5 (ultrasonography OR echography OR sonography OR ultrasound) AND (reliability OR reproducibility OR agreement OR Validity OR inter-observer OR interobserver OR intra-observer OR intraobserver OR inter-rater OR interrater OR intra-rater OR intrarater OR intra-examiner OR intraexaminer OR inter-examiner OR interexaminer OR interexplorer OR test-retest) AND (Achilles tendon).

PEDro

Búsqueda simple:

#6 ultrasonography AND Achilles tendon.

#7 sonography AND Achilles tendon.

#8 ultrasound* AND Achilles tendon.

#9 echography AND Achilles tendon.

Dialnet

Búsqueda por palabras en documentos:

#10 ultrasonography AND Achilles tendon.

#11 sonography AND Achilles tendon.

#12 ultrasound* AND Achilles tendon.

#13 echography AND Achilles tendon.

Criterios de selección

Solamente se incluyeron las publicaciones de revistas revisadas por pares que estudiaran la fiabilidad de la ecografía en el tendón de Aquiles (TA) en humanos vivos. Se descartaron aquellos trabajos en los que se utilizaba la sonoelastografía, que si bien tiene una base ecográfica tiene otras particularidades especiales, y los trabajos que evaluaban fiabilidad en la clasificación y diagnóstico cualitativo.

Grado de cumplimiento de la descripción del método ecográfico

Para evaluar el grado de descripción del método ecográfico que afecta directamente a la reproducibilidad de los estudios se elaboró una herramienta *ad hoc* no validada basada en la experiencia clínica e investigadora de los cinco autores.

Se describieron 9 ítems que se codificaron como sigue: 0 puntos si el ítem no se describe, 1 punto si el ítem se describe de forma dudosa o incompleta y por lo tanto no reproducible, 2 puntos si el ítem se describe de forma clara y por lo tanto reproducible (tabla 1).

Calidad metodológica de los estudios de fiabilidad

Para la evaluación de la calidad metodológica de los estudios de fiabilidad se utilizó la escala Quality Appraisal for Reliability Studies (QAREL)¹¹ que se compone de 11 preguntas (cumple, no cumple, dudoso o no aplicable) agrupadas en tres grupos que describen la validez interna (ítems 3-9), la validez externa (ítems 1, 2 y 10) y la pertinencia de los análisis estadísticos (ítem 11).

Tabla 1 Descripción de la escala de reproducibilidad del método

Ítem	Descripción
1: <i>Posición del sujeto</i>	Se describe la posición del sujeto durante la exploración así como la posición de la región anatómica que se explora.
2: <i>Región anatómica explorada</i>	Se describe la localización exacta del transductor en la estructura de la zona a explorar.
3: <i>Plano de exploración (corte)</i>	Se describe el plano o tipo de corte y si fue o no bilateral.
4: <i>Características del transductor</i>	Se describe el tipo de sonda y la frecuencia.
5: <i>Material de interposición</i>	Se describe el tipo de material de interposición utilizado y la cantidad en el caso de que se utilizara gel ecográfico. En este apartado también se incluye si se hace referencia a la presión aplicada con el transductor sobre la región explorada. Esta presión deberá ser la mínima posible para no deformar la estructura analizada.
6: <i>Características del equipo ecográfico</i>	Detalle de tanto el modelo como la marca utilizada, ya que no todos los equipos ecográficos obtienen la imagen con la misma resolución.
7: <i>Parámetros de optimización de la imagen</i>	Se valora la descripción de los ajustes, frecuencia de trabajo, foco y curvas de ganancia que se establecieron para llevar a cabo las exploraciones. La descripción en particular de estos ajustes es muy importante, ya que influirán directamente en la imagen ecográfica obtenida.
8: <i>Almacenado de imagen</i>	Se evalúa si las imágenes obtenidas en las exploraciones fueron guardadas fuera del ecógrafo para ser analizadas posteriormente y en qué formato se guardaron (PNG, JPEG, TIFF). Algunos formatos guardan las imágenes con pérdidas (compresión) que afectarán a la calidad y a la determinación de determinadas variables de análisis de imagen.
9: <i>Características del explorador</i>	Descripción de quién o quiénes llevaron a cabo las exploraciones ecográficas así como su experiencia en años en este ámbito.

Recientemente¹², los creadores de la escala han publicado valores sobre la fiabilidad de la misma que indican que nueve de los once ítems tienen una fiabilidad entre observadores buena o moderada y que deberían revisarse los otros dos.

Resultados y discusión

Selección de estudios

Tras localizar 227 artículos mediante las estrategias de búsqueda y búsqueda manual, se descartaron 72 trabajos mediante el gestor bibliográfico ZOTERO por estar duplicados. De los 155 restantes, se excluyeron 142 trabajos por no ajustarse a los criterios de selección (fig. 1). En los casos de dificultades o ausencia de consenso entre los dos revisores, un tercer revisor participó en la decisión final. Finalmente se trabajó con una selección de 13 trabajos¹³⁻²⁵.

Características generales de los estudios

Todos los trabajos incluidos en la revisión datan de una fecha posterior a 2003 y la mayoría son de los últimos 4 años^{13-15,20-22,24} lo que podría explicarse por la reciente conciencia de la necesidad de estudios de fiabilidad de calidad a la luz de las recomendaciones STARD²⁶, de forma similar a las PRISMA²⁷, para las revisiones sistemáticas, las STROBE²⁸ para los estudios observacionales o las ya consolidadas CONSORT²⁹ para los ensayos clínicos.

Hay que señalar que pese a que el TA es uno de los tendones que con mayor frecuencia se lesiona, en la mayoría de

los artículos revisados se estudia la fiabilidad de la exploración ecográfica en sujetos sanos asintomáticos, excepto en el artículo de Sunding et al.¹⁴, que estudia la fiabilidad de la ecografía en sujetos sintomáticos. Este es un aspecto interesante que afecta a la validez externa de los trabajos: es importante que los estudios de fiabilidad se realicen sobre muestras similares a la población sobre la que posteriormente se realizarán los análisis³⁰.

Descripción de la reproducibilidad del método ecográfico

En la tabla 2 se muestran los resultados obtenidos para los diferentes ítems en la escala de la elaboración propia para evaluar el grado de reproducibilidad del método ecográfico.

En cuanto a la posición del sujeto, todas las exploraciones se llevaron a cabo en decúbito prono, pero la posición del miembro inferior al llevar a cabo la exploración del TA no queda clara en un 38% de los trabajos. La posición del tobillo y de la rodilla debería describirse siempre con el máximo detalle, ya que el grosor o el área de sección transversal del TA, que son dos de las variables ecográficas que con más frecuencia se analizan, dependen de éstas.

La región anatómica en la cual se colocó la sonda para realizar las exploraciones ecográficas y el corte ecográfico elegido se describen de forma clara en aproximadamente tres cuartas partes de los trabajos (77%). La región anatómica en particular es muy importante, ya que el TA no posee las mismas características a lo largo de todo su recorrido. Por ejemplo, Intziagianni et al.²² llegan a la conclusión de que las medidas a 2 cm proximales al calcáneo son menos reproducibles que a 4 y 6 cm.

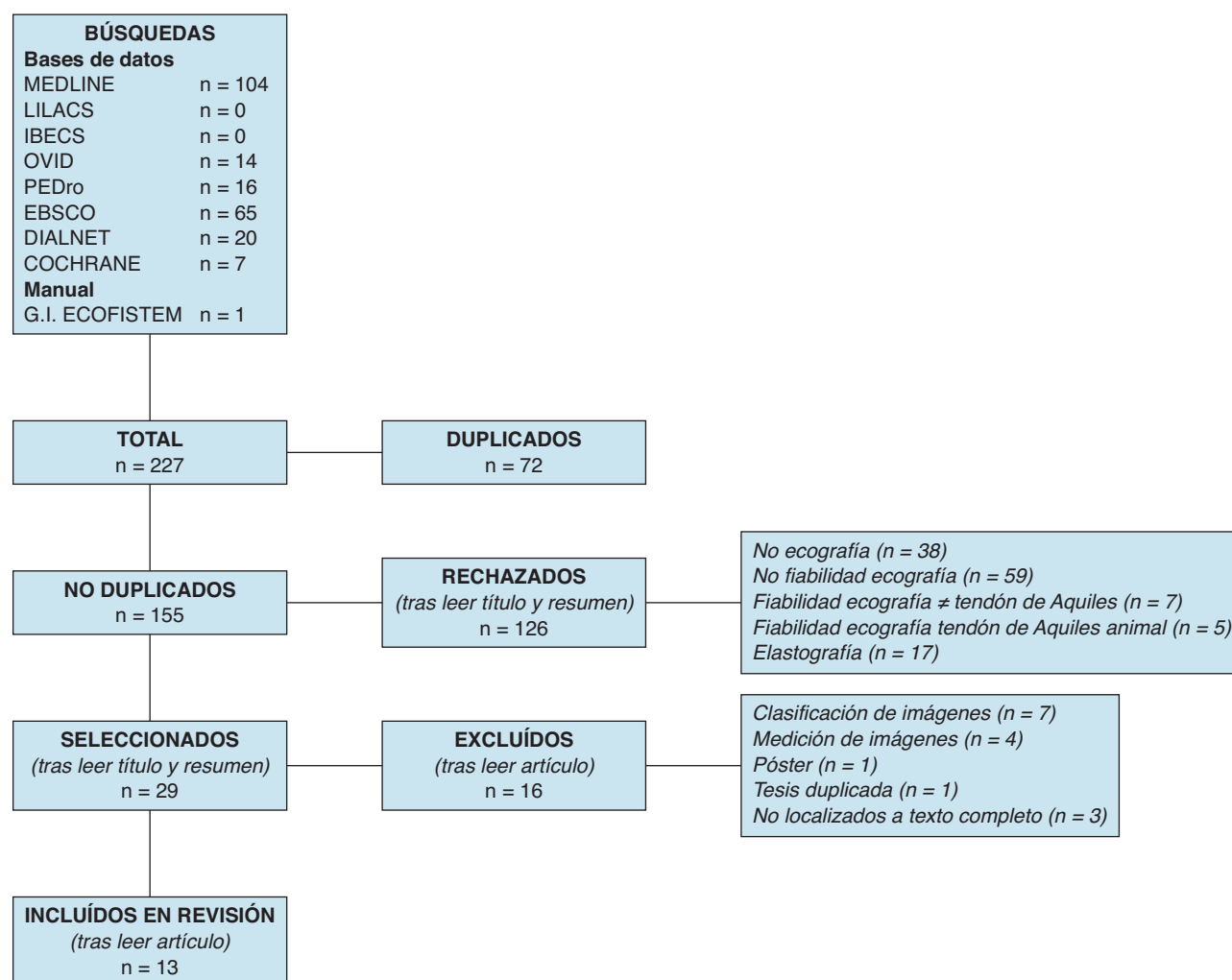


Figura 1 Flujograma del proceso de búsqueda.

El tipo y la frecuencia de la sonda es uno de los ítems que mejor se describen en todos los artículos (92%); una sonda lineal en todos los casos al tratarse de una exploración de tipo musculoesquelético. Además, en el caso del TA hay que tener en cuenta que la zona es irregular y cóncava lo

que hace imposible acoplar totalmente la sonda si ésta es demasiado larga³¹. En esta revisión hemos encontrado que la longitud de la sonda se describe en 6 trabajos^{13,15,16,20,24,25}.

La cantidad de gel y la presión ejercida para llevar a cabo la exploración ecográfica se describe solamente en la mitad de los trabajos. Hay que puntualizar que no tendría mucho sentido ni sería práctico fijar una cantidad cuantitativa de gel, por lo que la cantidad de gel idónea sería la suficiente para permitir el acoplamiento acústico que evite el rebote y la dispersión de los ultrasonidos entre el transductor y la piel. Los autores de los trabajos deberían describir de una u otra forma si al menos tuvieron en cuenta este aspecto.

Con la presión ejercida sobre el transductor ocurre algo similar: el explorador debería ser consciente de ejercer la mínima presión para asegurar la transducción y no modificar la forma de la estructura que afectará especialmente a variables como el grosor o a parámetros morfológicos que evalúen la forma o circularidad del mismo.

Debe destacarse que solamente Tweedell et al.²¹ describen la utilización de gel y la presión ejercida

El equipo ecográfico se describe de forma clara en todos los artículos. Este ítem no es el que más afecta a la reproducibilidad del estudio, pero sí puede interferir en la imagen obtenida y con ello en los resultados, en especial en aque-

Tabla 2 Frecuencias de cumplimiento de los ítems de explicación de la metodología ecográfica

Ítem	No	Dudoso	Sí
1. Posición del sujeto	0 (0%)	5 (38,5%)	8 (61,5%)
2. Región anatómica	0 (0%)	3 (23,1%)	10 (76,9%)
3. Corte ecográfico	0 (0%)	3 (23,1%)	10 (76,9%)
4. Sonda	1 (7,7%)	0 (0%)	12 (92,3%)
5. Gel	7 (53,8%)	5 (38,5%)	1 (7,7%)
6. Equipo	0 (0%)	0 (0%)	13 (100%)
7. Parámetros ecográficos	6 (46,2%)	5 (38,5%)	2 (15,4%)
8. Imagen	4 (30,8%)	5 (38,5%)	4 (30,8%)
9. Explorador	5 (38,5%)	6 (46,2%)	2 (15,4%)

llos casos en los que se cuantifique la intensidad de los niveles de gris de las estructuras.

Los parámetros de optimización de la imagen ecográfica no suelen describirse (46%), o se hace de forma dudosa (38%) y son críticos en los estudios ecográficos de cuantificación de la ecogenicidad o en los análisis texturales. Además, debería especificarse si estos se cambiaron durante las exploraciones o permanecieron constantes, ya que influirán directamente sobre la imagen obtenida. Los parámetros sólo fueron descritos de forma clara en dos artículos: Sunding et al.¹⁴ y Obst et al.¹⁵.

Respecto a la manipulación de los registros ecográficos no queda claro en muchos de los trabajos si las imágenes fueron guardadas en dispositivos de almacenamiento externo para su posterior análisis, ni en qué formato. En este sentido es importante tener en cuenta que existen formatos con pérdida y formatos sin pérdida. En el caso de que se vaya a realizar una cuantificación de la ecogenicidad, el formato de elección debería ser sin pérdida (DICOM, TIFF, BMP)³².

La experiencia del explorador o exploradores se describe de forma clara en aproximadamente una cuarta parte de los trabajos. Lo ideal, que sólo se explicita en dos de los artículos (Intziegianni et al.²² y Barfod et al.²⁴), es que se detallen los años de experiencia del explorador en este ámbito.

Aunque la lectura de las puntuaciones totales en este tipo de escalas debe hacerse con precaución porque no todos los ítems tienen el mismo peso, puede observarse en la tabla 3 que ninguno de los trabajos evaluados obtiene la puntuación máxima de 18 puntos. Sin embargo, los trabajos de Intziegianni et al.²² y de Obst et al.¹⁵ alcanzan los 15 puntos.

Parece claro que deberían detallarse más los métodos de exploración ecográfica y, quizá, desarrollar listas de recomendaciones consensuadas por expertos y específicas para el uso de la ecografía.

Calidad metodológica de los estudios de fiabilidad

En la tabla 4 y en la figura 2 se muestran los resultados del grado de cumplimiento de la escala QAREL.

En relación con la representatividad de la muestra, solamente en 7 estudios es similar a la población sobre la que luego se actuará. En este apartado tenemos que destacar que sólo Sunding et al.¹⁴ evaluaron la fiabilidad de la ecografía en pacientes con sintomatología. A nuestro parecer, deberían realizarse más estudios de fiabilidad en TA patológicos por las implicaciones obvias que esto tendría para la práctica clínica diaria.

En la mayoría de los trabajos tampoco se especifican las características de los exploradores, ni en cuanto a formación ni a los años de experiencia, ni si tienen la misma condición que los exploradores que implementarán la exploración en la práctica clínica (fig. 2).

La información que se proporciona sobre el cegamiento de los exploradores, en especial cuando hay más de uno, es insuficiente. Esto podría comprometer la calidad del estudio al sobreestimar el acuerdo entre exploradores, lo cual podría ser causa de invalidación del estudio. El cegamiento del explorador respecto a los criterios de selección y a posibles características clínicas de los pacientes debería ser el máximo posible. En los trabajos revisados hemos encontrado que sólo en un 38% se especifica el cegamiento interobservador y en un 23% el cegamiento intraobservador. Respecto a los demás ítems relativos al cegamiento prácticamente el 90% no lo cumple (fig. 2).

El ítem de control de la estabilidad de la variable se controló en cinco de los trabajos. Este ítem hace referencia a cómo los investigadores controlan los posibles cambios en las estructuras analizadas entre las distintas mediciones realizadas. En los estudios debe explicitarse qué procedimientos o pautas se siguieron para evitar o reducir estos posibles cambios. En cuatro trabajos que cumplen este criterio se advierte a los sujetos del estudio de que no realicen actividades vigorosas las 24-72h previas a la prueba^{15,20,21,23}, y en seis trabajos se deja descansar al sujeto entre 10 y 30 minutos para intentar controlar estos cambios en el tendón^{13,15,20,21,23}.

En todos los estudios excepto en uno se realizan análisis cuantitativos de la ecografía^{13-18,20-25}. Para determinar la fia-

Tabla 3 Matriz de datos sobre el grado de explicación del método ecográfico

Autor	Año	I1	I2	I3	I4	I5	I6	I7	I8	I9	Total
Tweedell et al. ²⁹	2016	2	2	2	2	2	2	1	1	0	14
Intziegianni et al. ²⁵	2015	2	2	2	2	0	2	1	2	2	15
Basfod et al. ²⁸	2014	2	2	2	0	1	2	1	0	2	12
Sunding et al. ²¹	2014	1	2	1	2	0	2	2	2	1	13
Syha et al. ²⁰	2014	1	2	1	2	0	2	0	2	1	11
Ryan et al. ²⁴	2013	2	2	2	2	1	2	1	1	1	14
Obst et al. ²²	2013	2	2	2	2	1	2	2	1	1	15
Kongsgaard et al. ³⁰	2011	2	2	2	2	1	2	0	1	0	12
Ríos-Díaz et al. ²⁵	2010	1	2	2	2	0	2	0	2	0	11
Brushøj et al. ¹⁷	2006	1	1	2	2	0	2	0	0	0	8
O'Connor et al. ²³	2004	2	1	1	2	0	2	0	0	1	9
Bjordal et al. ³¹	2003	2	2	2	2	0	2	0	1	0	11
Ying et al. ³²	2003	1	1	2	2	1	2	0	0	0	9

I: ítems valorados en el método ecográfico; 0: no cumple; 1: dudoso o incompleto; 2: sí cumple.

Tabla 4 Matriz de cumplimiento de las preguntas de la escala QAREL

Autor	Año	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11
Tweedell et al. ²⁹	2016	S	NC	NA	NC	NA	NC	NC	NC	S	S	S
Intziegianni et al. ²⁵	2015	NC	NC	NA	NC	NA	S	NC	NC	NC	S	S
Basfod et al. ²⁸	2014	N	S	S	S	NA	NC	NC	NC	NC	S	NC
Sunding et al. ²¹	2014	S	S	S	NA	NA	NC	NC	NC	NC	S	NC
Syha et al. ²⁰	2014	N	S	NC	NC	NA	NC	S	NC	S	S	NC
Ryan et al. ²⁴	2013	N	NC	NA	NC	NA	NC	NC	NC	S	S	NC
Obst et al. ²²	2013	S	NC	NA	N	NA	NC	NC	N	S	S	S
Kongsgaard et al. ³⁰	2011	N	NC	NA	NC	NA	NC	NC	NC	NC	S	N
Ríos-Díaz et al. ²⁵	2010	S	S	S	S	NA	NC	S	NC	NC	S	S
Brushøj et al. ¹⁷	2006	S	NC	NC	S	S	NC	NC	NC	NC	N	N
O'Connor et al. ²³	2004	N	S	S	NC	NA	NC	NC	S	S	NC	NC
Bjordan et al. ³¹	2003	S	S	NA	NC	NA	NC	NC	NC	NC	NC	N
Ying et al. ³²	2003	S	NC	S	NA	NA	NC	NC	NC	NC	S	NC

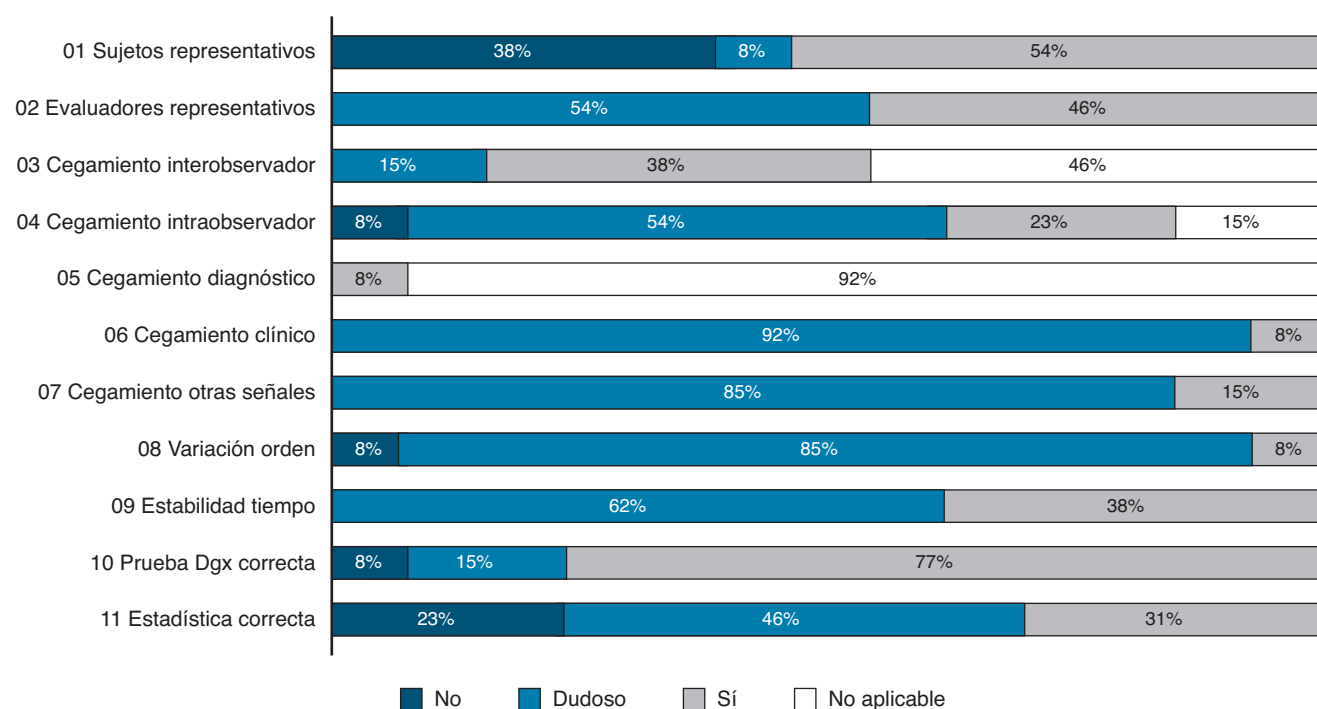
P: preguntas de la escala QAREL; S: sí cumple; N: no cumple; NC: no claro; NA: no aplicable.

bilidad de estas variables, los autores suelen utilizar el coeficiente de correlación intraclase (CCI)^{25,33}.

Aunque el uso de este coeficiente es correcto, diferentes autores afirman que el diseño del estudio de los estudios de fiabilidad es, con frecuencia, incorrecto o incompleto^{34,35}.

Por ejemplo, debería describirse cómo se estimó el tamaño de la muestra para asegurar una potencia estadística suficiente^{36,37}, o tener en cuenta que algunas de estas pruebas (el CCI también) se basan en el supuesto de la distribución normal de los datos, algo que debería comprobarse y, en caso de no cumplirse, realizar los ajustes necesarios^{38,39}.

En el caso del CCI, Shrout y Fleiss (1979)³⁹ describieron hasta seis formas diferentes que pueden resumirse en tres modelos generales y que deberían quedar perfectamente descritos en la sección de métodos. Solamente en cuatro de los trabajos revisados^{15,21,22,25} se explicita el tipo de modelo utilizado para el cálculo del CCI. Por ejemplo, los autores deberían indicar si han calculado el CCI buscando el acuerdo absoluto (sería lo ideal) o buscando la consistencia entre mediciones. El cálculo del CCI se basa en los modelos de análisis de la variancia (ANOVA) por lo que debería indicarse si se ha utilizado un ANOVA de una o dos vías y si los ICC se han calcu-

**Figura 2** Calidad metodológica de los estudios según la escala QAREL.

lado sobre mediciones simples o sobre la media de un conjunto de ellas. Básicamente, los distintos modelos del cálculo del CCI se diferencian en cómo se consideran los errores sistemáticos y aleatorios en la ecuación del modelo. Este aspecto, que puede parecer aparentemente banal, tiene implicaciones metodológicas y teóricas importantes. Además del propio trabajo de Shout y Fleiss³⁹, Weir³⁸ y McGraw y Wong⁴⁰ describen en sendos trabajos y de una forma exhaustiva las implicaciones que existen tras el cálculo de los CCI.

Aun así, el CCI no está exento de críticas^{38,40,41} y debería complementarse su uso con otros índices como el error estándar de la media y la mínima diferencia detectable que son independientes del CCI y tienen un interés práctico porque permiten conocer los límites de detección de los cambios del protocolo de medición. El método visual descrito por Bland y Altman (1986)⁴² también se considera útil para determinar la precisión de las medidas y detectar posibles sesgos sistemáticos en los registros.

En el trabajo de Bjordal et al.¹⁹ y en el de Kongsgaard et al.¹⁶ los autores utilizan el coeficiente de correlación lineal de Pearson como índice estadístico de reproducibilidad. Este índice estadístico no debe usarse con ese cometido porque, si bien indica consistencia entre las medidas, no tiene en cuenta las diferencias que puedan existir entre las medidas repetidas³³.

Finalmente, nos encontramos que Brushøj et al.¹⁷, para determinar la fiabilidad, calculan el coeficiente de variación. Este estadístico no puede interpretarse como indicador de concordancia entre medidas y sí como medida de la variabilidad de las observaciones respecto a la media grupal; tiene otros usos diferentes a la determinación de la fiabilidad entre medidas.

Pese a la creencia por parte de algunos investigadores de que la metodología ecográfica en las exploraciones del TA está perfectamente instaurada y validada, esto no es así y la mayoría de los estudios que la analizan adolecen de serios problemas metodológicos en aspectos tan relevantes como la representatividad de la muestra, el cegamiento de los exploradores y el adecuado análisis estadístico de los datos.

Finalmente concluimos que la descripción del método ecográfico, en general, no está suficientemente desarrollada, lo que afecta directamente a la reproducibilidad del método y a la comparación entre estudios. Sería recomendable generar guías de recomendación consensuadas por expertos y específicas para las exploraciones ecográficas.

Bibliografía

1. Arya S, Kulig K. Tendinopathy alters mechanical and material properties of the Achilles tendon. *J Appl Physiol* 2010;108(3):670-5.
2. Alfredson H. The chronic painful Achilles and Patellar tendon: research on basic biology and treatment. *Scand J Med Sci Sports* 2005;15:252-9.
3. Kader D, Saxena A, Movin T, Maffulli N. Achilles tendinopathy: some aspects of basic science and clinical management. *Br J Sports Med* 2002;36(4):239-49.
4. Jiménez Díaz JF. Diagnóstico clínico y ecográfico de las lesiones del deporte. Murcia: Universidad Católica San Antonio; 2002.
5. Minaya-Muñoz F, Valera-Garrido F, Benito-Domingo A. Ecografía musculoesquelética en fisioterapia. En: Valera F, Minaya F, Editores. *Fisioterapia invasiva*. Barcelona: Elsevier; 2013. p. 99-106.
6. Martínez Paya JJ. Anatomía ecográfica del hombro: herramienta de prevención, diagnóstico, investigación y validación de técnicas terapéuticas. Madrid: Editorial Médica Panamericana; 2008.
7. Hurley WL, Denegar CR, Hertel J. Research methods: a framework for evidence-based clinical practice. 1st ed. Philadelphia: Wolters Kluwer/Lippincott Williams & Wilkins Health; 2011.
8. Lexell JE, Downham DY. How to Assess the Reliability of Measurements in Rehabilitation: *Am J Phys Med Rehabil* 2005;84(9):719-23.
9. Charter RA, Feldt LS. Meaning of reliability in terms of correct and incorrect clinical decisions: the art of decision making is still alive. *J Clin Exp Neuropsychol* 2001;23(4):530-7.
10. Urrútia G, Bonfill X. Declaración PRISMA: una propuesta para mejorar la publicación de revisiones sistemáticas y metaanálisis. *Med Clínica* 2010;135(11):507-11.
11. Lucas NP, Macaskill P, Irwig L, Bogduk N. The development of a quality appraisal tool for studies of diagnostic reliability (QAREL). *J Clin Epidemiol* 2010;63(8):854-61.
12. Lucas N, Macaskill P, Irwig L, Moran R, Rickards L, Turner R, et al. The reliability of a quality appraisal tool for studies of diagnostic reliability (QAREL). *BMC Med Res Methodol* 2013;13(1):1.
13. Syha R, Grau S, Nieß AM, Hein T, Werner MK, Ketelsen D, et al. Computer-based quantification of the Achilles tendon thickness in sequential B-mode ultrasound images: a study of feasibility and reliability. *Arch Orthop Trauma Surg* 2014;134(10):1443-9.
14. Sunding K, Fahlström M, Werner S, Forssblad M, Willberg L. Evaluation of Achilles and patellar tendinopathy with greyscale ultrasound and colour Doppler: using a four-grade scale. *Knee Surg Sports Traumatol Arthrosc* 2016;24(6):1988-96.
15. Obst SJ, Newsham-West R, Barrett RS. In Vivo Measurement of Human Achilles Tendon Morphology Using Freehand 3-D Ultrasound. *Ultrasound Med Biol* 2014;40(1):62-70.
16. Kongsgaard M, Nielsen CH, Hegnsvad S, Aagaard P, Magnusson SP. Mechanical properties of the human Achilles tendon, in vivo. *Clin Biomech* 2011;26(7):772-7.
17. Brushøj C, Henriksen BM, Albrecht-Beste E, Hölmich P, Larsen K, Bachmann Nielsen M. Reproducibility of Ultrasound and Magnetic Resonance Imaging Measurements of Tendon Size. *Acta Radiol* 2006;47(9):954-9.
18. Ying M, Yeung E, Li B, Li W, Lui M, Tsoi C-W. Sonographic evaluation of the size of Achilles tendon: the effect of exercise and dominance of the ankle. *Ultrasound Med Biol* 2003;29(5):637-42.
19. Bjordal JM, Demmink JH, Ljunggren AE. Tendon thickness and depth from skin for supraspinatus, common wrist and finger extensors, patellar and Achilles tendons: ultrasonography study of healthy subjects. *Physiotherapy* 2003;89(6):375-83.
20. Ryan ED, Rosenberg JG, Scharville MJ, Sobolewski EJ, Thompson BJ, King GE. Test-Retest Reliability and the Minimal Detectable Change for Achilles Tendon Length: A Panoramic Ultrasound Assessment. *Ultrasound Med Biol* 2013;39(12):2488-91.
21. Tweedell AJ, Ryan ED, Scharville MJ, Rosenberg JG, Sobolewski EJ, Kleinberg CR. The influence of ultrasound measurement techniques on the age-related differences in Achilles tendon size. *Exp Gerontol* 2016;76:68-71.
22. Intzeigianni K, Cassel M, König N, Müller S, Fröhlich K, Mayer F. Ultrasonography for the assessment of the structural properties of the Achilles tendon in asymptomatic individuals: An intra-rater reproducibility study. *Isokinet Exerc Sci* 2015;23(4):263-70.
23. O'Connor PJ, Grainger AJ, Morgan SR, Smith KL, Waterton JC, Nash AFP. Ultrasound assessment of tendons in asymptomatic volunteers: a study of reproducibility. *Eur Radiol* 2004;14(11):1968-73.
24. Barfod KW, Riecke AF, Boesen A, Hansen P, Maier JF, Døssing S, et al. Validation of a novel ultrasound measurement of achilles tendon length. *Knee Surg Sports Traumatol Arthrosc* 2015;23(11):3398-406.

25. Ríos-Díaz J, Martínez-Payá JJ, del-Baño-Aledo ME, de-Groot-Ferrando A. Fiabilidad y reproducibilidad intra e inter-observador de un método semiautomático de análisis ecográfico del tendón de Aquiles. *Cuest Fisioter* 2010;39(3):190-8.
26. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. The STARD statement for reporting studies of diagnostic accuracy: Explanation and elaboration. *Clin Chem* 2003;49:7e18.
27. Moher D, Liberati A, Tetzlaff J, Altman DG. The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA Statement. *PLoS Med* 2008;6:e1000097.
28. Vandenbroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and elaboration. *PLoS Med* 4 2007;e297.
29. Moher D, Schulz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA* 2001;285:1987-91.
30. Knottnerus JA, Muris JW. Assessment of the accuracy of diagnostic tests: the cross-sectional study. En: Knottnerus JA, Buntinx F, Editores. *The Evidence Base of Clinical Diagnosis: Theory and Methods of Diagnostic Research*. 2nd ed. Oxford: Blackwell Publishing Ltd; 2009. p. 42-62.
31. Matas RB, Villarrubia MR, Aleu AC. Ecografía muscular de la extremidad inferior: sistemática de exploración y lesiones en le deporte. Madrid: Elsevier; 2005.
32. Ríos-Díaz J, Martínez-Payá JJ, de Groot-Ferrando A, del Baño-Aledo ME. Procesamiento y análisis de imagen en ecografía. En: Valera F, Minaya F, Editores. *Fisioterapia Invasiva*. Barcelona: Elsevier; 2013. p. 359-74.
33. Pita Fernández S, Pérttega Díaz S, Rodríguez Maseda E. La fiabilidad de las mediciones clínicas: el análisis de concordancia para variables numéricas. *Cad Aten Primaria* 2003;10(4): 290-6.
34. English C, Fisher L, Thoires K. Reliability of real-time ultrasound for measuring skeletal muscle size in human limbs in vivo: a systematic review. *Clin Rehabil* 2013;26:934-44.
35. Zaki R, Bulgiba A, Ismail R, Ismail NA. Statistical Methods Used to Test for Agreement of Medical Instruments Measuring Continuous Variables in Method Comparison Studies: A Systematic Review. *PLoS ONE* 2012;7: e37908.
36. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Med* 1998;17:101-10.
37. Shieh G. Optimal sample sizes for the design of reliability studies: Power consideration. *Behav Res Methods* 2014;46: 772-85.
38. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res Natl Strength Cond Assoc* 2005;19:231-40.
39. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979;86: 420-8.
40. McGraw KO, Wong S. Forming inferences about some intraclass correlation coefficients. *Psychol Methods* 1996;1:30-46.
41. Rankin G, Stokes M. Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses. *Clin Rehabil* 1998;12:187-99.
42. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307-10.