ASCE MAGAZINE ISSN: 3073–1178

Revista ASCE Magjjazine, Periodicidad: Trimestral, Volumen: 5, Número: 3, Año: 2026 páginas 2323-2344

Doi: https://doi.org/10.70577/asce.v5i3.1085

Recibido: 2026-07-29

Aceptado: 2026-08-12

Publicado: 2026-08-26

Exploración clínica en las patologías del codo: revisión sistemática sobre la sensibilidad y especificidad de los tests ortopédicos

Clinical examination in elbow pathologies: a systematic review on the sensitivity and specificity of orthopedic tests

Autor(s)

María De Los Ángeles Alulima Cuenca[1]
Facultad de Ciencias de la Salud
mdalulima@utpl.edu.ec
https://orcid.org/0009-0004-3615-7736
Universidad Técnica Particular de Loja
Loja - Ecuador

Britanny Nicole Castillo Veintimilla[2]
Facultad de Ciencias de la Salud
bncastillo2@utpl.edu.ec
https://orcid.org/0009-0002-6679-5478
Universidad Técnica Particular de Loja
Loja - Ecuador

Fernando Benjamín Herrera Chacón[3]
Facultad de Ciencias de la Salud
fbherrera2@utpl.edu.ec
https://orcid.org/0009-0006-4785-876X
Universidad Técnica Particular de Loja
Loja - Ecuador

Dylan Geovanny Erique Valdivieso[4]
Facultad de Ciencias de la Salud
dgerique@utpl.edu.ec
https://orcid.org/0009-0009-5045-3466
Universidad Técnica Particular de Loja
Loja - Ecuador

Geovanny Patricio Cárdenas Caicedo[5]
Facultad de Ciencias de la Salud
gpcardenas6@utpl.edu.ec
https://orcid.org/0009-0009-3496-7467
Universidad Técnica Particular de Loja
Loja – Ecuador

Cómo citar

Alulima Cuenca, M. D. L. Ángeles, Castillo Veintimilla, B. N., Herrera Chacón, F. B., Erique Valdivieso, D. G., & Cárdenas Caicedo, G. P. (2026). Exploración clínica en las patologías del codo: revisión sistemática sobre la sensibilidad y especificidad de los tests ortopédicos. *ASCE MAGAZINE*, *5*(3), 2323–2344. https://doi.org/10.70577/asce.v5i3.1085

Esta obra está bajo una Licencia Creative Commons Atribución-No Comercial-Compartir Igual 4.0 Internacional

Resumen

Las patologías del codo constituyen una causa frecuente de dolor, discapacidad funcional y ausentismo laboral; la exploración clínica es el primer filtro diagnóstico, pero la precisión de sus pruebas es heterogénea y varía según la patología y la maniobra empleada. El objetivo fue analizar la evidencia científica sobre la sensibilidad y especificidad de los principales tests ortopédicos empleados en la valoración de las patologías del codo, con el fin de identificar las pruebas de mayor rendimiento diagnóstico y fortalecer el razonamiento clínico basado en la evidencia.

Se realizó una revisión sistemática de estudios de precisión diagnóstica adherida a las guías PRISMA 2020 y PRISMADTA, con búsqueda en PubMed y PubMed Central (2015–2025) combinando términos MeSH y libres con operadores booleanos; se incluyeron revisiones sistemáticas, estudios de precisión diagnóstica y revisiones clínicas que reportaran sensibilidad, especificidad o rendimiento diagnóstico de pruebas de exploración física del codo, valorando la calidad metodológica con QUADAS2 y realizando síntesis narrativa.

Se incluyeron diez estudios el Hook Test rindió mejor en roturas completas del bíceps distal (sensibilidad y especificidad del 100% contra cirugía, 96% y 67% en validación prospectiva), Moving Valgus Stress Test alcanzó 100% y 75% de sensibilidad y especificidad en lesiones del ligamento colateral medial, el Elbow Extension Test alcanzó 98,6% de sensibilidad para descartar fracturas; las maniobras clásicas para epicondilalgia lateral (Cozen, Mill, Maudsley) mostraron rendimiento variable y la prensión sin dolor aportó 78–83% y 80–90% de sensibilidad y especificidad.

Se concluyo que ninguna prueba aislada es diagnósticamente suficiente; la precisión aumenta al combinar maniobras con la anamnesis y cuando procede, la imagen, por lo que la selección del test debe responder a la sospecha clínica y a la mejor evidencia disponible en cada contexto asistencial.

Palabras clave: Codo; exploración física; pruebas ortopédicas; sensibilidad y especificidad.

Esta obra está bajo una Licencia Creative Commons Atribución-No Comercial-Compartir Igual 4.0 Internacional

Abstract

Elbow disorders represent a frequent cause of pain, functional disability, and occupational absenteeism. While clinical examination serves as the primary diagnostic screen, the accuracy of its physical tests remains heterogeneous and varies depending on the specific pathology and maneuver performed. The objective of this study was to analyze the scientific evidence on the sensitivity and specificity of the main orthopedic tests used in the assessment of elbow pathology, in order to identify the maneuvers with the highest diagnostic yield and strengthen evidencebased clinical reasoning.

A systematic review of diagnostic accuracy studies was conducted following the PRISMA 2020 and PRISMADTA guidelines. Comprehensive searches were performed in PubMed and PubMed Central (2015–2025) using a combination of MeSH and freetext terms joined by Boolean operators. Systematic reviews, diagnostic accuracy studies, and clinical reviews reporting the sensitivity, specificity, or diagnostic performance of physical examination tests for the elbow were included. Methodological quality was evaluated using the QUADAS2 tool, and a narrative synthesis was performed.

Ten studies met the inclusion criteria. The Hook Test demonstrated the highest diagnostic performance for complete distal biceps tendon ruptures (100% sensitivity and specificity when compared to surgical findings, and 96% sensitivity and 67% specificity in prospective validation). The Moving Valgus Stress Test achieved 100% sensitivity and 75% specificity for medial collateral ligament injuries, while the Elbow Extension Test showed 98.6% sensitivity for ruling out fractures. Classic maneuvers for lateral epicondylalgia (Cozen's, Mill's, and Maudsley's tests) exhibited variable diagnostic performance, whereas painfree grip strength yielded a sensitivity of 78–83% and a specificity of 80–90%.

In conclusion, no single test is diagnostically sufficient on its own. Diagnostic accuracy increases when physical maneuvers are combined with a detailed clinical history and, when appropriate, diagnostic imaging. Therefore, test selection should align with the clinical suspicion and the best available evidence in each healthcare setting.

Keywords: Elbow; physical examination; orthopedic tests; sensitivity and specificity.

Esta obra está bajo una Licencia Creative Commons Atribución-No Comercial-Compartir Igual 4.0 Internacional

Introducción

El codo es la articulación intermedia del miembro superior, encargada de transmitir y disipar cargas entre el hombro y la mano, y se encuentra expuesta de forma sostenida a sobrecargas laborales y deportivas que generan afecciones musculoesqueléticas de alto impacto funcional (Zwerus et al., 2018). El deterioro de las estructuras tendinosas, ligamentarias y neurales compromete actividades cotidianas y laborales levantamiento de peso, prensión, movimientos repetitivos de muñeca y antebrazo y se traduce en absentismo laboral y disminución de la calidad de vida de la población adulta activa (Curti et al., 2021).

Desde el punto de vista epidemiológico, el estudio poblacional de Shiri et al. (2006), sobre una muestra representativa de 4783 personas de 30 a 64 años, estimó una prevalencia de epicondilitis lateral del 1,3% y de epicondilitis medial del 0,4%, con pico entre los 45 y los 54 años y asociación independiente con tabaquismo, movimientos repetitivos y actividades de fuerza, lo que confirma la relevancia sanitaria y ocupacional de estas entidades.

Anatómicamente, el codo constituye un complejo integrado por las articulaciones humerocubital, humerorradial y radiocubital proximal, englobadas en una única cápsula, y su estabilidad pasiva depende de la congruencia ósea y de los ligamentos colaterales: el complejo ligamentario medial resiste el estrés en valgo, especialmente demandante en los gestos de lanzamiento, mientras que el complejo lateral controla la inestabilidad rotatoria posterolateral (Pottanat et al., 2025; Zwerus et al., 2018).

La estabilidad dinámica recae en los grupos flexores y extensores y en la musculatura epicondílea, cuya contracción genera compresión articular protectora; cuando la demanda repetitiva ocupacional o deportiva supera la capacidad tisular, se desencadenan microtraumatismos y degeneración (Curti et al., 2021; Kheiran et al., 2021).

En coherencia con este modelo, Curti et al. (2021) hallaron evidencia limitada de relación causal entre la sobrecarga biomecánica ocupacional y la tendinopatía lateral del codo con valores de Strain Index superiores a 5–6,1 que podrían duplicar el riesgo e insuficiente para la tendinopatía medial. El espectro patológico del codo es amplio: entre las tendinopatías, la epicondilalgia lateral es la alteración más frecuente y la medial resulta sustancialmente

Esta obra está bajo una Licencia Creative Commons Atribución-No Comercial-Compartir Igual 4.0 Internacional

menos habitual (Di Filippo et al., 2022; Karanasios et al., 2022), y ambas comparten un trasfondo degenerativo más que inflamatorio (Curti et al., 2021; Kheiran et al., 2021).

En el extremo agudo se sitúan la rotura distal del bíceps braquial, cuyo diagnóstico precoz es imprescindible para evitar el déficit de supinación y flexión (Baylor et al., 2023; Caekebeke et al., 2022), las lesiones del ligamento colateral medial y el dolor medial no traumático en lanzadores (Adani et al., 2025; Pottanat et al., 2025), las fracturas pericubitales tras traumatismos (Breda et al., 2023) y el compromiso neural, representado por el síndrome del túnel cubital (Novak et al., 1994; Zwerus et al., 2018).

Frente a esta diversidad, la anamnesis y la exploración física constituyen el primer filtro diagnóstico antes de indicar estudios de imagen (Curti et al., 2021; Kheiran et al., 2021) y los tests ortopédicos pretenden reproducir de forma estandarizada el dolor, la laxitud o el déficit neurológico, cuantificándose su utilidad mediante indicadores de precisión diagnóstica.

La sensibilidad (Se) representa la proporción de individuos enfermos identificados correctamente por una prueba clínica; por tanto, un resultado negativo en una prueba altamente sensible permite descartar la patología, por su parte, la especificidad (Sp) corresponde a la proporción de sujetos sanos clasificados de manera precisa, de modo que un resultado positivo en un test con alta especificidad contribuye a confirmar el diagnóstico.

A partir de ambas se derivan el valor predictivo positivo (VPP) y el valor predictivo negativo (VPN), probabilidades posteriores a la prueba que dependen de la prevalencia del trastorno en el contexto asistencial, y las razones de verosimilitud, que cuantifican cuánto modifica un resultado la probabilidad preprueba, dichas razones ($RV+ = Se/(1-Sp)$; $RV- = (1-Se)/Sp$) son independientes de la prevalencia, por lo que constituyen el indicador más robusto para comparar pruebas entre estudios (McInnes et al., 2018), y en términos prácticos solo pruebas con RV+ elevada o RV− muy próxima a cero modifican sustancialmente la decisión clínica.

La literatura muestra, sin embargo, una marcada disparidad de rendimiento: para el dolor lateral, las maniobras de Cozen, Mill y Maudsley están ampliamente extendidas, pero Karanasios et al. (2022) no identificaron ninguna prueba con sensibilidad aislada suficiente

Esta obra está bajo una Licencia Creative Commons Atribución-No Comercial-Compartir Igual 4.0 Internacional

para descartar la tendinopatía, lo que exige integrarlas con la prensión sin dolor y otros hallazgos (Di Filippo et al., 2022).

En contraste, el Hook Test ha mostrado sensibilidad y especificidad de hasta el 100% en avulsiones completas del bíceps distal (O'Driscoll et al., 2007), aunque su exactitud decae en roturas parciales (Caekebeke et al., 2022), y el Moving Valgus Stress Test alcanzó sensibilidad del 100% y especificidad del 75% en lesiones del ligamento colateral medial (O'Driscoll et al., 2005).

El Elbow Extension Test se posiciona como la maniobra más útil para descartar fracturas agudas (Breda et al., 2023), mientras que en el síndrome del túnel cubital el rendimiento de las pruebas provocativas aisladas es moderado y mejora sustancialmente al combinar compresión y flexión mantenida (Novak et al., 1994).

A pesar de ello, las revisiones disponibles advierten heterogeneidad metodológica y terminológica: Di Filippo et al. (2022) identificaron diez denominaciones distintas para el dolor lateral del codo y doce tests diferentes en las revisiones analizadas, lo que dificulta la comparación y genera incertidumbre en la selección de pruebas en la práctica fisioterapéutica. Existe, por tanto, un vacío de síntesis actualizada y jerarquizada que ordene el valor real de cada maniobra según su precisión verificada, y la pregunta de investigación fue: ¿cuál es la sensibilidad y especificidad de los principales tests ortopédicos utilizados en la exploración clínica de las patologías del codo en la población adulta, según la evidencia publicada entre 2015 y 2025?

En consecuencia, el objetivo de esta revisión sistemática fue analizar la sensibilidad y la especificidad de los principales tests clínicos empleados en el diagnóstico de las patologías del codo, con el fin de identificar las pruebas de mayor rendimiento diagnóstico y fortalecer el razonamiento clínico fisioterapéutico basado en la evidencia.

Esta obra está bajo una Licencia Creative Commons Atribución-No Comercial-Compartir Igual 4.0 Internacional

Material y métodos

Se realizó una revisión sistemática de estudios de precisión diagnóstica, conducida y reportada con adherencia a la declaración PRISMA 2020 (Page et al., 2021) y a su extensión específica para estudios de exactitud diagnóstica, PRISMADTA (McInnes et al., 2018); dado el objetivo sintetizar sensibilidad y especificidad de pruebas de exploración física, este marco garantizó la declaración explícita de la pregunta, la estrategia de búsqueda, los estándares de referencia y los índices de precisión extraídos.

La pregunta de investigación se estructuró en formato PICO del siguiente modo:

- **Población (P):** personas adultas con sospecha clínica de patología del codo (tendinopatías epicondíleas lateral o medial, rotura distal del bíceps, lesión del ligamento colateral medial, fractura aguda o síndrome del túnel cubital).
- **Pruebas índice (I):** tests ortopédicos de exploración física del codo (entre otros, Cozen, Mill, Maudsley, prensión sin dolor, Hook Test, Biceps Crease Interval, Moving Valgus Stress Test, estrés en valgo, Milking Maneuver, Elbow Extension Test, Elbow Flexion Test, signo de Tinel y pruebas de compresión neural).
- **Comparador/estándar de referencia (C):** confirmación quirúrgica, resonancia magnética, ecografía, radiografía convencional o estudios electrodiagnósticos, según la patología.
- **Resultados (O):** sensibilidad, especificidad y, cuando estuvieron disponibles, valores predictivos y razones de verosimilitud.

La búsqueda bibliográfica se ejecutó en las bases de datos PubMed y PubMed Central, con publicaciones comprendidas entre enero de 2015 y diciembre de 2025 (última búsqueda ejecutada en diciembre de 2025), y se aplicaron filtros de idioma (inglés y español) y de tipo de publicación acordes con los criterios de elegibilidad.

Los terminos de búsqueda combinó términos controlados y libres mediante operadores booleanos: ((“elbow”[MeSH Terms] OR “elbow”[Title/Abstract]) AND (“physical examination”[MeSH Terms] OR “clinical examination”[Title/Abstract] OR “diagnostic

Esta obra está bajo una Licencia Creative Commons Atribución-No Comercial-Compartir Igual 4.0 Internacional

test"[Title/Abstract] OR "orthopedic test"[Title/Abstract]) AND ("sensitivity and specificity"[MeSH Terms] OR "sensitivity"[Title/Abstract] OR "specificity"[Title/Abstract] OR "diagnostic accuracy"[Title/Abstract])) OR ("Cozen test"[Title/Abstract] OR "Mill test"[Title/Abstract] OR "Maudsley test"[Title/Abstract] OR "Hook test"[Title/Abstract] OR "Moving Valgus Stress Test"[Title/Abstract] OR "Elbow Extension Test"[Title/Abstract] OR "lateral elbow tendinopathy"[Title/Abstract] OR "medial elbow pain"[Title/Abstract]).

Criterios de elegibilidad

Se incluyeron artículos originales de precisión diagnóstica (prospectivos o retrospectivos), revisiones sistemáticas, *overviews* de revisiones y revisiones clínicas con base sistemática de literatura, publicados entre 2015 y 2025, en inglés o español, que evaluaran pruebas de exploración clínica del codo y reportaran información sobre sensibilidad, especificidad o rendimiento diagnóstico frente a un estándar de referencia explícito.

Se excluyeron artículos duplicados, publicaciones sin acceso al texto completo, cartas al editor, resúmenes de congreso, estudios centrados exclusivamente en tratamiento sin datos diagnósticos y trabajos que no abordaran pruebas de exploración física del codo; las referencias clásicas anteriores a 2015 (por ejemplo, estudios de validación original de las maniobras) no fueron elegibles como estudios incluidos, pero se citaron como evidencia de apoyo cuando las revisiones incluidas las sintetizaban.

Proceso de selección de estudios

Tras la eliminación de duplicados, dos revisores cribaron de forma independiente los títulos y resúmenes; los registros potencialmente elegibles se evaluaron a texto completo y las discrepancias se resolvieron por consenso con un tercer revisor, documentándose el proceso en la Figura 1 (diagrama de flujo PRISMA 2020).

Se identificaron 286 registros (PubMed, n = 174; PubMed Central, n = 112) y se eliminaron 41 duplicados; de los 245 registros cribados por título y resumen se excluyeron 198 por no cumplir los criterios temáticos o de diseño, y se evaluaron 47 informes a texto completo.

De esos informes, 37 fueron excluidos con motivos (ausencia de datos de sensibilidad/especificidad, n = 14; estudio centrado en tratamiento, n = 9; sin estándar de

Esta obra está bajo una Licencia Creative Commons Atribución-No Comercial-Compartir Igual 4.0 Internacional

referencia válido, n = 7; tipo de publicación no elegible, n = 4; idioma distinto de inglés/español, n = 3) y, finalmente, se incluyeron 10 estudios en la síntesis cualitativa.

Mediante una plantilla estandarizada se extrajeron, para cada estudio incluido, autor y año de publicación, diseño, tamaño muestral o número de estudios agregados, objetivo, patología evaluada, pruebas clínicas analizadas, estándar de referencia, valores de sensibilidad y especificidad (con intervalos de confianza cuando se reportaban) y principales hallazgos, verificando la extracción un segundo revisor.

La calidad metodológica y el riesgo de sesgo de los estudios incluidos se valoraron con la herramienta QUADAS2 (Whiting et al., 2011), que examina cuatro dominios selección de pacientes, prueba índice, estándar de referencia y flujo y tiempos junto con la aplicabilidad.

Dada la heterogeneidad clínica (patologías y pruebas muy diversas), metodológica (diseños primarios y síntesis secundarias) y estadística (estándares de referencia y formas de reporte dispares) de los estudios incluidos, se optó por una síntesis narrativa cualitativa con tabulación de los valores de precisión, descartándose el metaanálisis cuantitativo por la imposibilidad de agrupar estimadores comparables sin incurrir en conclusiones espurias.

Resultados

Selección de los estudios

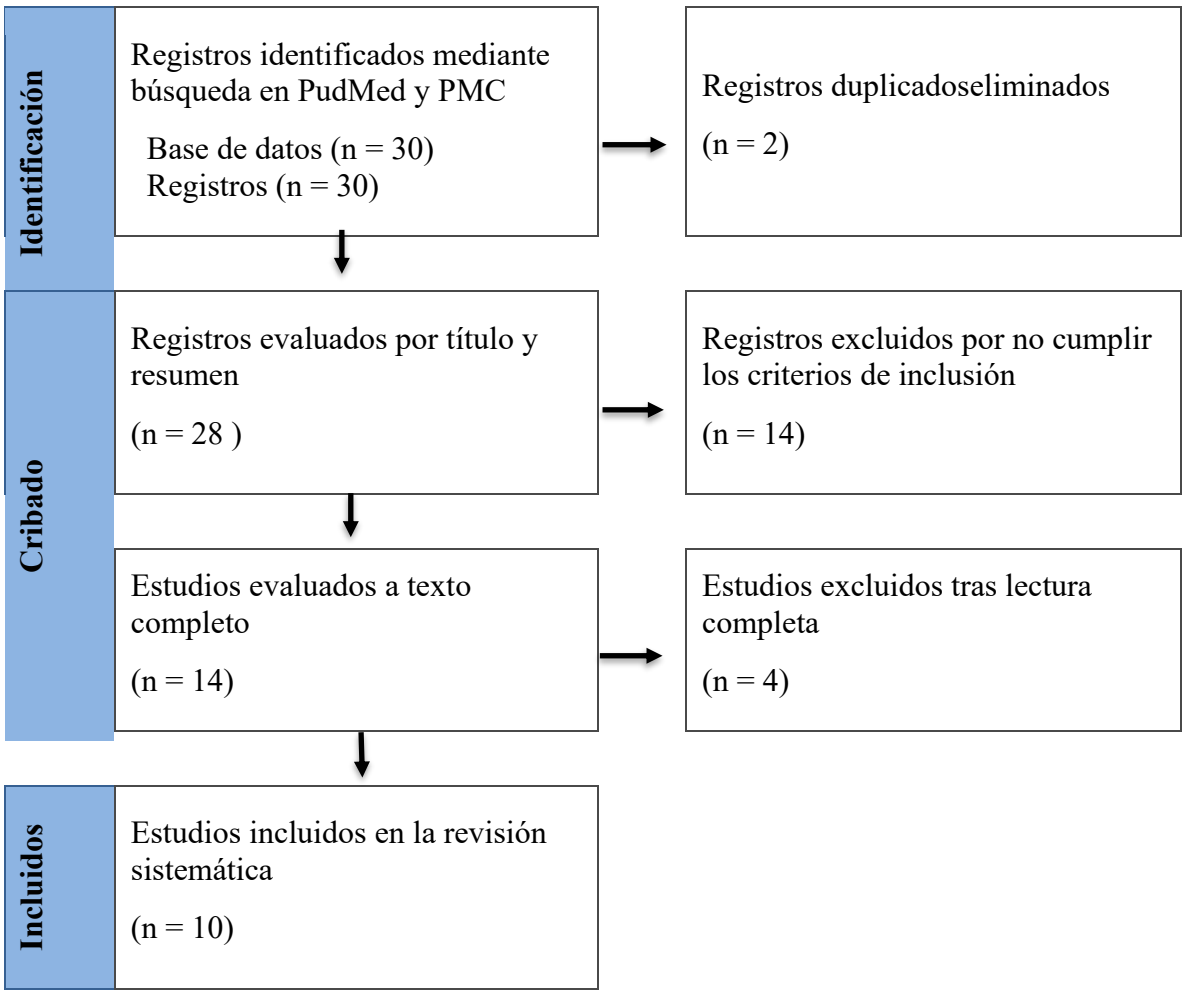
El proceso de identificación y selección se resume en la Figura 1: de los 286 registros identificados en las dos bases de datos, el cribado por título y resumen y la evaluación a texto completo condujeron a la inclusión final de 10 estudios que cumplieron íntegramente los criterios de elegibilidad.

Figura 1:

Diagrama de flujo PRISMA 2020 del proceso de identificación y selección de los estudios incluidos — Identificación: registros identificados en PubMed (n = 174) y PubMed Central (n = 112), y registros duplicados eliminados antes del cribado (n = 41).

Esta obra está bajo una Licencia Creative Commons Atribución-No Comercial-Compartir Igual 4.0 Internacional

Identificación y selección de estudios sistemáticos en las bases de datos PubMed y PubMed Central (20152024)



Características de los estudios incluidos

La Tabla 1 presenta las características generales de los diez estudios incluidos: predominaron las revisiones sistemáticas y las revisiones clínicas con base en literatura, junto con dos estudios primarios de precisión diagnóstica (Baylor et al., 2023; Caekebeke et al., 2022), y las patologías cubiertas abarcaron la epicondilalgia lateral y medial, la rotura distal del bíceps, las lesiones del compartimento medial, las fracturas agudas y el síndrome del túnel cubital.

Esta obra está bajo una Licencia Creative Commons Atribución-No Comercial-Compartir Igual 4.0 Internacional

Tabla 1.

Características de los estudios incluidos en la revisión.

Autor y año	Tipo de estudio	Muestra	Objetivo	Patología	Tests clínicos
Zwerus et al. (2018)	Revisión sistemática	10 estudios (24 tests descritos; 14 con datos de precisión)	Revisar la evidencia sobre la precisión diagnóstica de las pruebas de exploración física del codo	Epicondilalgia lateral y medial, lesión del ligamento colateral medial, rotura distal del bíceps, túnel cubital, inestabilidad posterolateral	Cozen, Mill, Maudsley, Hook Test, BCI (*Biceps Crease Interval*), BCR (*Biceps Crease Ratio*), Moving Valgus Stress Test, Elbow Flexion Test, Chair Test, entre otros
Karanasios et al. (2022)	Revisión sistemática	24 estudios, 1.370 participantes	Evaluar la precisión diagnóstica de las pruebas clínicas y de imagen en la tendinopatía lateral del codo	Epicondilalgia lateral	Cozen, Mill, Maudsley, prensión sin dolor (painfree grip), pruebas de provocación del dolor
Breda et al. (2023)	Revisión sistemática	12 estudios, 4.485 pacientes	Analizar la precisión de las pruebas clínicas para descartar fractura de codo en traumatismo agudo	Fracturas de codo	Elbow Extension Test, Elbow full ROM test, extensión + point tenderness test

Esta obra está bajo una Licencia Creative Commons Atribución-No Comercial-Compartir Igual 4.0 Internacional

Curti et al. (2021)	Revisión sistemática con síntesis de mejor evidencia	4 estudios (2 de cohortes, 2 transversales)	Analizar la relación entre sobrecarga biomecánica ocupacional y tendinopatías del codo	Tendinopatías lateral y medial	Exploración clínica funcional y criterios diagnósticos descritos en los estudios incluidos
Kheiran et al. (2021)	Revisión clínica (narrativa)	Revisión de la literatura	Sintetizar la evidencia sobre las tendinopatías frecuentes del codo: diagnóstico y opciones terapéuticas	Epicondilalgia medial y lateral, tendinopatía distal del bíceps y del tríceps	Cozen, Mill, Maudsley, extensión resistida de muñeca, exploración funcional
Di Filippo et al. (2022)	*Overview* de revisiones sistemáticas	25 revisiones sistemáticas (227 ECA)	Analizar criterios diagnósticos, terminología y tratamiento del dolor lateral del codo	Dolor lateral del codo (epicondilalgia lateral)	Palpación dolorosa, Cozen, Mill, Maudsley, painfree grip strength, otras pruebas de provocación
Pottanat et al. (2025)	Revisión clínica (narrativa con base en evidencia)	Revisión de la literatura	Actualizar la evaluación clínica del codo en deportistas lanzadores: patologías, tests físicos y estudios complementarios	Lesión del ligamento colateral medial, epicondilalgia medial, lesión del bíceps distal, neuropatía cubital, inestabilidad	Moving Valgus Stress Test, Valgus Stress Test, Milking Maneuver, Hook Test, Tinel, Elbow Flexion Test

Esta obra está bajo una Licencia Creative Commons Atribución-No Comercial-Compartir Igual 4.0 Internacional

Caekebeke et al. (2022)	Estudio de precisión diagnóstica	Pacientes con sospecha de patología distal del bíceps comparados con controles	Determinar la utilidad diagnóstica de pruebas clínicas para roturas parciales y tendinitis del bíceps distal	Rotura parcial y tendinitis del bíceps distal	Biceps provocation test, resisted hook test, tilt sign
Baylor et al. (2023)	Estudio prospectivo de fiabilidad y validez	28 pacientes quirúrgicos con sospecha de rotura distal del bíceps	Evaluar la fiabilidad y validez del Hook Test en el diagnóstico de roturas del bíceps distal	Rotura distal del bíceps braquial	Hook Test
Adani et al. (2025)	Revisión clínica comprensiva	Revisión de la literatura	Revisar causas, estrategias diagnósticas y tratamiento del dolor medial no traumático del codo	Dolor medial del codo: lesión del ligamento colateral medial, epicondilalgia medial, túnel cubital	Moving Valgus Stress Test, Valgus Stress Test, Milking Maneuver, Tinel, Elbow Flexion Test

La Tabla 2 sintetiza los valores de sensibilidad y especificidad reportados por cada estudio incluido frente a su estándar de referencia, junto con los principales hallazgos, y sirve de base para la síntesis narrativa por patología que se desarrolla a continuación.

Tabla 2.

Precisión diagnóstica de las pruebas clínicas del codo en los estudios incluidos.

Autor y año	Estándar de referencia	Sensibilidad	Especificidad	Hallazgos principales
Zwerus et al. (2018)	RM, ecografía, radiografía, hallazgos quirúrgicos según patología	Hook Test 81–100%; BCI 88–92%; Moving Valgus Stress Test 100% (O'Driscoll et al., 2005); grip strength	Hook Test 100%; BCI 50–100%; Moving Valgus Stress Test 75%; grip strength: 80–90%	Evidencia limitada para muchas pruebas; el mejor rendimiento se obtiene al combinar maniobras con

Esta obra está bajo una Licencia Creative Commons Atribución-No Comercial-Compartir Igual 4.0 Internacional

		(descenso 5–10%): 78–83%		anamnesis y métodos complementarios
Karanasios et al. (2022)	Diagnóstico clínico, ecografía y RM	Test de Cozen 91%; diferencia de fuerza de prensión 5–10%: 78–83%	Prensión sin dolor: 80–90%	Ninguna prueba aislada con sensibilidad suficiente para descartar la tendinopatía; la combinación de pruebas mejora el rendimiento; 97% de los estudios con riesgo de sesgo incierto o alto
Breda et al. (2023)	Radiografía convencional	Hasta 98,6% (IC 95%: 95,0–99,8) en adultos	24,0% (IC 95%: 19,0–30,0) a 69,4% (IC 95%: 57,3–79,5)	El Elbow Extension Test es útil para descartar fractura cuando es negativo y puede reducir radiografías innecesarias
Curti et al. (2021)	Diagnóstico clínico, ecografía y RM según estudio	No reporta valores de precisión de pruebas	No reporta	Strain Index > 5–6,1 podría duplicar el riesgo de epicondilitis lateral; evidencia causal limitada (lateral) e insuficiente (medial); la exploración clínica sigue siendo la primera herramienta
Kheiran et al. (2021)	Exploración clínica, ecografía y RM	No reporta valores únicos	Variable según prueba y patología	Las pruebas clínicas orientan el diagnóstico, pero ninguna debe usarse como criterio único; la

Esta obra está bajo una Licencia Creative Commons Atribución-No Comercial-Compartir Igual 4.0 Internacional

				combinación de anamnesis, exploración e imagen mejora la precisión
Di Filippo et al. (2022)	Criterios reportados en las revisiones incluidas	No reporta valores de precisión	No reporta	Diez términos distintos para el dolor lateral («lateral epicondylitis», 51,6%); tests más citados: palpación 46,8%, Cozen 42,1%, Maudsley 22,2%, painfree grip 19,0%; la heterogeneidad dificulta la comparación
Pottanat et al. (2025)	RM, ecografía y evaluación quirúrgica según patología	Moving Valgus Stress Test cercana al 100% (según estudios citados)	Aproximadamente 75% para el Moving Valgus Stress Test	La exploración física es el primer paso diagnóstico en lanzadores; la combinación de pruebas mejora significativamente la precisión
Caekebeke et al. (2022)	RM y confirmación quirúrgica	Biceps provocation test 95%; resisted hook test 78%; tilt sign 58%; BPT + resisted hook en paralelo 98%	Biceps provocation test 97%; resisted hook test 76%; tilt sign 55%; combinación en paralelo 73%	En roturas parciales y tendinitis, el BPT es la prueba más precisa; la combinación en paralelo maximiza la sensibilidad (98%)
Baylor et al. (2023)	RM y hallazgos quirúrgicos	Hook Test 96% para roturas completas	Hook Test 67%	Fiabilidad interevaluador sustancial (κ = 0,71); se advierte contra el uso del Hook Test como

Esta obra está bajo una Licencia Creative Commons Atribución-No Comercial-Compartir Igual 4.0 Internacional

				única prueba diagnóstica
Adani et al. (2025)	RM, ecografía y estudios electrodiagnósticos	Moving Valgus Stress Test aproximadamente 100% (según estudios citados)	Aproximadamente 75%	El diagnóstico del dolor medial exige integrar anamnesis, exploración y estudios complementarios; ninguna prueba debe utilizarse de forma aislada

Síntesis narrativa por patología

Epicondilalgia lateral: las pruebas clásicas de provocación Cozen, Mill y Maudsley son las más empleadas en la práctica (Di Filippo et al., 2022), pero su precisión aislada es insuficiente; Karanasios et al. (2022) reportaron para el test de Cozen una sensibilidad del 91%, sin datos de especificidad robustos.

Para el test de prensión sin dolor (diferencia de fuerza del 5–10%) se reportó sensibilidad de 78–83% y especificidad de 80–90%, valores coherentes con los rangos sintetizados por Zwerus et al. (2018); ningún test aislado permitió descartar la tendinopatía con seguridad, por lo que la combinación de maniobras con la anamnesis se consolida como estrategia recomendable (Di Filippo et al., 2022; Kheiran et al., 2021).

Rotura distal del bíceps braquial: para las avulsiones completas, el Hook Test es la maniobra de mayor rendimiento, con sensibilidad y especificidad del 100% frente a la confirmación quirúrgica (O'Driscoll et al., 2007) y validación prospectiva con sensibilidad del 96%, especificidad del 67% y fiabilidad interevaluador sustancial ($\kappa = 0{,}71$) (Baylor et al., 2023).

Zwerus et al. (2018) sitúan su sensibilidad entre el 81 y el 100% con especificidad del 100% y reportan para el Biceps Crease Interval sensibilidad de 88–92% y especificidad de 50–100%; en las roturas parciales y la tendinitis distal, el rendimiento del resisted hook test desciende (sensibilidad 78%, especificidad 76%), siendo el biceps provocation test la prueba más precisa (sensibilidad 95%, especificidad 97%) y la combinación en paralelo de ambas la estrategia más sensible (98%) (Caekebeke et al., 2022).

Esta obra está bajo una Licencia Creative Commons Atribución-No Comercial-Compartir Igual 4.0 Internacional

Lesión del ligamento colateral medial y dolor medial: el Moving Valgus Stress Test mostró sensibilidad del 100% y especificidad del 75% frente a confirmación quirúrgica (O'Driscoll et al., 2005), valores reproducidos en las síntesis incluidas (Adani et al., 2025; Pottanat et al., 2025; Zwerus et al., 2018); el Valgus Stress Test convencional presenta rendimiento inferior y más variable, por lo que se recomienda complementarlo con el Milking Maneuver y la evaluación de imagen en lanzadores (Adani et al., 2025; Pottanat et al., 2025).

Fracturas agudas: el Elbow Extension Test alcanzó sensibilidad de hasta 98,6% (IC 95%: 95,0–99,8) en adultos, con especificidad variable entre 24,0% y 69,4%, lo que lo consolida como prueba de exclusión: un resultado negativo reduce sustancialmente la probabilidad de fractura y puede evitar radiografías innecesarias, sin que un resultado positivo sustituya la evaluación radiológica (Breda et al., 2023).

Síndrome del túnel cubital: la evidencia de apoyo indica que el signo de Tinel presenta sensibilidad del 70% y especificidad del 98%, mientras que la combinación de compresión directa sobre el nervio con flexión mantenida del codo durante 60 segundos eleva la sensibilidad al 98% y la especificidad al 99%, superando a cada maniobra aislada (Novak et al., 1994; sintetizado por Zwerus et al., 2018).

Discusión

La presente revisión sistemática sintetizó la evidencia disponible sobre la sensibilidad y especificidad de los tests ortopédicos del codo y permite extraer tres conclusiones de fondo: en primer lugar, existe una jerarquía clara de rendimiento diagnóstico según la patología, de modo que el Hook Test para las avulsiones completas del bíceps distal, el Moving Valgus Stress Test para las lesiones del ligamento colateral medial y el Elbow Extension Test para descartar fracturas agudas constituyen las maniobras con mayor respaldo cuantitativo.

En segundo lugar, las pruebas clásicas de provocación para la epicondilalgia lateral Cozen, Mill y Maudsley, pese a ser las más utilizadas (Di Filippo et al., 2022), no alcanzan por sí solas una precisión suficiente ni para confirmar ni para descartar el diagnóstico (Karanasios et al., 2022).

Esta obra está bajo una Licencia Creative Commons Atribución-No Comercial-Compartir Igual 4.0 Internacional

En tercer lugar, de forma transversal a todas las patologías, la evidencia coincide en que ninguna prueba debe emplearse como criterio diagnóstico único y que la combinación de maniobras con la anamnesis, la exploración física integral y los métodos de imagen incrementa la precisión y la fiabilidad del razonamiento clínico (Adani et al., 2025; Kheiran et al., 2021; Pottanat et al., 2025; Zwerus et al., 2018).

El contraste entre ambos grupos de pruebas es ilustrativo desde la perspectiva de la teoría de la precisión diagnóstica: el Hook Test, con sensibilidad y especificidad del 100% en la serie de validación original (O'Driscoll et al., 2007) y del 96% y 67%, respectivamente, en validación prospectiva independiente (Baylor et al., 2023), ofrece una razón de verosimilitud positiva alta que modifica sustancialmente la probabilidad postprueba de avulsión completa.

No obstante, la caída de especificidad observada por Baylor et al. (2023) y la advertencia expresa de sus autores contra su uso aislado recuerdan que incluso las mejores pruebas degradan su rendimiento fuera del contexto de validación, especialmente con prevalencias distintas y examinadores menos experimentados.

El Moving Valgus Stress Test (sensibilidad 100%, especificidad 75%; O'Driscoll et al., 2005) funciona, por su perfil, como excelente prueba de cribado de la insuficiencia del ligamento colateral medial, mientras que su especificidad moderada aconseja confirmación por imagen antes de decisiones quirúrgicas (Adani et al., 2025; Pottanat et al., 2025). El Elbow Extension Test, con sensibilidad de hasta 98,6% pero especificidad de apenas 24,0–69,4% (Breda et al., 2023), es el ejemplo paradigmático de prueba de exclusión, pues su valor clínico reside en el resultado negativo y no en el positivo.

Frente a ellas, las maniobras clásicas para la epicondilalgia lateral presentan un rendimiento más modesto y dependiente del contexto: el test de Cozen alcanzó una sensibilidad del 91% (Karanasios et al., 2022), pero sin especificidad robusta, mientras que el test de prensión sin dolor aporta el perfil más equilibrado (sensibilidad 78–83%, especificidad 80–90%) y emerge como candidato a integrar los protocolos de evaluación (Karanasios et al., 2022; Zwerus et al., 2018).

Esta disparidad apoya con fuerza el concepto de diagnóstico por combinación o *clusters* de pruebas: la aplicación de maniobras en paralelo maximiza la sensibilidad como demostraron

Esta obra está bajo una Licencia Creative Commons Atribución-No Comercial-Compartir Igual 4.0 Internacional

Caekebeke et al. (2022) al combinar el biceps provocation test con el resisted hook test, elevando la sensibilidad al 98%, mientras que la interpretación en serie incrementa la especificidad.

El caso del síndrome del túnel cubital es especialmente elocuente: la combinación de compresión y flexión mantenida elevó la sensibilidad del 32–75% de las pruebas aisladas al 98% con especificidad del 99% (Novak et al., 1994), principio que los estudios incluidos extienden a todo el espectro patológico del codo.

Conclusiones

La evidencia sintetizada en esta revisión sistemática confirma que las pruebas de exploración clínica constituyen una herramienta fundamental, accesible y de bajo costo para la evaluación inicial de las patologías del codo, pero ninguna prueba posee por sí sola sensibilidad y especificidad suficientes para establecer un diagnóstico definitivo, y su máxima utilidad se alcanza al integrarlas en combinaciones de maniobras junto con la anamnesis, la exploración física completa y, cuando la sospecha lo exige, los métodos de imagen.

La revisión permitió jerarquizar el rendimiento diagnóstico de las principales pruebas según la patología: el Hook Test es la maniobra de referencia para las avulsiones completas del tendón distal del bíceps (sensibilidad y especificidad de hasta el 100%, con validación prospectiva del 96% y 67%), aunque su rendimiento desciende en las roturas parciales, donde el biceps provocation test y las combinaciones en paralelo ofrecen mayor precisión.

El Moving Valgus Stress Test presenta el mejor perfil de cribado para las lesiones del ligamento colateral medial (sensibilidad 100%, especificidad 75%), y el Elbow Extension Test constituye la prueba de exclusión más útil en las fracturas agudas (sensibilidad hasta 98,6%); en la epicondilalgia lateral, las maniobras clásicas de Cozen, Mill y Maudsley mantienen utilidad orientativa con rendimiento variable, mientras que el test de prensión sin dolor aporta el equilibrio más favorable entre sensibilidad (78–83%) y especificidad (80–90%).

La heterogeneidad metodológica y terminológica identificada entre los estudios diferencias de diseño, población, criterios diagnósticos y estándares de referencia pone de manifiesto la

Esta obra está bajo una Licencia Creative Commons Atribución-No Comercial-Compartir Igual 4.0 Internacional

necesidad de investigaciones con protocolos estandarizados, muestras amplias y reporte completo de los indicadores de precisión, incluidas las razones de verosimilitud.

Desde el ámbito fisioterapéutico, la selección de cada test debe responder a la sospecha clínica, al contexto funcional y ocupacional del paciente y a la mejor evidencia disponible, fortaleciendo así un razonamiento clínico riguroso y una valoración basada en la evidencia.

Referencias bibliográficas

Adani, N., Azalia, X., Gani, K. S., Mitchel, M., & Kholinne, E. (2025). Nontraumatic medialsided elbow pain: A comprehensive review of etiologies, diagnostic strategies, and treatment approaches. *Cureus, 17*(10), Article e94701. https://doi.org/10.7759/cureus.94701

Baylor, J. L., Rae, M., Manzar, S., Pallis, M., Olsen, H. P., Akoon, A., & Grandizio, L. C. (2023). Reliability and validity of the hook test for diagnosis of distal biceps tendon ruptures. *The Journal of Hand Surgery, 48*(11), 1091–1097. https://doi.org/10.1016/j.jhsa.2023.07.004

Breda, G., De Marco, G., Cesaraccio, P., & Pillastrini, P. (2023). Diagnostic accuracy of clinical tests to rule out elbow fracture: A systematic review. *Clinics in Shoulder and Elbow, 26*(2), 182–190. https://doi.org/10.5397/cise.2022.00948

Caekebeke, P., Meglic, U., van den Bekerom, M. P. J., & van Riet, R. (2022). Evaluation of clinical tests for partial distal biceps tendon ruptures and tendinitis. *Journal of Shoulder and Elbow Surgery, 31*(3), 532–536. https://doi.org/10.1016/j.jse.2021.10.012

Curti, S., Mattioli, S., Bonfiglioli, R., Farioli, A., & Violante, F. S. (2021). Elbow tendinopathy and occupational biomechanical overload: A systematic review with bestevidence synthesis. *Journal of Occupational Health, 63*(1), Article e12186. https://doi.org/10.1002/13489585.12186

Di Filippo, L., Vincenzi, S., Pennella, D., & Maselli, F. (2022). Treatment, diagnostic criteria and variability of terminology for lateral elbow pain: Findings from an overview of systematic reviews. *Healthcare, 10*(6), Article 1095. https://doi.org/10.3390/healthcare10061095

Karanasios, S., Korakakis, V., Moutzouri, M., Drakonaki, E., Koci, K., Pantazopoulou, V., Tsepis, E., & Gioftsos, G. (2022). Diagnostic accuracy of examination tests for lateral elbow

Esta obra está bajo una Licencia Creative Commons Atribución-No Comercial-Compartir Igual 4.0 Internacional

tendinopathy (LET) – A systematic review. *Journal of Hand Therapy, 35*(4), 541–551. https://doi.org/10.1016/j.jht.2021.02.002

Kheiran, A., Pandey, A., & Pandey, R. (2021). Common tendinopathies around the elbow; what does current evidence say? *Journal of Clinical Orthopaedics and Trauma, 19*, 216–223. https://doi.org/10.1016/j.jcot.2021.05.021

McInnes, M. D. F., Moher, D., Thombs, B. D., McGrath, T. A., Bossuyt, P. M., the PRISMADTA Group, Clifford, T., Cohen, J. F., Deeks, J. J., Gatsonis, C., Hooft, L., Hunt, H. A., Hyde, C. J., Korevaar, D. A., Leeflang, M. M. G., Macaskill, P., Reitsma, J. B., Rodin, R., … Willis, B. H. (2018). Preferred reporting items for a systematic review and metaanalysis of diagnostic test accuracy studies: The PRISMADTA statement. *JAMA, 319*(4), 388–396. https://doi.org/10.1001/jama.2017.19163

Novak, C. B., Lee, G. W., Mackinnon, S. E., & Lay, L. (1994). Provocative testing for cubital tunnel syndrome. *The Journal of Hand Surgery, 19*(5), 817–820. https://doi.org/10.1016/03635023(94)901937

O'Driscoll, S. W., Goncalves, L. B. J., & Dietz, P. (2007). The hook test for distal biceps tendon avulsion. *The American Journal of Sports Medicine, 35*(11), 1865–1869. https://doi.org/10.1177/0363546507305016

O'Driscoll, S. W., Lawton, R. L., & Smith, A. M. (2005). The "moving valgus stress test" for medial collateral ligament tears of the elbow. *The American Journal of Sports Medicine, 33*(2), 231–239. https://doi.org/10.1177/0363546504267804

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., MayoWilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, Article n71. https://doi.org/10.1136/bmj.n71

Pottanat, P. J., Ashy, C., Reis, R. J., Bailey, E. P., Friedman, R. J., Rogalski, B. L., & Eichinger, J. K. (2025). Evaluation of the elbow in throwing athletes. *Video Journal of Sports Medicine, 5*(2). https://doi.org/10.1177/26350254241291590

Esta obra está bajo una Licencia Creative Commons Atribución-No Comercial-Compartir Igual 4.0 Internacional

Shiri, R., ViikariJuntura, E., Varonen, H., & Heliövaara, M. (2006). Prevalence and determinants of lateral and medial epicondylitis: A population study. *American Journal of Epidemiology, 164*(11), 1065–1074. https://doi.org/10.1093/aje/kwj325

Whiting, P. F., Rutjes, A. W. S., Westwood, M. E., Mallett, S., Deeks, J. J., Reitsma, J. B., Leeflang, M. M. G., Sterne, J. A. C., Bossuyt, P. M. M., & the QUADAS2 Group. (2011). QUADAS2: A revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine, 155*(8), 529–536. https://doi.org/10.7326/0003481915582011101800009

Zwerus, E. L., Somford, M. P., Maissan, F., Heisen, J., Eygendaal, D., & van den Bekerom, M. P. J. (2018). Physical examination of the elbow, what is the evidence? A systematic literature review. *British Journal of Sports Medicine, 52*(19), 1253–1260. https://doi.org/10.1136/bjsports2016096712

Conflicto de intereses:

Los autores afirman que no existe ningún conflicto de interés relacionados con la elaboración y publicación de este artículo, garantizando la independencia y objetividad de los hallazgos reportados en la presente revisión sistemática

Financiamiento:

Los autores declaran que no recibieron financiamiento ajeno para la realización de esta investigación

Agradecimiento:

Los autores expresan su agradecimiento a la Universidad Técnica Particular de Loja por el soporte académico y los recursos institucionales proporcionados.

Nota:

El artículo no es el resultado de una publicación previa

Esta obra está bajo una Licencia Creative Commons Atribución-No Comercial-Compartir Igual 4.0 Internacional