



Universidad de Valladolid

FACULTAD DE CIENCIAS DE LA SALUD DE SORIA

GRADO EN FISIOTERAPIA

TRABAJO FIN DE GRADO

Propiedades psicométricas de escalas de evaluación observacional de la marcha en pacientes con parálisis cerebral. Una revisión sistemática.

Presentado por: Alba Ferré Fernández

Tutores: Ricardo Medrano de la Fuente y Silvia Lahuerta Martín

Soria, a 13 de junio de 2024

Terminar este TFG se me hizo más duro de lo que pensaba, pero estoy agradecida de que mi madre y mi hermano no dejaran de apoyarme en ningún momento. Esto va por ti papá, y por vosotros abuelos, porque espero que me sigáis viendo desde arriba y estéis orgullosos.

RESUMEN

Introducción: La evaluación de la marcha en pacientes en edad pediátrica con parálisis cerebral (PC) es crucial para diseñar un tratamiento efectivo. Sin embargo, la fiabilidad y validez de las escalas observacionales utilizadas en esta población requieren una evaluación para garantizar que sean adecuadas.

Objetivos: Se planteó como objetivo principal analizar la fiabilidad y validez de las escalas observacionales de la marcha en pacientes con PC en edad pediátrica.

Metodología: Se siguió la metodología Cosmin y se realizó una revisión sistemática aplicando criterios de inclusión y exclusión para seleccionar los estudios adecuados. Se incluyeron estudios que analizaran la fiabilidad y validez de escalas observacionales de la marcha. Se evaluó la calidad metodológica de los estudios incluidos.

Resultados: Se incluyeron 12 estudios con una calidad metodológica entre adecuada y muy buena. Los resultados de fiabilidad variaron mostrando coeficientes de correlación intraclase superiores a 0,90 e inferiores a 0,70. Respecto a la validez de criterio se observaron valores kappa de 0,21 a 0,71 y porcentajes de acuerdo del 49,5% al 80%. Por último, el único estudio que evaluó la validez de constructo obtuvo correlaciones de 0,69 a 0,73.

Conclusiones: En la fiabilidad y validez de las escalas observacionales se observaron discrepancias en los resultados. Se destaca la necesidad de aumentar la calidad de los procedimientos de evaluación y los métodos estadísticos para mejorar la evaluación de la marcha en pacientes con PC en edad pediátrica; y, así, proporcionar una base sólida en fisioterapia con escalas fiables y validadas.

ÍNDICE

1.	INTRODUCCIÓN	5
1.1	Clasificación	5
1.2	Clínica.....	5
1.3	Valoración de la marcha.....	7
1.4	Fiabilidad y validez	7
2.	JUSTIFICACIÓN.....	7
3.	OBJETIVOS	8
4.	MATERIAL Y MÉTODOS.....	8
4.1	Estrategia de búsqueda.....	8
4.2	Criterios de inclusión y exclusión	8
4.3	Selección de estudios.....	9
4.4	Evaluación de la calidad metodológica de los estudios	9
4.5	Evaluación de los criterios para buenas propiedades psicométricas.....	9
5.	RESULTADOS	10
5.1	Resultados de las búsquedas	10
5.2	Características sociodemográficas y clínicas de los estudios	10
5.3	Calidad metodológica de los estudios	10
5.4	Análisis de los resultados	11
5.4.1	Fiabilidad	11
5.4.1.1	Fiabilidad intra e interobservadores	11
5.4.1.2	Error de medición.....	11
5.4.2	Validez.....	11
5.5	Evaluación de los criterios para buenas propiedades psicométricas	11
6.	DISCUSIÓN.....	12
7.	CONCLUSIONES	14
	BIBLIOGRAFÍA	15
	ANEXOS	18

LISTADO DE ABREVIATURAS

CCI: coeficiente de correlación intraclase

CMD: cambio más pequeño (mínimo) detectable

CV: coeficiente de variación

EEM: error estándar de medición

EME: escala de la marcha de Edimburgo

EOM: escala de observación de la marcha

ES: escala de Salford

EVE: escala visual de Edimburgo

EVEI: escala visual de Edimburgo por intervalos

EVM: escala visual de la marcha

LdA: límites de acuerdo

PC: Parálisis cerebral

SCFMG: Sistema de clasificación de funciones motoras gruesas

TFG: trabajo fin de grado

1. INTRODUCCIÓN

La parálisis cerebral (PC) se define como un grupo de trastornos que impactan en el desarrollo, el control del movimiento y la postura, que puede dar como resultado limitaciones significativas en la capacidad de llevar a cabo actividades cotidianas. Se trata de una lesión no progresiva que puede ocurrir durante el período fetal, en el momento del parto, o durante los dos primeros años de vida (1). Se estima una prevalencia constante durante la última década que oscila entre 2-2,5/1000 niños nacidos vivos (2).

Existen factores de riesgo asociados al desarrollo de PC: factores prenatales, factores perinatales y factores postnatales. Los factores prenatales comprenden aspectos relacionados con la madre, alteraciones de la placenta y condiciones del feto. Los perinatales están asociados con el periodo que va desde el final del embarazo hasta los primeros días de vida del recién nacido. Por otro lado, los factores postnatales incluyen influencias que impactan al niño después del nacimiento. Identificar estos factores se considera esencial para implementar medidas preventivas y proporcionar intervenciones que contribuyan al bienestar y desarrollo saludable de los pacientes (3).

1.1 Clasificación

Los pacientes con PC pueden clasificarse en base al trastorno motor predominante. De esta manera, los pacientes pueden presentar PC espástica, PC discinética, PC atáxica y formas mixtas. La forma espástica se caracteriza por rigidez muscular, aumento de los reflejos y limitación del movimiento voluntario, lo que puede provocar contracturas y deformidades musculares. La forma discinética se manifiesta con movimientos involuntarios, cambios repentinos en el tono muscular y reflejos persistentes. La forma atáxica se caracteriza por falta de coordinación motora, temblores intencionales y problemas de equilibrio. Por último, los pacientes con PC de formas mixtas pueden presentar características de varios tipos mencionados anteriormente (4,5). Por otro lado, el sistema de clasificación de funciones motoras gruesas (SCFMG) en inglés "*Gross Motor Function Classification System*" se considera el *gold standard* para conocer el nivel de funcionalidad del paciente (6). El SCFMG fue diseñado con el propósito de proporcionar un enfoque sencillo para categorizar a niños con PC menores de 12 años en función de sus capacidades y limitaciones funcionales. Se compone de cinco niveles y se divide en cuatro grupos de edad. Las distinciones entre estos niveles reflejan diferencias en habilidades motoras gruesas que se consideran significativas en la vida cotidiana de los niños que padecen PC. En el nivel I del SCFMG se encuentran aquellos niños que caminan sin limitaciones, seguidos por el nivel II que comprende aquellos que caminan con ciertas limitaciones. El nivel III incluye a los niños que caminan con la ayuda de un dispositivo manual, mientras que el nivel IV abarca a aquellos con movilidad limitada que pueden utilizar dispositivos motorizados. El nivel V engloba a los niños transportados en silla de rueda. La clasificación por edad considera grupos de edades inferiores a los 2 años, entre los 2 y 4 años, entre los 4 y 6 años, entre los 6 y 12 años y un grupo adicional entre los 12 y 18 años introducido por Palisano y colaboradores en 2008 (7).

1.2 Clínica

Los pacientes con PC se caracterizan en gran medida por el retraso en el desarrollo motor, lo que puede ocasionar estrategias motoras inmaduras. Este retraso no solo afecta la adquisición de nuevas habilidades motoras, sino que también se puede manifestar con una

persistencia de reflejos primitivos. Así mismo, los pacientes con PC se pueden caracterizar por la presencia de espasticidad y distonía, además de otras alteraciones en el control motor, debilidad y alteraciones sensoriales y del equilibrio. La espasticidad se caracteriza por un aumento anormal en la respuesta muscular al estiramiento que puede manifestarse como contracciones musculares exageradas, espasmos o incluso episodios de contracciones muscular rítmicas y repetitivas llamadas clonus; mientras que la distonía corresponde con contracciones sostenidas o intermitentes que forman movimientos anormales y repetitivos (3,4). Las alteraciones en el control motor pueden afectar a la capacidad del cuerpo para coordinar y ejecutar movimientos voluntarios de manera eficaz. La debilidad muscular puede influir en la capacidad para realizar actividades motoras, mientras que las alteraciones sensoriales y del equilibrio pueden dificultar la percepción y la estabilidad propia durante el movimiento. A consecuencia de todo ello se producen deformidades musculoesqueléticas que pueden ocasionar alteraciones de la marcha. Las deformidades más comunes son respecto el pie y desplazamiento de la cadera, hecho que puede influir en un apoyo del pie anormal y una afectación de la función de la marcha (8).

Las alteraciones de la marcha en pacientes con PC pueden variar en función del tipo de PC. En relación a los pacientes con PC espástica se pueden agrupar en patrones de marcha espástica unilateral y espástica bilateral. La marcha espástica unilateral se subdivide en cuatro patrones de marcha: el tipo 1 caracterizado por un pie caído, tipo 2 que es el más común se observa equino con una rodilla en extensión o recurvatum y la cadera extendida. El tercer tipo presenta una alteración de la dorsiflexión de tobillo con una rodilla rígida y flexionada, y por último el cuarto tipo tiene un patrón similar al de la diplejía, sin embargo, dado que la afectación es unilateral existe un pie equino con rodilla flexionada y rígida, cadera flexionada e inclinación pélvica anterior. La marcha espástica bilateral también se subdivide en cuatro patrones: el denominado verdadero equino caracterizado por el tobillo en flexión plantar que puede ser difícil de identificar si se desarrolla el recurvatum en la rodilla; la marcha de salto o segunda forma, con el tobillo en equino, rodilla y cadera en flexión e inclinación pélvica anterior con lordosis lumbar acentuada. En tercer lugar, el equino aparente donde el equino disminuye al aumentar la flexión de cadera y rodilla y por último la marcha en cuclillas que se caracteriza por una flexión excesiva tanto del tobillo como de la rodilla y de la cadera (9).

La PC en su variedad discinética se caracteriza por patrones de movimiento extrapiramidales, los cuales causan anomalías en el tono, alteraciones en el control postural y déficits en la coordinación. Estos movimientos discinéticos pueden incluir atetosis, corea, coreoatetosis y distonía. La atetosis se manifiesta mediante movimientos lentos y ondulantes, que puede interferir con la estabilidad durante la marcha. La corea puede causar movimientos bruscos e irregulares que dificultan el mantenimiento de un patrón de marcha regular. Por otro lado, la distonía con posturas anormales puede afectar a la alineación corporal y el equilibrio durante la marcha (10).

La PC atáxica se identifica por una marcha con problemas de equilibrio y movimientos descoordinados, lo que dificulta mantener una marcha estable. Los pacientes tienden a caminar de forma inestable, con un aumento de la base de sustentación. Estos desafíos contribuyen a una marcha insegura que puede afectar a la capacidad para realizar actividades de la vida diaria (10).

1.3 Valoración de la marcha

Dado que la marcha representa un desafío para pacientes con PC, la observación y análisis de las posibles alteraciones es fundamental para plantear un adecuado tratamiento para conseguir la mejora de su movilidad y funcionalidad (11). El análisis observacional de la marcha se centra en comprender la gravedad y extensión de los déficits funcionales, lo que ayuda a determinar qué tipo de intervención puede ser más adecuada. Además de la evaluación inicial, realizar un seguimiento del patrón de marcha después del tratamiento permite evaluar el progreso y los resultados de manera objetiva (12).

La evaluación de la marcha se puede dividir en tres categorías: métodos cualitativos, métodos cuantitativos y análisis clínico de la marcha. Los métodos cualitativos engloban tanto la observación como el empleo de cuestionarios de manera directa y subjetiva, mientras que los métodos cuantitativos se basan en la comparación de datos numéricos de la marcha para identificar áreas de normalidad. En el caso del análisis clínico de la marcha, se busca identificar y comprender las anomalías en la marcha, generalmente mediante test funcionales. El análisis observacional de la marcha implica conocer la biomecánica de la marcha y el observador genera un informe con una puntuación basándose en parámetros como longitud del paso y ángulo de las articulaciones (13,14).

1.4 Fiabilidad y validez

Las escalas de evaluación observacional de la marcha son instrumentos de medida que presentan diversas propiedades psicométricas inherentes, como la fiabilidad y la validez (15). La fiabilidad se refiere a la estabilidad de las medidas obtenidas mediante una herramienta de evaluación. Dentro de la fiabilidad se pueden encontrar diferentes tipos: fiabilidad intraevaluador y fiabilidad entre evaluadores, y error de medición. La fiabilidad intraevaluador se obtiene de las mediciones de los mismos investigadores evaluadas en diferentes situaciones, en cambio la fiabilidad entre evaluador se analiza entre distintos investigadores, pero en la misma situación. El error de medición, dentro de la fiabilidad, es el error sistemático y aleatorio en la puntuación observada (16).

Por otro lado, la validez se refiere a la medida en que la herramienta de valoración realmente mide lo que se propone medir, por tanto, evalúa si la herramienta consigue medir la variable o constructo específico que se pretende evaluar. La validez es fundamental para garantizar que las conclusiones derivadas de su uso sean precisas y confiables y así asegurar la efectividad de las intervenciones terapéuticas basadas en la evaluación. Se distinguen tipos de validez: la validez de constructo, la validez de criterio y la validez de contenido. La validez de constructo es el grado de consistencia de la hipótesis con los resultados mientras que la validez de criterio es el grado de acuerdo de las puntuaciones del instrumento a medir con un *gold standard*. A su vez, la validez de contenido es la claridad de los ítems, tareas, observaciones o parámetros de un instrumento de medición y la manera que concuerda con el constructo que se va a medir. Utilizar métodos con una fiabilidad y validez adecuada es una condición esencial en la práctica clínica de fisioterapia (15-17).

2. JUSTIFICACIÓN

La utilización de escalas observacionales para la valoración de la marcha en pacientes en edad pediátrica con PC es hoy en día una práctica común para los fisioterapeutas. Por tanto, es

fundamental asegurar la fiabilidad y validez de las escalas y así garantizar su utilidad clínica y científica. Aunque existen varias escalas observacionales de la marcha utilizadas en la práctica clínica, su fiabilidad y validez en pacientes pediátricos con PC aún no está ampliamente investigada.

Zanudin y colaboradores (18) realizaron una revisión sistemática en el año 2017 enfocada a analizar las propiedades psicométricas de diversos instrumentos de medida utilizadas para evaluar la calidad de la marcha y el rendimiento al caminar de jóvenes con PC, dejando fuera del análisis a niños con edades inferiores. Por tanto, se necesitan estudios que analicen las propiedades psicométricas de instrumentos de medida englobando toda la edad pediátrica con PC.

Por todo ello, se considera relevante la realización del presente Trabajo Fin de Grado (TFG) que pretende analizar la fiabilidad y validez de las escalas observacionales de valoración de la marcha en pacientes en edad pediátrica con PC y así contribuir a una mejora en las prácticas de fisioterapia.

3. OBJETIVOS

El objetivo principal de esta revisión sistemática es analizar la fiabilidad y validez de las escalas observacionales de la marcha utilizadas en la valoración de pacientes en edad pediátrica con PC.

Como objetivos específicos se plantean:

- Identificar escalas fiables y validas, así como conocer sus limitaciones.
- Analizar la calidad metodológica de los estudios incluidos en la revisión sistemática.
- Evaluar la fiabilidad de las escalas observacionales de la marcha en pacientes en edad pediátrica con PC.
- Analizar la validez de las escalas observacionales de la marcha en población pediátrica con PC.

4. MATERIAL Y MÉTODOS

4.1 Estrategia de búsqueda

Durante los meses de septiembre y octubre de 2023 se llevó a cabo una búsqueda bibliográfica en diversas bases de datos especializadas. Las bases de datos seleccionadas fueron *Pubmed (Medline)*, *Cochrane*, *Cinahl* y *Web of Science*. Las búsquedas se realizaron usando una combinación de los siguientes *Medical Subjects Heading (MSH)*: "*cerebral palsy*", "*gait*", "*gait analysis*", "*gait disorders, neurologic*", "*reproducibility of results*", "*validity*" y "*reliability*", y palabras claves: "*outcomes measurement*", "*observational methods*", y "*assessment*", unidos con los operadores booleanos *AND* y *OR*; sin límite de fecha de publicación. En el Anexo 1 se detalla la estrategia de búsqueda realizada.

4.2 Criterios de inclusión y exclusión

Los estudios que se consideraron aptos para ser incluidos son aquellos que cumplían la pregunta de investigación según la metodología Cosmin (16):

- Población: pacientes de 2 a 18 años con PC.

- Constructo: valoración observacional de la marcha en pacientes en edad pediátrica con PC.
- Instrumento: escalas de evaluación observacional de la marcha.
- Propiedades psicométricas: fiabilidad y validez.

Así mismo, se fijaron los siguientes criterios de exclusión:

- Revisiones sistemáticas.
- Estudios que no incluían pacientes pediátricos.
- Estudios que analizaban otros instrumentos de medición de la marcha que no fueran de observación.

4.3 Selección de estudios

Tras llevar a cabo la búsqueda, se procedió a realizar una selección de estudios mediante varias etapas de filtrado. En primer lugar, fueron eliminados los estudios duplicados. Posteriormente, se aplicó un cribado según el título, el resumen y las palabras clave con el fin de identificar los que estuviesen relacionados con el tema. Finalmente, se realizó una lectura completa para verificar si cumplían los criterios de inclusión establecidos. Ante cualquier duda durante este proceso, se consultó con los tutores de TFG.

4.4 Evaluación de la calidad metodológica de los estudios

En el análisis de la calidad metodológica de los estudios, se evaluó el riesgo de sesgo a través de la lista de verificación de Cosmin (19,20). Esta lista se compone de nueve apartados para cada una de las propiedades psicométricas con varios ítems sobre estándares metodológicos del diseño y enfoque estadístico. La calidad metodológica general se basó en la puntuación más baja obtenida respecto a fiabilidad, error de medición, validez de criterio y validez de constructo en cualquiera de los ítems (19,20).

Para la fiabilidad se tuvieron en cuenta la estabilidad temporal, el intervalo de tiempo entre mediciones, la similitud en las condiciones de medición, el cegamiento del profesional y en la asignación de puntuaciones, y por último los métodos estadísticos utilizados como el coeficiente de correlación intraclase (CCI) y el valor kappa. Por otro lado, en el análisis del error de medición se aplicaron los mismos criterios que la fiabilidad excepto los métodos estadísticos en los cuales se tiene en cuenta el error estándar de medición (EEM), el cambio mínimo detectable (CMD), los límites de acuerdo (LdA), el coeficiente de variación (CV) y el porcentaje de acuerdo. En cuanto a la validez de criterio se analizó las correlaciones o área bajo la curva para puntuaciones continuas, la sensibilidad y especificidad para puntuaciones dicotómicas, y si hay errores en el diseño o en los métodos estadísticos. Por último, para la validez de constructo se observó la claridad en la medida y la suficiencia de las propiedades de medición del instrumento comparador, así como las posibles debilidades en el diseño o en los métodos estadísticos (19,20).

4.5 Evaluación de los criterios para buenas propiedades psicométricas

Se analizaron los criterios para buenas propiedades psicométricas. Para ello, se observó la fiabilidad intra e interobservador y el error de medición; la validez de criterio y la validez de constructo. Los resultados se dividieron obteniendo una puntuación de “positivo” (+), “dudoso o indeterminado” (?) y “negativo” (-) (21) (Tabla 1). Para la validez de criterio se tuvieron en cuenta la adecuación de las mediciones en relación a un *gold standard*. Por otro lado, la validez

de constructo se evaluó mediante la concordancia de los resultados con la hipótesis establecida. Por último, el análisis de la fiabilidad, se realizó teniendo en cuenta el CCI o el valor kappa ponderada como indicativos de ser adecuado y el error de medición a raíz del EEM, el CMD, los LdA, el CV o el porcentaje de acuerdo específico (19-21).

5. RESULTADOS

5.1 Resultados de las búsquedas

Tras la realización de las búsquedas se identificaron un total de 798 estudios, de los cuales 184 estaban duplicados y fueron eliminados. Por lectura de título y resumen fueron eliminados 598 estudios, dejando un total de 16 estudios. Posteriormente 4 fueron descartados teniendo en cuenta los criterios de inclusión y exclusión por lo que finalmente 12 estudios fueron incluidos en esta revisión. En la Figura 1 se muestra el diagrama de flujo del proceso de selección de estudios.

Todos los estudios analizaban la fiabilidad, ya sea fiabilidad intraexaminador (22-24,26-29,33), fiabilidad entre examinadores (22-30) o error de medición (22,24-26,30,31,33). Además, seis estudios analizaban la validez de criterio (23,25-27,29,32) y un estudio la validez de constructo (22).

5.2 Características sociodemográficas y clínicas de los estudios

La población de los estudios incluidos mostró características heterogéneas como se observa en la Tabla 2. El tamaño de la muestra varió desde 4 pacientes (25) hasta 62 (22). En cuanto al rango de edad de la población pediátrica que participaba en los estudios era de 4 a 16 años (23-27,29-33) excepto un artículo que el rango de edad es desde los 4 hasta los 18 años (22) y un artículo que incluye desde los 3 años hasta los 10 años (28). Ocho de los estudios especifican cuantos hombres y mujeres forman parte de la muestra (22-24,27-30,32). Todos los estudios tenían una mayoría de hombres que de mujeres.

Todos los pacientes fueron diagnosticados de PC de tipo espástica. Cuatro de los estudios incluyeron pacientes con hemiplejía o diplejía (22,23,28,30), otro estudio incluía, además, monoplejía (24). Por otro lado, dos estudios incluyeron pacientes con cuadriplejía además de los dichos ya anteriormente (30,31). El resto de los estudios incluyeron pacientes con hemiplejía (25,27), otros dos estudios pacientes con diplejía (29,33) y uno de ellos no especificaba el tipo de PC (26). Los estudios utilizaron el SCFMG para clasificar a los pacientes en niveles (22-24,26,27), mientras que el estudio de Brown y colaboradores (25) usó el índice de Gillette.

Tres estudios utilizaron la escala visual de Edimburgo (EVE) (22,24,26) y otros tres la escala de Salford (ES) (30-32). La escala de observación de la marcha (EOM) fue utilizada por dos estudios (23,29) así como la escala visual de la marcha (EVM) (25,27). Solo un estudio utilizó la escala visual de Edimburgo por intervalos (EVEI) (28) y otro la escala de la marcha de Edimburgo (EME) (33). Todas ellas fueron analizadas mediante la utilización de vídeos de los pacientes (22-33).

5.3 Calidad metodológica de los estudios

La calidad metodológica siguiendo la lista de verificación de riesgo de sesgo de Cosmin (19,20) se dividió según la propiedad psicométrica analizada: fiabilidad, error de medición, validez de criterio y validez de constructo. En la Tabla 3 se muestran los resultados respecto a la calidad metodológica en relación a la fiabilidad donde cinco obtuvieron una calidad muy buena

(22,25,29,30,33) y otros cinco una calidad adecuada (23,24,26-28). La calidad metodológica respecto al error de medición obtuvo resultados muy buenos para 5 estudios (22,25,30,31,33) y adecuados en dos estudios (24,26) que se pueden observar en la Tabla 4. La evaluación de la calidad metodológica en referencia a la validez de criterio, plasmada en la Tabla 5, fue inadecuada en todos los estudios en los que se midió (23,25-27,29,32). Para la validez de constructo la calidad fue muy buena en el estudio en el que se analizó (22) (Tabla 6).

5.4 Análisis de los resultados

5.4.1 Fiabilidad

5.4.1.1 Fiabilidad intra e interobservadores

Los resultados de fiabilidad se dividieron tanto la fiabilidad intra e interobservadores, así como el error de medición de los instrumentos (Tabla 7). El estudio de Abe y colaboradores (22) que analiza la EVE mostró una fiabilidad con CCI superiores a 0,90 pero sin embargo otros estudios con esta misma escala se observaron resultados inferiores a 0,70 (24,26). En relación a la EOM, los resultados para la fiabilidad intraobservadores variaron de 0,61 a 0,74 en los estudios que se analizó (23,29). La fiabilidad entre observadores en ambos estudios que analizaban la EOM fue menor a 0,70 (23,29). Los estudios que analizaban la EVM observaron resultados inferiores a 0,70 (26) o un coeficiente de repetibilidad de 5,56 para observadores expertos y 5,93 para observadores inexpertos (25). Con respecto a la ES los resultados obtenidos en la fiabilidad entre observadores para fisioterapeutas pediátricos y no pediátricos fueron superiores a 0,70. Sin embargo, para estudiantes de fisioterapia fueron menores (30). El único estudio que analizó la EME obtuvo un valor de CCI de 0,59 (33).

5.4.1.2 Error de medición

En relación al error de medición, solo dos estudios calcularon el EEM junto con el CMD (22,30). El resto de estudios representaron sus resultados en porcentajes de acuerdo (24-26,31,33) (Tabla 7). Solo en el estudio de Aroojis y colaboradores (24), el porcentaje de acuerdo fue superior al 80% en uno de los dos observadores en el error de medición intraobservador.

5.4.2 Validez

Respecto a la validez, 6 estudios (23,25-27,29,32) analizan la validez de criterio y solo uno de ellos la validez de constructo (22) (Tabla 8). Tres de los estudios dieron un valor de kappa ponderada de 0,40 (23), para un observador 0,21 mientras que otro observador 0,22 (27) y el último estudio para un observador 0,52 y el otro observador 0,71 (29). En los estudios en los que se analizó a través del porcentaje de acuerdo los resultados variaron desde 57,5% para observadores expertos y 49,5 observadores inexpertos (25), un 61,06% (26) y un 58% para categorías de calificación mientras que 80% en estimaciones subjetivas (32). El único estudio que analizó la validez de constructo arrojó resultados para un primer evaluador $r=0,73$ y otros dos observadores de $r=0,71$ y $r=0,69$ (22). Estos resultados fueron estadísticamente significativos con una $p<0,001$ indicando una fuerte correlación entre la medida de función motora gruesa – 66 y la EVE.

5.5 Evaluación de los criterios para buenas propiedades psicométricas

Los resultados de fiabilidad, error de medición y validez demostraron una variabilidad entre estudios, reflejando una calidad de las propiedades psicométricas heterogénea, con

calificaciones que oscilaron entre positivas y negativas como se observa en la Tabla 9. En relación a la fiabilidad, algunos estudios exhibieron una concordancia alta y positiva (22), mientras que otros mostraron resultados menos consistentes (23,24,29). Respecto al error de medición, ciertos estudios no proporcionaron mediciones específicas, generando dudas sobre su calidad (22,30). En otros casos los resultados fueron negativos, ya que el porcentaje de acuerdo específico no fue mayor al 80% (24,26,31,33). Para la validez de criterio y constructo la calidad de ambas propiedades fue dudosa (22,23,25-27,29,32), ya que no proporcionaron información suficiente o no fueron bien descritos los métodos estadísticos utilizados.

6. DISCUSIÓN

El objetivo principal de este TFG fue analizar la fiabilidad y validez de las escalas observacionales de la marcha en pacientes en edad pediátrica con PC. Para lograrlo, se plantearon objetivos específicos que incluyeron identificar escalas fiables y válidas, así como evaluar la calidad metodológica de los estudios incluidos.

Con respecto a la muestra de los estudios incluidos, conviene destacar que según Cosmin (34) se acepta como una calidad metodológica muy buena un tamaño de la muestra de al menos 100 pacientes y se considera adecuado si se incluyen entre 50-99 pacientes, dudoso si incluyen entre 30-49 paciente e inadecuado menos de 30 pacientes. La mayoría de los estudios incluidos en esta revisión utilizaron una muestra menor de 30 pacientes (23,25,28-33), el estudio con mayor muestra es de 62 pacientes (22) y los restantes tiene una muestra de 30 pacientes (24,26) y de 31 (27). Esto coincide con revisiones anteriores (18) indicando que los estudios que analizan las propiedades psicométricas por lo general tienen menos de 30 participantes. El tamaño de la muestra es un factor importante a valorar a la hora de realizar investigaciones, por lo que para mejorar la calidad metodológica se recomienda utilizar una muestra más grande. Sin embargo, el manual de Cosmin (20) para revisiones sistemáticas sugiere que no es necesario exigir un tamaño de muestro específico en estudios individuales cuando se pueden combinar los resultados de varios estudios pequeños porque juntos pueden proporcionar suficiente evidencia. Sin embargo, en los estudios cuyo objetivo sea valorar la validez dado que sus resultados no se pueden combinar, sí que requieren un tamaño muestral superior para alcanzar resultados más sólidos (20).

Los resultados de los estudios incluidos (22-33) presentan una muestra heterogénea en términos de edad. Esto refleja la diversidad de la población pediátrica afectada por PC. Dado el amplio rango de edades, es importante tener en cuenta que el desarrollo motor y las habilidades motrices pueden diferir a lo largo de la infancia (10). Por lo tanto, para futuras investigaciones sería beneficioso dividir la muestra en grupos de edad más específicos para abordar de manera más precisa la evaluación de la marcha en cada etapa del desarrollo en edad pediátrica. La disponibilidad de datos relacionados con la validez y la fiabilidad divididos en grupos de edad permitiría una mejora en la práctica clínica ya que sería posible trabajar de manera específica en función de la edad de los pacientes en una etapa donde se experimentan tantos cambios relevantes relacionados con el control motor y la movilidad (10).

Todos los estudios se centraron en pacientes diagnosticados en PC espástica, principalmente hemiplejía y diplejía (22-33). Sin embargo, un estudio no especifica el tipo de PC espástica (26) por lo que dificulta comparar los resultados entre estudios y aplicar estos resultados a otras poblaciones o situaciones. La homogeneidad en las características de la población nos ayuda a

garantizar la comparabilidad de los resultados entre estudios y su aplicabilidad en diferentes contextos.

La calidad metodológica en cuanto a la fiabilidad varió significativamente entre los estudios incluidos (22-31,33). Algunos estudios no delimitaron adecuadamente las características tanto durante la observación como en la toma de puntuaciones (23,28). Es fundamental realizar la toma de puntuaciones de manera ciega a otros observadores y repetir los procesos en el tiempo con un margen de tiempo adecuado para garantizar resultados consistentes. Se recomienda un intervalo de una a dos semanas entre las mediciones repetidas, pero es importante considerar que puede variar dependiendo de la estabilidad de los participantes (19,20). Además, hay estudios que no describieron adecuadamente los métodos estadísticos o no especificaron los resultados obtenidos. Los estudios que no definieron cómo administraron las medidas o asignaron las puntuaciones como aquellos con métodos estadísticos poco claros obtuvieron una calidad adecuada (24,26-28) mientras que el resto obtuvieron una calidad muy buena (22,25,29-31,33). Es crucial en el análisis de la fiabilidad que los observadores no estén al tanto de las puntuaciones previas ni de la forma en que otros observadores pasaron cada escala en los mismos pacientes tanto en situaciones iguales como diferentes. Así mismo, también es importante que las medidas estadísticas de los resultados estén bien descritas (19,20).

En relación al error de medición, los estudios con una calidad metodológica adecuada (24,26) podrían mejorar si los métodos estadísticos calcularan el porcentaje de acuerdo específico ya que valora cada categoría individualmente en lugar de limitarse al porcentaje de acuerdo general (19,20). Este enfoque permitiría una calidad metodológica respecto al error de medición más alta.

A la hora de evaluar el riesgo de sesgo para la validez de criterio todos los resultados fueron inadecuados, según la metodología Cosmin únicamente se valoran puntuaciones continuas y dicotómicas. Sin embargo, todos los estudios arrojaron los resultados mediante la puntuación de kappa ponderada (23,25-27,29,32), por lo que para puntuaciones ordinales podría considerarse esta adecuada y así obtener una calidad muy buena. Estos resultados concuerdan con los observados en la revisión de Zanudin y colaboradores (18), donde los estudios que obtuvieron una calidad adecuada estuvieron influenciados por un análisis estadístico inadecuado.

Cuando se aplicaron los criterios para buenas propiedades psicométricas la fiabilidad varió desde concordancia alta y positiva (22), hasta resultados menos consistentes, encontrando diferencias tanto en la fiabilidad intra e interobservadores (23,24,29). En términos generales en diversos estudios los resultados fueron diferentes tanto para la fiabilidad intraobservadores y la fiabilidad entre observadores (23,24,29). Hubo diferencias en la fiabilidad intraobservadores de diferentes observadores (29) y en la fiabilidad entre observadores (30). Esto puede deberse a la experiencia de los observadores ya que influye en la capacidad para realizar evaluaciones. También la manera de aplicar los criterios puede variar según los observadores y la manera de interpretarlos pudieron llevarlos a resultados diferentes. Se observaron discrepancias en los resultados según el tipo de observador, como fisioterapeutas pediátricos, no pediátricos o estudiantes de fisioterapia (30). Es posible que los estudiantes de fisioterapia, al no haberse incorporado aún al mercado laboral y haber tenido menos oportunidades de experiencia en la práctica clínica, mostraran niveles de fiabilidad inferiores en la evaluación de resultados en comparación con

profesionales experimentados. Respecto al error de medición un porcentaje de acuerdo bajo indica una falta de consistencia y precisión en las mediciones entre observadores o en diferentes momentos. Esto puede variar dependiendo de cada observador en la manera de interpretar o incluso errores a la hora de medir además de la experiencia. Para la validez de criterio y constructo los resultados son mixtos, un único estudio indica una alta validez (22) mientras que en los demás baja (23,24-27,29,32); y en cuanto a los criterios para buenas propiedades en relación a la validez fue dudosa (22,23,25-27,29,32) porque los estudios no proporcionan suficiente información clara sobre cómo realizaron los análisis además de no describir los métodos estadísticos utilizados.

Como se ha detallado, en los estudios incluidos en la presente revisión se han observado limitaciones en el tamaño de la muestra, la definición del tipo de PC y la calidad metodológica de los estudios revisados. Esto indica la importancia de futuras investigaciones que utilicen un tamaño de muestra más amplio, además de un diseño de estudio más completo y métodos estadísticos validados y adecuados para mejorar la evaluación y así proporcionar una base sólida para prácticas clínicas fiables y validadas.

7. CONCLUSIONES

Los resultados de este TFG muestran que existe una falta de estudios con una calidad metodológica alta que analicen la fiabilidad y validez de las escalas observacionales de la marcha en pacientes en edad pediátrica con PC.

La EVE es la única escala que presenta tanto una buena fiabilidad como una buena validez. La ES y la EOM parecen ser fiables. Para el resto de las escalas, EME, EVM y EVEI, los resultados indican una falta de fiabilidad y validez.

BIBLIOGRAFÍA

1. Macías L, Fagoaga J. *Fisioterapia en Pediatría*. 2ª Edición. Barcelona: Editorial Médica Panamericana; 2021.
2. Espinoza Diaz C. I, Amaguaya Maroto G, Culqui Barrionuevo M, Espinosa Moya J, Silva Acosta J, Angulo Procel A, Rivera Pérez J, Avilés Jaya A. C. Prevalencia, factores de riesgo y características clínicas de la parálisis cerebral infantil. *Archivos Venezolanos de Farmacología y Terapéutica*. 2019;38(6):778-789.
3. Graham HK, Rosenbaum P, Paneth N, Dan B, Lin JP, Damiano DL, et al. Cerebral palsy. *Nat Rev Dis Primers*. 2016;2(1).
4. Grigoriu AI, Lempereur M, Bouvier S, Padure L, Brochard S. Characteristics of newly acquired gait in toddlers with unilateral cerebral palsy: Implications for early rehabilitation. *Ann Phys Rehabil Med*. 2021;64(3):101333.
5. Hurtado IL. La parálisis cerebral. Actualización del concepto, diagnóstico y tratamiento. *Pediatría Integral*. 2007; 11: 687-698.
6. Ruiz Brunner M de LM, Escobar Zuluaga J, Cieri ME, Ayllón C, Cuestas E. Sistemas de clasificación para niños, niñas y adolescentes con parálisis cerebral: su uso en la práctica clínica. *Rev Fac Cien Med Univ Nac Cordoba*. 2020;77(3):191–8.
7. Palisano RJ, Rosenbaum P, Bartlett D, Livingston MH. Content validity of the expanded and revised Gross Motor Function Classification System. *Dev Med Child Neurol*. 2008;50(10):744–50.
8. Raposo MR, Ricardo D, Teles J, Veloso AP, João F. Gait Analysis in Children with Cerebral Palsy: Are Plantar Pressure Insoles a Reliable Tool? *Sensors (Basel)*. 2022;22(14):5234.
9. Rodda J, Graham HK. Classification of gait patterns in spastic hemiplegia and spastic diplegia: a basis for a management algorithm. *Eur J Neurol*. 200;8(s5):98–108.
10. Vázquez VCC, Vidal RCA. Parálisis cerebral infantil: definición y clasificación a través de la historia. *Rev Mex Ortop Ped*. 2014;16(1):6-10.
11. Rathinam C, Bateman A, Peirson J, Skinner J. Observational gait assessment tools in paediatrics--a systematic review. *Gait Posture*. 2014;40(2):279–85.
12. Baker R, Esquenazi A, Benedetti MG, Desloovere K. Gait analysis: clinical facts. *Eur J Phys Rehabil Med*. 2016;52(4):560–74.
13. Chakraborty S, Nandy A, Kesar TM. Gait deficits and dynamic stability in children and adolescents with cerebral palsy: A systematic review and meta-analysis. *Clin Biomech (Bristol, Avon)*. 2020;71:11–23.

14. Armand S, Decoulon G, Bonnefoy-Mazure A. Gait analysis in children with cerebral palsy. *EFORT open Rev.* 2016;1(12):448–60.
15. Argibay J. C, TECNICAS PSICOMETRICAS. CUESTIONES DE VALIDEZ Y CONFIABILIDAD. *Subjetividad y Procesos Cognitivos.* 2006; (8):15-33.
16. Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res.* 2018;27(5):1147–57.
17. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol.* 2010;63(7):737–45.
18. Zanudin A, Mercer TH, Jagadamma KC, van der Linden ML. Psychometric properties of measures of gait quality and walking performance in young people with Cerebral Palsy: A systematic review. *Gait Posture.* 2017;58:30–40.
19. Mokkink LB, Boers M, van der Vleuten CPM, Bouter LM, Alonso J, Patrick DL, et al. COSMIN Risk of Bias tool to assess the quality of studies on reliability or measurement error of outcome measurement instruments: a Delphi study. *BMC Med Res Methodol.* 2020;20(1):293.
20. Mokkink LB, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, de Vet HCW, et al. COSMIN methodology for systematic reviews of Patient - Reported Outcome Measures (PROMs) - User manual. 2018.
21. Dobson F, Hinman RS, Hall M, Terwee CB, Roos EM, Bennell KL. Measurement properties of performance-based measures to assess physical function in hip and knee osteoarthritis: a systematic review. *Osteoarthr Cartil.* 2012;20(12):1548–62.
22. Abe H, Koyanagi S, Kusumoto Y, Himuro N. Intra-rater and inter-rater reliability, minimal detectable change, and construct validity of the Edinburgh Visual Gait Score in children with cerebral palsy. *Gait Posture* 2022;94:119–23.
23. Araújo PA, Kirkwood RN, Figueiredo EM. Validade e confiabilidade intra e interexaminadores da Escala Observacional de Marcha para crianças com paralisia cerebral espástica. *Brazilian J Phys Ther.* 2009;13(3):267–73.
24. Aroojis A, Sagade B, Chand S. Usability and Reliability of the Edinburgh Visual Gait Score in Children with Spastic Cerebral Palsy Using Smartphone Slow-Motion Video Technology and a Motion Analysis Application: A Pilot Study. *Indian J Orthop.* 2021;55(4):931–8.
25. Brown CR, Hillman SJ, Richardson AM, Herman JL, Robb JE, CR B, et al. Reliability and validity of the Visual Gait Assessment Scale for children with hemiplegic cerebral palsy

- when used by experienced and inexperienced observers. *Gait Posture*. 2008;27(4):648–52.
26. Del Pilar Duque Orozco M, Abousamra O, Church C, Lennon N, Henley J, Rogers KJ, et al. Reliability and validity of Edinburgh visual gait score as an evaluation tool for children with cerebral palsy. *Gait Posture*. 2016;49:14–8.
 27. Dickens WE, Smith MF. Validation of a visual gait assessment scale for children with hemiplegic cerebral palsy. *Gait Posture*. 2006;23(1):78–82.
 28. Maathuis KGB, van der Schans CP, van Iperen A, Rietman HS, Geertzen JHB. Gait in children with cerebral palsy: observer reliability of Physician Rating Scale and Edinburgh Visual Gait Analysis Interval Testing scale. *J Pediatr Orthop*. 2005;25(3):268–72.
 29. Mackey AH, Lobb GL, Walt SE, Stott NS. Reliability and validity of the Observational Gait Scale in children with spastic diplegia. *Dev Med Child Neurol*. 2003;45(1):4–11.
 30. Paci M, Mini G, Marchettini M, Ferrarello F. The Salford Gait Tool: Does the clinical experience of the raters influence the inter-rater reliability? *Dev Neurorehabil*. 2018;21(2):131–2.
 31. Toro B, Nester CJ, Farren PC. Inter- and intraobserver repeatability of the Salford Gait Tool: an observation-based clinical gait assessment tool. *Arch Phys Med Rehabil*. 2007;88(3):328–32.
 32. Toro B, Nester CJ, Farren PC. The development and validity of the Salford Gait Tool: an observation-based clinical gait assessment tool. *Arch Phys Med Rehabil*. 2007;88(3):321–7.
 33. Viehweger E, Pfund LZ, Hélix M, Rohon MA, Jacquemier M, Scavarda D, et al. Influence of clinical and gait analysis experience on reliability of observational gait analysis (Edinburgh Gait Score Reliability). *Ann Phys Rehabil Med*. 2010;53(9):535–46.
 34. Terwee CB, Prinsen CAC, Chiarotto A, de Vet HCW, Bouter LM, Alonso J, et al. COSMIN methodology for assessing the content validity of Patient - Reported Outcome Measures PROMs- User manual. 2018.

ANEXOS

Anexo 1. Estrategia de búsqueda

Pubmed: (("cerebral palsy"[MeSH Terms] OR ("cerebral"[All Fields] AND "palsy"[All Fields]) OR "cerebral palsy"[All Fields]) AND ("gait"[MeSH Terms] OR "gait"[All Fields] OR ("gait"[MeSH Terms] OR "gait"[All Fields]) AND "analysis"[All Fields]) OR ("gait disorders, neurologic"[MeSH Terms] OR ("gait"[All Fields] AND "disorders"[All Fields] AND "neurologic"[All Fields]) OR "neurologic gait disorders"[All Fields] OR "gait disorders neurologic"[All Fields])) AND (((("outcome"[All Fields] OR "outcomes"[All Fields]) AND ("measurability"[All Fields] OR "measurable"[All Fields] OR "measurably"[All Fields] OR "measure s"[All Fields] OR "measureable"[All Fields] OR "measured"[All Fields] OR "measurement"[All Fields] OR "measurement s"[All Fields] OR "measurements"[All Fields] OR "measurer"[All Fields] OR "measurers"[All Fields] OR "measuring"[All Fields] OR "measurings"[All Fields] OR "measurment"[All Fields] OR "measurments"[All Fields] OR "weights and measures"[MeSH Terms] OR ("weights"[All Fields] AND "measures"[All Fields]) OR "weights and measures"[All Fields] OR "measure"[All Fields] OR "measures"[All Fields])) OR ("observational"[All Fields] AND ("method s"[All Fields] OR "methods"[MeSH Terms] OR "methods"[All Fields] OR "method"[All Fields] OR "methods"[MeSH Subheading])) OR ("assess"[All Fields] OR "assessed"[All Fields] OR "assesment"[All Fields] OR "assesses"[All Fields] OR "assessing"[All Fields] OR "assessment"[All Fields] OR "assessment s"[All Fields] OR "assessments"[All Fields])) AND ("reproducibility of results"[MeSH Terms] OR ("reproducibility"[All Fields] AND "results"[All Fields]) OR "reproducibility of results"[All Fields] OR ("valid"[All Fields] OR "validate"[All Fields] OR "validated"[All Fields] OR "validates"[All Fields] OR "validating"[All Fields] OR "validation"[All Fields] OR "validational"[All Fields] OR "validations"[All Fields] OR "validator"[All Fields] OR "validators"[All Fields] OR "validities"[All Fields] OR "validity"[All Fields]) OR ("reliabilities"[All Fields] OR "reliability"[All Fields] OR "reliable"[All Fields] OR "reliability"[All Fields] OR "reliably"[All Fields]) OR ("sensitivity and specificity"[MeSH Terms] OR ("sensitivity"[All Fields] AND "specificity"[All Fields]) OR "sensitivity and specificity"[All Fields] OR "specificity"[All Fields] OR "specific"[All Fields] OR "specifically"[All Fields] OR "specification"[All Fields] OR "specifications"[All Fields] OR "specificities"[All Fields] OR "specifics"[All Fields] OR "specificities"[All Fields] OR "specifity"[All Fields]) OR ("sensitive"[All Fields] OR "sensitively"[All Fields] OR "sensitives"[All Fields] OR "sensitivities"[All Fields] OR "sensitivity and specificity"[MeSH Terms] OR ("sensitivity"[All Fields] AND "specificity"[All Fields]) OR "sensitivity and specificity"[All Fields] OR "sensitivity"[All Fields])))) AND (ffrft[Filter])

Cochrane: (cerebral palsy) AND (gait OR gait análisis OR gait disorders, neurologic) AND (outcomes measurement OR observational methods OR assessment) AND (reproducibility of results OR validity OR reliability OR specificity OR sensitivity)

Cinahl: (cerebral palsy) AND (gait OR gait análisis OR gait disorders, neurologic) AND (outcomes measurement OR observational methods OR assessment) AND (reproducibility of results OR validity OR reliability OR specificity OR sensitivity)

Web Of Science: (cerebral palsy) AND (gait OR gait análisis OR gait disorders, neurologic) AND (outcomes measurement OR observational methods OR assessment) AND (reproducibility of results OR validity OR reliability OR specificity OR sensitivity)

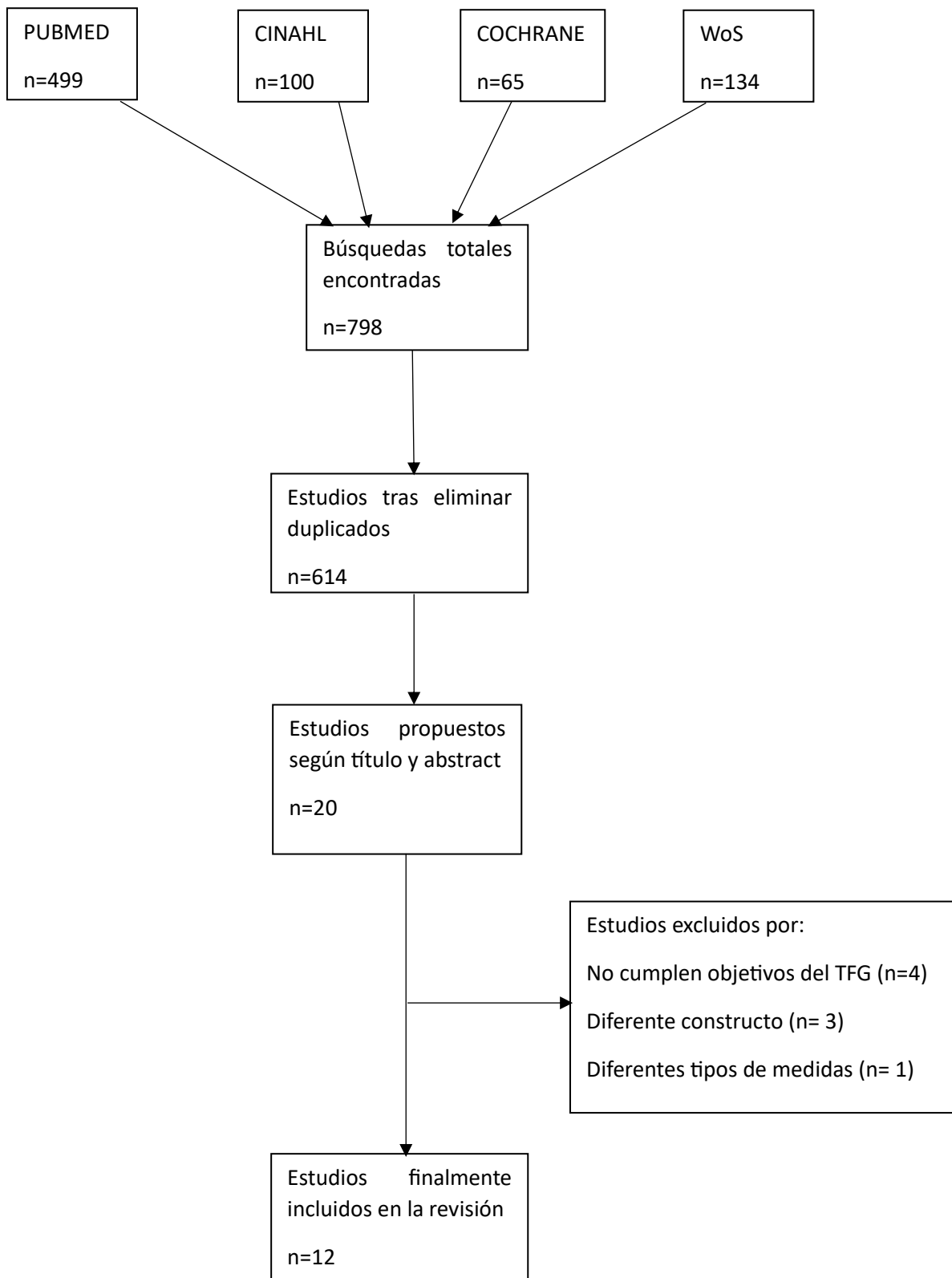


Figura 1: Diagrama de flujo de la selección de estudios según la metodología Cosmin

TABLA 1. EVALUACIÓN DE LOS CRITERIOS PARA BUENAS PROPIEDADES PSICOMÉTRICAS

Propiedad	Definición	Clasificación	Criterio de calidad
Validez de constructo	Grado en que las puntuaciones de un instrumento se relacionan con las puntuaciones de un instrumento se relacionan con un "gold standard"	+	El resultado está de acuerdo con la hipótesis
		?	No se especifica ninguna hipótesis
		-	El resultado no está de acuerdo con la hipótesis
Validez de criterio	Grado en que las puntuaciones de un instrumento se relacionan con un "gold standard"	+	Correlación con el estándar de oro es igual o mayor que 0,70 o ABC es igual o mayor que 0,70
		?	No se especifica el resultado
		-	Correlación inferior a 0,70 o ABC menor que 0,70
Fiabilidad	Grado de consistencia y precisión de las mediciones, nos dice si podemos confiar en obtener los mismos resultados cada vez que se realice una medición	+	CCI o Kappa ponderada es igual o superior de 0,70
		?	No se especifica el resultado
		-	CCI o Kappa ponderada menor de 0,70
Error de medición	Grado de error en la puntuación de los pacientes que no están relacionados con cambios reales en lo que estamos midiendo	+	CMD o LdA o $CV \cdot \sqrt{2} \cdot 1.96 < CMI$ o % de acuerdo específico $> 80\%$
		?	CMI no definido
		-	CMD o LdA o $CV \cdot \sqrt{2} \cdot 1.96 > CMI$; % de acuerdo específico $< 80\%$
+: positivo -: negativo ?: dudoso ABC: área bajo la curva; CCI: coeficiente correlación intraclase; CMD: cambio mínimo detectable; LdA: límites de acuerdo; CV: coeficiente de variación; CMI: cambio mínimo importante			

TABLA 2. CARACTERÍSTICAS SOCIODEMOGRÁFICAS Y CLÍNICAS DE LOS PACIENTES

INSTRUMENTO DE MEDIDA	ESTUDIO	POBLACIÓN			CARACTERÍSTICAS DE LA ENFERMEDAD	
		N	EDAD	SEXO	SCFMG	TIPO DE PC
Escala visual de Edimburgo (EVE)	Abe et al. (2022)	62	4 – 18 años	22 mujeres 40 hombres	<ul style="list-style-type: none"> Nivel I → 32 niños Nivel II → 25 niños Nivel III → 5 niños 	<ul style="list-style-type: none"> Unilateral espástica Bilateral espástica
	Aroojis et al. (2021)	30	4 – 14 años	9 mujeres 21 hombres	<ul style="list-style-type: none"> Nivel I → 1 Nivel II → 13 Nivel III → 16 	<ul style="list-style-type: none"> Diplejía (21) Hemiplejía (8) Monoplejía (1)
	Del pilar Duque Orotzco et al. (2016)	30	6 – 12 años	-	<ul style="list-style-type: none"> Nivel I → 10 Nivel II → 10 Nivel III → 10 	PC espástica
Escala observacional de la marcha (EOM)	Araújo et al. (2009)	18	6 – 12 años	7 mujeres 11 hombres	<ul style="list-style-type: none"> Nivel I → 11 niños Nivel II → 7 niños 	12 diplejía espástica 6 hemiplejía espástica
	Mackey et al. (2003)	20	6 – 12 años	9 mujeres 11 hombres	-	PC diplejía espástica
Escala visual de la marcha (EVM)	Brown et al. (2008)	4	4 – 12 años	-	-	Hemiplejía
	Dickens et al. (2006)	31	5 – 17 años	11 mujeres 20 hombres	<ul style="list-style-type: none"> Nivel I y II → 24 Nivel III y IV → 7 	Hemiplejía espástica

TABLA 2. CONTINUACIÓN

Escala visual de Edimburgo por intervalos (EVEI)	Maathuis et al. (2005)	24	3 – 10 años	6 mujeres 18 hombres	-	15 diplejía espástica 9 hemiplejía espástica (8 pierna afecta dcha. y 1 pierna afecta izqda.)
Escala de Salford (ES)	Paci et al. (2018)	7	8 – 14 años	1 mujer 6 hombres	-	4 diplejía espástica 3 hemiplejía espástica
	Toro et al. (2007)	13	6 – 16 años	-	-	PC espástica hemiplejía, diplejía y cuadriplejía
	Toro et al. (2007)	10	5 – 10 años	4 mujeres 6 hombres	-	PC espástica hemiplejía (3 niños), diplejía (3 niños) y cuadriplejía (4 niños)
Escala de la marcha de Edimburgo (EME)	Viehweger et al. (2010)	10	9 – 16 años	-	-	PC diplejía espástica
PC: parálisis cerebral; SCFMG: sistema de clasificación de función motoras gruesas						

TABLA 3. EVALUACIÓN DEL RIESGO DE SESGO PARA LA FIABILIDAD

ESTUDIO	1	2	3	4	5	6	7	8	9	PUNTUACIÓN FINAL
Abe et al.(2022)	NA	Muy bueno	Muy bueno	Muy bueno	Muy bueno	Muy bueno	Muy bueno	NA	NA	Muy bueno
Araújo et al. (2009)	NA	Muy bueno	Muy bueno	Adecuado	Adecuado	Muy bueno	NA	Muy bueno	NA	Adecuado
Aroojis et al. (2021)	NA	Muy bueno	Muy bueno	Muy bueno	Muy bueno	Muy bueno	NA	Adecuado	NA	Adecuado
Brown et al. (2008)	NA	Muy bueno	Muy bueno	Muy bueno	Muy bueno	Muy bueno	Adecuado	NA	NA	Muy bueno
Del Pilar Duque Orozco et al. (2016)	NA	Muy bueno	Muy bueno	Muy bueno	Muy bueno	Muy bueno	NA	Adecuado	NA	Adecuado
Dickens et al. (2006)	NA	Muy bueno	Muy bueno	Muy bueno	Muy bueno	Muy bueno	NA	Adecuado	NA	Adecuado
Maathuis et al. (2005)	NA	Muy bueno	Muy bueno	Adecuado	Adecuado	Muy bueno	NA	Adecuado	NA	Adecuado
Mackey et al. (2003)	NA	Muy bueno	Muy bueno	Muy bueno	Muy bueno	Muy bueno	NA	Muy bueno	NA	Muy bueno
Paci et al. (2018)	NA	NA	Muy bueno	Muy bueno	Muy bueno	Muy bueno	Muy bueno	NA	NA	Muy bueno
Viehweger et al. (2010)	NA	Muy bueno	Muy bueno	Muy bueno	Muy bueno	Muy bueno	Muy bueno	NA	NA	Muy bueno
<p>1.ESTABILIDAD DEL PACIENTE: ¿Se mantuvieron estables los pacientes en el tiempo transcurrido entre las mediciones del constructo? Al ser grabaciones de vídeo no aplica. 2.INTERVALO DE TIEMPO: ¿Fue apropiado el intervalo de tiempo entre las mediciones? 3.CONDICIONES DE MEDIDAS: ¿Fueron similares las condiciones de medición excepto la condición que se evalúa como fuente de variación? 4.ADMINISTRACIÓN DE MEDIDAS: ¿El profesional o los profesionales administraron la medición sin conocer las puntuaciones o valores de otra(s) medición(es) repetida(s) en los mismos pacientes? 5.ASIGNACIÓN DE PUNTUACIÓN: ¿El profesional o los profesionales asignaron puntuaciones o determinaron valores sin conocer las puntuaciones o valores de otras mediciones repetidas en los mismos pacientes? 6.OTROS DEFECTOS: ¿Hubo algún otro fallo importante en el diseño o en los métodos estadísticos del estudio? 7.MÉTODOS ESTADÍSTICOS PARA PUNTUACIONES CONTINUAS: Para las puntuaciones continuas, ¿se calculó un coeficiente de correlación intraclase (CCI)? 8.MÉTODOS ESTADÍSTICOS PARA PUNTUACIONES ORDINALES: Para las puntuaciones ordinales, ¿se calculó el valor kappa (ponderada)? 9.MÉTODOS ESTADÍSTICOS PARA PUNTUACIONES DICOTÓMICAS O NOMINALES: Para las puntuaciones dicotómicas/nominales, ¿se calculó el valor kappa para cada categoría frente a las otras categorías combinadas? NA: no aplicable</p>										

TABLA 4. EVALUACIÓN DEL RIESGO DE SESGO PARA EL ERROR DE MEDICIÓN

ESTUDIO	1	2	3	4	5	6	7	8	PUNTUACIÓN FINAL
Abe et al.(2022)	NA	Muy bueno	Muy bueno	Muy bueno	Muy bueno	Muy bueno	Muy bueno	NA	Muy bueno
Aroojis et al. (2021)	NA	Muy bueno	Muy bueno	Muy bueno	Muy bueno	Muy bueno	NA	Adecuado	Adecuado
Brown et al. (2008)	NA	Muy bueno	Muy bueno	Muy bueno	Muy bueno	Muy bueno	NA	Muy bueno	Muy bueno
Del Pilar Duque Orozco et al. (2016)	NA	Muy bueno	Muy bueno	Muy bueno	Muy bueno	Muy bueno	NA	Adecuado	Adecuado
Paci et al. (2018)	NA	NA	Muy bueno	Muy bueno	Muy bueno	Muy bueno	Muy bueno	NA	Muy bueno
Toro et al. (2007)	NA	Muy bueno	Muy bueno	Muy bueno	Muy bueno	Muy bueno	NA	Muy bueno	Muy bueno
Viehweger et al. (2010)	NA	Muy bueno	Muy bueno	Muy bueno	Muy bueno	Muy bueno	NA	Muy bueno	Muy bueno
<p>1.ESTABILIDAD DEL PACIENTE: ¿Se mantuvieron estables los pacientes en el tiempo transcurrido entre las mediciones del constructo? Al ser grabaciones de vídeo no aplica. 2.INTERVALO DE TIEMPO: ¿Fue apropiado el intervalo de tiempo entre las mediciones? 3.CONDICIONES DE MEDIDAS: ¿Fueron similares las condiciones de medición excepto la condición que se evalúa como fuente de variación? 4.ADMINISTRACIÓN DE MEDIDAS: ¿El profesional o los profesionales administraron la medición sin conocer las puntuaciones o valores de otra(s) medición(es) repetida(s) en los mismos pacientes? 5.ASIGNACIÓN DE PUNTUACIÓN: ¿El profesional o los profesionales asignaron puntuaciones o determinaron valores sin conocer las puntuaciones o valores de otras mediciones repetidas en los mismos pacientes? 6.OTROS DEFECTOS: ¿Hubo algún otro fallo importante en el diseño o en los métodos estadísticos del estudio? 7.MÉTODOS ESTADÍSTICOS PARA PUNTUACIONES CONTINUAS: Para las puntuaciones continuas, ¿se calculó el error estándar de medición (EEM), el cambio más pequeño detectable (CMD), los límites de acuerdo (LdA) o el coeficiente de variación (CV)? 8.MÉTODOS ESTADÍSTICOS PARA PUNTUACIONES DICOTÓMICAS, NOMINALES U ORDINALES: Para las puntuaciones dicotómicas/nominales/ordinales, ¿se calculó el porcentaje de acuerdo específico?</p> <p>NA: no aplica</p>									

TABLA 5. EVALUACIÓN DEL RIESGO DE SESGO SOBRE LA VALIDEZ DE CRITERIO

ESTUDIO	MÉTODOS ESTADÍSTICOS PARA PUNTUACIONES CONTINUAS	MÉTODOS ESTADÍSTICOS PARA PUNTUACIONES DICOTÓMICAS	OTROS DEFECTOS	PUNTUACIÓN FINAL
Araújo et al. (2009)	Inadecuado	NA	Muy bueno	Inadecuado
Brown et al. (2008)	Inadecuado	NA	Muy bueno	Inadecuado
Del Pilar Duque Orozco et al. (2016)	Inadecuado	NA	Muy bueno	Inadecuado
Dickens et al. (2006)	Inadecuado	NA	Muy bueno	Inadecuado
Mackey et al. (2003)	Inadecuado	NA	Muy bueno	Inadecuado
Toro et al. (2007)	Inadecuado	NA	Muy bueno	Inadecuado
<p>1: Para puntuaciones continuas, ¿se calcularon correlaciones o el área bajo la curva? Si se calcularon se considera adecuado; si no se calcularon se considera inadecuado.</p> <p>2: Para puntuaciones dicotómicas, ¿se determinaron la sensibilidad y especificidad? Si se calcularon, se considera adecuado; si no se calcularon se considera inadecuado.</p> <p>3: ¿Hubo otros errores importantes en el diseño o en los métodos estadísticos del estudio? Si no hubo otros errores metodológicos importantes se considera muy bueno y si hubo otros errores metodológicos importantes es inadecuado.</p> <p>NA: no aplica</p>				

TABLA 6. EVALUACIÓN DEL RIESGO DE SESGO PARA LA VALIDEZ DE CONSTRUCTO

ESTUDIO	CLARIDAD MEDICIÓN INSTRUMENTO COMPARADOR	CALIDAD INSTRUMENTO COMPARADOR	OTROS DEFECTOS	PUNTUACIÓN FINAL
Abe et al.(2022)	Muy bueno	Muy bueno	Muy bueno	Muy bueno
1: ¿Está claro qué mide el/los instrumento(s) comparador(es)? 2: ¿Fueron suficientes las propiedades de medición del/los instrumento(s) comparador(es)? 3: Evalúa si existen otras debilidades importantes en el diseño o los métodos estadísticos del estudio que podrían sesgar los resultados.				

TABLA 7. RESULTADOS PARA FIABILIDAD

INSTRUMENTO DE MEDIDA	ESTUDIO	IDIOMA DEL CUESTIONARIO	FIABILIDAD			ERROR DE MEDICIÓN		
			N	CALIDAD	RESULTADOS	N	CALIDAD	RESULTADOS
Escala visual de Edimburgo (EVE)	Abe et al. (2022)	Versión japonesa	62	Muy bueno	<ul style="list-style-type: none"> Intraexaminadores: <ul style="list-style-type: none"> Calificador A: CCI=0,96 Calificador B: CCI=0,90 Calificador C: CCI=0,97 Entre examinadores: CCI=0,91 	62	Muy bueno	Calificador A: <ul style="list-style-type: none"> EEM=1,7 CMD=4 Calificador B: <ul style="list-style-type: none"> EEM=2,6 CMD=6 Calificador C: <ul style="list-style-type: none"> EEM=1,5 CMD=3,6
	Aroojis et al. (2021)	Versión original (inglés)	30	Adecuado	<ul style="list-style-type: none"> Intraexaminadores: <ul style="list-style-type: none"> Evaluador 1: CdR=0,98 Evaluador 2: CdR=0,65 Entre examinadores: Kappa ponderada=0,47 	30	Adecuado	<ul style="list-style-type: none"> Intraexaminadores: <ul style="list-style-type: none"> Observador 1: % de acuerdo de 77,35 Observador 2: % de acuerdo de 88,62 Entre examinadores: % de acuerdo de 71,92

TABLA 7. CONTINUACIÓN

Escala visual de Edimburgo (EVE)	Del Pilar Duque Orozco et al. (2016)	Versión original (inglés)	30	Adecuado	<ul style="list-style-type: none"> • Intraobservadores Kappa ponderada=0,50 • Entre observadores Kappa ponderada= 0,46 	30	Adecuado	<ul style="list-style-type: none"> • Intraobservadores Acuerdo del 73,94% • Entre observadores Acuerdo del 72,88%
Escala observacional de la marcha (EOM)	Araújo et al. (2009)	Versión original (inglés)	18	Adecuado	<ul style="list-style-type: none"> • Intraobservadores Kappa ponderada=0,74 • Entre observadores Kappa ponderada=0,56 	-	-	-
	Mackey et al. (2003)	Versión original (inglés)	20	Muy bueno	<ul style="list-style-type: none"> • Intraobservadores Observador 1: pierna derecha=0,61kp y pierna izquierda=0,60kp Observador 2: pierna derecha=0,69kp y pierna izquierda=0,75kp • Entre observadores Sesión 1: pierna derecha=0,66kp y pierna izquierda=0,65kp Sesión 2: pierna derecha=0,60kp y pierna izquierda=0,65kp 	-	-	-

TABLA 7. CONTINUACIÓN

Escala visual de la marcha (EVM)	Brown et al. (2008)	Versión original (inglés)	4	Adecuado	<ul style="list-style-type: none"> Entre examinadores <p>Observadores expertos: CdR=6,76 (sesión 1) y CdR=4,36 (sesión 2)</p> <p>Observadores inexpertos: CdR=6,06 (sesión 1) y CdR=5,79 (sesión 2)</p>	4	Muy bueno	<ul style="list-style-type: none"> Intraexaminadores <p>Observadores expertos: acuerdo del 71,46%</p> <p>Observadores inexpertos: acuerdo del 60,71%</p> <ul style="list-style-type: none"> Entre examinadores <p>Observadores expertos: acuerdo del 64,3% sesión 1 y del 60,7% sesión 2</p> <p>Observadores inexpertos: acuerdo del 53,79% sesión 1 y del 62,14% sesión 2</p>
	Dickens et al. (2006)	Versión original (inglés)	31	Adecuado	<ul style="list-style-type: none"> Intraobservadores <p>Observador 1: kappa ponderada=0,52</p> <p>Observador 2: kappa ponderada=0,54</p> <ul style="list-style-type: none"> Entre observadores <p>Sesión 1: kappa ponderada=0,66</p> <p>Sesión 2: kappa ponderada=0,59</p>	-	-	-

TABLA 7. CONTINUACIÓN

Escala visual de Edimburgo por intervalos (EVEI)	Maathuis et al. (2005)	Versión original (inglés)	22	Adecuado	<ul style="list-style-type: none"> • Intraobservadores Observador 1: pierna derecha=9,1 y pierna izquierda=7,8 Observador 2: pierna derecha=10,3 y pierna izquierda=8,5 Observador 3: pierna derecha=13,5 y pierna izquierda 11,1 • Entre observadores Pierna derecha: $p < 0,001$ Pierna izquierda: $p = 0,003$ 	-	-	-
Escala de Salford (ES)	Paci et al. (2018)	Versión original (inglés)	7	Muy bueno	<ul style="list-style-type: none"> • Entre observadores Fisioterapeutas pediátricos: CCI=0,792 Fisioterapeutas no pediátricos: CCI=0,748 Estudiantes de fisioterapia: CCI=0,613 	7	Muy bueno	<ul style="list-style-type: none"> • Entre observadores Fisioterapeutas pediátricos: <ul style="list-style-type: none"> - EEM=0,55 - CMD=1,52 Fisioterapeutas no pediátricos: <ul style="list-style-type: none"> - EEM=0,55 - CMD=1,50 Estudiantes de fisioterapia_ <ul style="list-style-type: none"> - EEM=1,03 - CMD=2,83

TABLA 7. CONTINUACIÓN

Escala de Salford (ES)	Toro et al. (2007)	Versión original (inglés)	-	-	-	13	Muy bueno	<ul style="list-style-type: none"> • Intraobservadores: Acuerdo total 75% • Entre observadores: Acuerdo total 77%
Escala de la marcha de Edimburgo (EME)	Viehweger et al. (2010)	Versión francesa	10	Muy bueno	<ul style="list-style-type: none"> • Intraobservadores CCI=0,59 	10	Muy bueno	<ul style="list-style-type: none"> • Entre observadores: Acuerdo total 57,1%
CCI: coeficiente correlación intraclase; EEM: error estándar de medición; CMD: cambio mínimo detectable; CdR: coeficiente de repetibilidad kp: kappa ponderada								

TABLA 8. RESULTADOS VALIDEZ

INSTRUMENTO DE MEDIDA	ESTUDIO	IDIOMA DEL CUESTIONARIO	VALIDEZ DE CRITERIO			VALIDEZ DE CONSTRUCTO		
			N	CALIDAD	RESULTADOS	N	CALIDAD	RESULTADOS
Escala visual de Edimburgo (EVE)	Abe et al. (2022)	Versión japonesa				62	Muy bueno	Evaluador A: r=-0,73 Evaluador B: r=-0,71 Evaluador C: r=0,69 P<0,001
	Del Pilar Duque Orozco et al. (2016)	Versión original (inglés)	30	Inadecuado	Acuerdo con 3DGA del 61,06%	-	-	-
Escala observacional de la marcha (EOM)	Araújo et al. (2009)	Versión original (inglés)	18	Inadecuado	Kappa ponderada = 0,40 P<0,05	-	-	-
	Mackey et al. (2003)	Versión original (inglés)	20	Inadecuado	Concordancia con 3DGA: Observador 1: <ul style="list-style-type: none">• Pierna derecha: kp=0,52• Pierna izquierda; kp=0,51 Observador 2: <ul style="list-style-type: none">• Pierna derecha: kp=0,68• Pierna izquierda: kp=0,74	-	-	-

TABLA 8. CONTINUACIÓN

Escala visual de la marcha (EVM)	Brown et al. (2008)	Versión original (inglés)	4	Inadecuado	<p>Acuerdo con 3DGA:</p> <ul style="list-style-type: none"> • Observadores expertos: <p>Sesión 1: acuerdo 60% y sesión 2: 55% de acuerdo</p> <ul style="list-style-type: none"> • Observadores inexpertos: <p>Sesión 1: acuerdo 49% y sesión 2:50% de acuerdo</p>	-	-	-
	Dickens et al. (2006)	Versión original (inglés)	31	Inadecuado	<p>Concordancia con 3DGA</p> <p>Observador 1:</p> <ul style="list-style-type: none"> • Sesión 1: kappa ponderada=0,24 • Sesión 2: kappa ponderada=0,17 <p>Observador 2:</p> <ul style="list-style-type: none"> • Sesión 1: kappa ponderada=0,21 • Sesión 2: kappa ponderada=0,23 	-	-	-
Escala de Salford (ES)	Toro et al. (2007)	Versión original (inglés)	10	Inadecuado	<p>Acuerdo promedio del 58% en las categorías de calificación y acuerdo promedio del 80% basado en estimaciones subjetivas</p>	-	-	-
kp: kappa ponderada; 3DGA: 3D Gait Analysis								

TABLA 9. RESULTADOS EVALUACIÓN DE CALIDAD DE PROPIEDADES PSICOMÉTRICAS

INSTRUMENTO DE MEDIDA	DE	ESTUDIO	EVALUACIÓN FIABILIDAD	EVALUACIÓN ERROR DE MEDICIÓN	EVALUACIÓN VALIDEZ CRITERIO	DE	EVALUACIÓN VALIDEZ DE CONSTRUCTO
Escala visual de Edimburgo (EVE)	de	Abe et al.(2022)	+	?	NA		?
		Aroojis et al. (2021)	Intraobservadores: ? Entre observadores: -	Intraobservadores: <ul style="list-style-type: none"> • Observador 1: - • Observador 2: + Entre observadores: -	NA		NA
		Del Pilar Duque Orozco et al. (2016)	-	-	?		NA
Escala observacional de la marcha (EOM)		Araújo et al. (2009)	Intraobservadores: + Entre observadores: -	NA	?		NA
		Mackey et al. (2003)	Intraobservadores: <ul style="list-style-type: none"> • Observador 1: - • Observador 2: + Entre observadores: <ul style="list-style-type: none"> • Sesión 1: - • Sesión 2: - 	NA	?		NA
Escala visual de la marcha (EVM)		Brown et al. (2008)	?	-	?		NA
		Dickens et al. (2006)	-	NA	?		NA

TABLA 9. CONTINUACIÓN

Escala visual de Edimburgo por intervalos (EVEI)	Maathuis et al. (2005)	?	NA	NA	NA
Escala de Salford (ES)	Paci et al. (2018)	Entre observadores: <ul style="list-style-type: none"> • Fisioterapeutas pediátricos: + • Fisioterapeutas no pediátricos: + • Estudiantes de fisioterapia: - 	Entre observadores: <ul style="list-style-type: none"> • Fisioterapeutas pediátricos: ? • Fisioterapeutas no pediátricos: ? • Estudiantes de fisioterapia: ? 	NA	NA
	Toro et al. (2007)	NA	-	NA	NA
	Toro et al. (2007)	NA	NA	?	NA
Escala de la marcha de Edimburgo (EME)	Viehweger et al. (2010)	-	-	NA	NA
+: positivo -: negativo ?: dudoso NA: no aplicable					