



UNIVERSIDAD
DE MÁLAGA

ED-UMA
Escuela de Doctorado
UNIVERSIDAD DE MÁLAGA



FACULTAD DE
CIENCIAS DE LA SALUD
Universidad de Málaga

**ESCUELA DE DOCTORADO DE LA UNIVERSIDAD DE
MÁLAGA**

Programa de Doctorado en Ciencias de la Salud

FACULTAD DE CIENCIAS DE LA SALUD

**ANÁLISIS PSICOMÉTRICO Y
VALIDACIÓN DE
INSTRUMENTOS DE
EVALUACIÓN FUNCIONAL Y DE
DESEMPEÑO OCUPACIONAL
DESDE TERAPIA OCUPACIONAL
EN PERSONAS QUE HAN
SUFRIDO UN ICTUS**

TESIS *DOCTORAL*

2025

DOCTORANDO/A: PABLO DE BLAS ZAMORANO

DIRECTOR/A: DR. DAVID PÉREZ CRUZADO
DR. JOSE ANTONIO MERCHAN BAEZA

TUTOR/A: DR. DAVID PÉREZ CRUZADO



Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

El hombre que mueve montañas empieza apartando piedras.

La gota de agua perfora la roca no por su fuerza, sino por su constancia.



UNIVERSIDAD
DE MÁLAGA

AUTOR: Pablo de Blas Zamorano



<https://orcid.org/0009-0001-4242-6654>

EDITA: Publicaciones y Divulgación Científica. Universidad de Málaga



Esta obra está bajo una licencia de Creative Commons

Reconocimiento- NoComercial-SinObraDerivada 4.0

Internacional: <https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>

Cualquier parte de esta obra se puede reproducir sin atribución de los autores. No se puede hacer uso comercial de la obra y no se puede alterar, transformar o hacer obras derivadas.

Esta Tesis Doctoral está depositada en el Repositorio

Institucional de la Universidad de Málaga: riuma.uma.es

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

**DECLARACIÓN DE AUTORÍA Y ORIGINALIDAD DE LA TESIS
PRESENTADA PARA OBTENER EL TÍTULO DE DOCTOR**

D./D^a: **PABLO DE BLAS ZAMORANO**

Estudiante del Programa de Doctorado en Ciencias de la Salud de la Universidad de Málaga, y autor/a de la tesis presentada para la obtención del título de doctor/a por la Universidad de Málaga, titulada:

Realizada bajo la tutorización de **DAVID PÉREZ CRUZADO** y la dirección de director/a **DAVID PÉREZ CRUZADO** y de **JOSE ANTONIO MERCHÁN BAEZA**.

DECLARO QUE:

La tesis presentada es una obra original que no infringe los derechos de propiedad intelectual, ni los derechos de propiedad industrial u otros, conforme al ordenamiento jurídico vigente (Real Decreto Legislativo 1/1996 de 12 de abril, por el que se aprueba el texto refundido de la Ley de Propiedad Intelectual, regularizando, aclarando y armonizando las disposiciones legales vigentes sobre la materia), modificado por la Ley 2/2019 de 1 de marzo.

Igualmente, asumo ante la Universidad de Málaga y ante cualquier otra instancia, la responsabilidad que pudiera derivarse en caso de plagio de contenidos en la tesis presentada, conforme al ordenamiento jurídico vigente.

En Málaga, a 25 de Septiembre de 2025

Fdo.: **PABLO DE BLAS ZAMORANO**

(Doctorando/a)

Fdo.: **DAVID PÉREZ CRUZADO**

(Tutor/a)

Fdo: **DAVID PÉREZ CRUZADO**

(Director/a)

Fdo: **JOSÉ ANTONIO MERCHÁN BAEZA**

(Director/a)

DECLARACIÓN DE CONFLICTO DE INTERÉS

Tanto el doctorando como los directores de la tesis declaran no tener ningún conflicto ni interés derivado con terceros como consecuencia del desarrollo de este estudio.

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

D./D^a: **DAVID PÉREZ CRUZADO**

Como director/a de la tesis doctoral titulada:

Presentada para su admisión a trámite por D./D^a: **PABLO DE BLAS ZAMORANO**

CERTIFICA

Que dicho trabajo ha sido realizado bajo su dirección y consideran que reúne los requisitos y calidad científica necesaria para ser defendido y juzgado por el tribunal de tesis correspondiente, a fin de optar al Grado de Doctor/a por la Universidad de Málaga.

Y para que conste a los efectos oportunos, en cumplimiento de las disposiciones vigentes,

Expide y firma el presente certificado en Málaga a 25 de Septiembre de 2025

Fdo.:

Dr. **DAVID PÉREZ CRUZADO**

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

D./D^a: **JOSÉ ANTONIO MERCHÁN BAEZA**

Como director/a de la tesis doctoral titulada:

Presentada para su admisión a trámite por D./D^a: **PABLO DE BLAS ZAMORANO**

CERTIFICA

Que dicho trabajo ha sido realizado bajo su dirección y consideran que reúne los requisitos y calidad científica necesaria para ser defendido y juzgado por el tribunal de tesis correspondiente, a fin de optar al Grado de Doctor/a por la Universidad de Málaga.

Y para que conste a los efectos oportunos, en cumplimiento de las disposiciones vigentes,

Expide y firma el presente certificado en Málaga a 25 de Septiembre de 2025

Fdo.:

Dr. JOSÉ ANTONIO MERCHÁN BAEZA

INFORME DE LA DIRECCIÓN SOBRE LAS PUBLICACIONES QUE AVALAN LA TESIS

D./D^a: **DAVID PÉREZ CRUZADO** y D/D^a. **JOSÉ ANTONIO MERCHÁN BAEZA**

Como director/a de la tesis:

Presentada por D./D^a: **PABLO DE BLAS ZAMORANO**

INFORMA

- Que dichas publicaciones han sido presentadas, publicadas o aceptadas para su publicación, muestra o reproducción con posterioridad a la fecha de matrícula del doctorando/a en el Programa de Doctorado.
- Que en ellas consta la Universidad de Málaga, a través de la afiliación de sus directores y el/la doctorando/a.
- Que el/la doctorando/a consta como primer/a o segundo/a autor/a de todas ellas¹.
- Que dichas publicaciones NO han sido utilizadas en tesis anteriores.

Lo que firma a los efectos oportunos, en Málaga a 25 de Septiembre de 2025.

Fdo.:

Fdo.:

Dr. DAVID PÉREZ CRUZADO

Dr. JOSÉ ANTONIO MERCHÁN BAEZA

Director/a

Director/a

¹ Según el art. 19 del Reglamento de Doctorado de la Universidad de Málaga, sólo en casos realmente excepcionales, la Comisión Académica del Programa de Doctorado podrá autorizar que en una contribución que avale una tesis, el doctorando figure en una posición posterior a la segunda, o que la contribución se haya producido en un periodo de investigación previo a la matrícula del doctorando en el Programa, a la vista de las justificaciones presentadas y con el visto bueno de la Comisión de Posgrado.

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

AGRADECIMIENTOS

Hay mucho que agradecer y a multitud de personas. Antes de comenzar a recordar las personas que me han rodeado durante este proceso, debo pedir disculpas por las otras que olvidaré mencionar.

En primer lugar, quiero agradecer a mis directores de tesis, Dr. David Pérez Cruzado, por acogerme sin conocerme, por su confianza, dedicación absoluta, disponibilidad ilimitada, soporte, guía, paciencia y cercanía que han facilitado enormemente este arduo proceso. A mi director Dr. José Antonio Merchán Baeza, por su gran conocimiento, su incansable entrega, su paciencia conmigo, su meticulosidad y perfeccionismo, y su afán de querer que aprenda. Ambos han sido una gran guía durante este proceso, aportando luz y esperanza en momentos de flaqueza, y dándome la oportunidad de aprender, crecer y mejorar; sabiendo llevar no solo un proyecto investigador, sino a la persona que hay detrás, que en ocasiones es más difícil que lo primero. Han sido ejemplos a seguir y gracias a ellos me siento realizado personal y profesionalmente. Ojalá esto no suponga un punto y final, sino un punto y seguido y podamos seguir trabajando juntos.

Ellos simbolizan la punta del iceberg, ya que me gustaría agradecer a todos mis profesores desde los comienzos de mi carrera como terapeuta ocupacional, pues cada uno puso su granito de arena para que hoy haya llegado hasta aquí.

Reconocer también a todas las personas que han colaborado en este proyecto, ya sean como colaboradoras en investigaciones y artículos; colegas con los que tener una charla enriquecedora sobre la TO, la investigación, o simplemente para desahogarse; compañeros que me han abierto las puertas de sus centros; y también mencionar a los pacientes que han participado en el estudio, ya que sin su predisposición y amabilidad no hubiese sido posible; así como agradecer a los lugares en los que he trabajado y que me han hecho crecer como persona y como profesional (tanto clínico, como investigador, como docente).

Es el momento de acordarse de amigos y compañeros que la vida cuza en tu camino y que todos ellos forman parte de tí, algunos acompañan durante diferentes etapas de la vida, otros se quedan en el camino, pero todos ellos te enseñan y, en una medida u otra, me han aportado y ayudado para llegar hasta aquí. Una mención especial para aquellos que ya no están entre nosotros.

Como no, mencionar a las personas que han estado conmigo desde que nací, dándome apoyo y amor, mi familia. Cualquier agradecimiento se queda corto para todo lo que merecen. Gracias a mis tíos, a mis primos, por ser siempre un ejemplo, sacarnos hacia delante, mantenernos unidos y desearnos siempre lo mejor. Este camino ha sido largo, pero ellos lo han amenizado dándome un

chute de energía cada vez que nos veíamos. Mención especial a los 4 ángeles que me guían y me cuidan desde el cielo, gracias a mis cuatro abuelos (Felix, Vicenta, Rubén y Aurora). Para mí, ellos siempre simbolizan el amor puro, como el que nos transmitieron a sus hijos y nietos; ellos me han dado fuerzas en muchos momentos y han estado presentes en mi mente y corazón durante toda mi vida y toda la tesis.

Quiero terminar mis agradecimientos con las cuatro personas más importantes de mi vida:

- Comenzaré con María, ya que ella me ha aguantado en los peores momentos de la tesis, cuando necesitaba parar y desconectar; también dando apoyo en momentos de máximo estrés; siempre reforzando el día a día dando mucho cariño, con una sonrisa y un halo de esperanza; siendo un pilar fundamental durante este duro proceso mientras construíamos un bonito futuro junto a Teo, Loki y Dana. Agradecer también a toda su familia por la acogida, el apoyo, el cariño, el humor y el buen rollo.
- Continúo con mis padres, Rubén y Graci, que saben cómo me siento antes de que yo pueda darme cuenta, saben lo que necesito antes de pedirlo, me conocen como nadie y siempre se han desvivido por sus hijos, madrugando y trabajando muy duro para sacarnos adelante, haciendo sacrificios y dando el amor más grande y real que conozco de forma incondicional. Sois un apoyo en todo momento y el mayor ejemplo para vuestros hijos y gente que os rodea. Imposible agradecerlos ni en mil años todo lo que habéis hecho por mí.
- Para finalizar, voy a hablar de la niña de mis ojos, mi hermana Lucía Ming, una persona que llegó a mi vida para cambiármela por completo, para salvarme, para hacerme feliz, para tener a alguien con quien reír, llorar, enfadarme, cotillear, conspirar o ir a ver al Rayo. La hermana que a pesar de los kilómetros siempre está, a la que quiero con locura y agradezco toda la alegría que nos has dado.

Es difícil encontrar las palabras y la manera de agradecerlos todo lo recibido a los cuatro, solo desearos siempre lo mejor y deciros que esta tesis es tanto mía como vuestra, ya que lo habéis vivido desde dentro, y sin vosotros no hubiera sido posible nada de lo que he logrado. Sois mi gran suerte; sois mi vida. Os quiero.

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

ÍNDICE

Tabla de contenido

1. Índice de siglas y abreviaturas	20
2. Índice de figuras y tablas.....	24
3. Estructura general de la tesis	26
4. Marco teórico.....	29
a. Ictus.....	30
i. Definición	30
ii. Tipos	31
iii. Epidemiología: incidencia y prevalencia	33
iv. Factores de riesgo	35
1. Factores de riesgo en supervivientes de ictus	36
v. Consecuencias.....	37
1. Alteraciones motoras	37
2. Alteraciones sensitivas	38
3. Alteraciones cognitivas	39
4. Alteraciones perceptivas.....	40
5. Alteraciones emocionales	41
6. Discapacidad e impacto.....	42
7. Gasto socioeconómico y sanitario	43
vi. Fases de evolución clínica y pronóstico	44
1. Fases tras el ictus.....	44
2. Pronóstico tras el ictus.....	46
a. Factores pronósticos.....	46
b. Protocolos de predicción de recuperación.....	50
b. Cuestionarios de evaluación en Terapia Ocupacional	52
i. Necesidad de evaluación en Terapia Ocupacional	52
ii. Definición de cuestionario	55
iii. Creación de cuestionarios	56
1. Definición del constructo.....	56
2. Finalidad del cuestionario	56
3. Elaboración de los ítems	57
4. Número de ítems	57
5. Contenido	57
6. Definición y disposición de los ítems	58
7. Identificación de sesgos	58
8. Clasificación de las respuestas.....	58
9. Puntuación de los ítems	58
10. Fase de validación.....	59
11. Prueba piloto o pre-test	59
12. Evaluación de las propiedades psicométricas.....	59
iv. Evaluación del vestido.....	59

1. Dificultades comunes en la actividad del vestido	61
a. El vestido tras un ictus	61
2. Nottingham Stroke Dressing Assessment.....	64
v. Evaluación de la función sensoriomotora del miembro superior	65
1. Escalas de evaluación	66
2. Fugl-Meyer Assessment	67
c. Análisis psicométrico y revisión de los cuestionarios de evaluación	70
i. Normativa COSMIN	70
1. Lista de verificación COSMIN	70
2. Guía COSMIN para revisiones sistemáticas	72
a. Búsqueda de la literatura	72
b. Evaluación de las propiedades de medición.....	74
c. Selección del cuestionario.....	77
ii. Criterios de Terwee y Prinsen	79
1. Propiedades psicométricas: calidad de los resultados ..	82
a. Validez de contenido.....	82
b. Validez estructural.....	84
c. Consistencia interna.....	85
d. Fiabilidad	86
e. Error de medición.....	86
f. Prueba de hipótesis para la validez de constructo	87
g. Validez transcultural/invarianza de medición.....	87
h. Validez de criterio.....	88
i. Sensibilidad al cambio	88
iii. Método Grade	89
1. Evaluación de la calidad de la evidencia y elaboración de la recomendación	90
d. Validación de cuestionarios de evaluación.....	96
i. Traducción y adaptación transcultural de cuestionarios	96
1. Guía ISPOR.....	96
ii. Propiedades psicométricas de los instrumentos de medida	99
1. Principios y propiedades psicométricas de las escalas ..	99
2. Validez de contenido	100
a. Primera etapa	100
b. Segunda etapa.....	102
c. Tercera etapa.....	102
3. Validez estructural	103
4. Consistencia interna	104
5. Fiabilidad	105
a. Fiabilidad intra-evaluador o fiabilidad test-retest	105
b. Fiabilidad inter-evaluador	106
6. Error de medición	107

a. Parámetros estimados del SEM	107
b. Etapas del SEM	107
7. Validez de criterio convergente	109
8. Validez de criterio	109
9. Prueba de hipótesis	111
10. Sensibilidad al cambio	111
11. Validez transcultural	111
5. Objetivos.....	114
a. Objetivo general	115
b. Objetivos específicos	115
6. Marco empírico.....	117
a. Revisión psicométrica de cuestionarios de evaluación	118
i. Índice de Barthel y Medida de Independencia Funcional en pacientes con ictus: una revisión sistemática psicométrica	119
1. Abstract	122
2. Resumen	123
3. Introducción	125
4. Material y Métodos	127
5. Resultados	145
6. Discusión	230
7. Conclusión	235
8. Bibliografía	237
ii. Escala Fugl-Meyer Assessment para miembro superior en ictus: una revisión sistemática psicométrica	246
1. Abstract	249
2. Resumen	250
3. Introducción	253
4. Material y Métodos	255
5. Resultados	264
6. Discusión	338
7. Conclusión	341
8. Bibliografía	343
b. Investigación clínica	352
i. Actualización, traducción, adaptación transcultural y validación de Nottingham Stroke Dressing Assessment al español	354
1. Abstract	357
2. Resumen	358
3. Introducción	359
4. Material y Métodos	362
5. Resultados	367
6. Discusión	377
7. Conclusión	381
8. Bibliografía	382
ii. Validación de la versión española de Fugl-Meyer Assessment en población española tras un ictus	389

1. Abstract	392
2. Resumen	393
3. Introducción	394
4. Material y Métodos	396
5. Resultados	399
6. Discusión	405
7. Conclusión	408
8. Bibliografía	410
7. Discusión	418
a. Análisis psicométrico y revisión de cuestionarios de evaluación	420
i. Normativa COSMIN	421
ii. Criterios de Terwee y Prinsen	423
iii. Método GRADE	426
iv. Revisión psicométrica del Índice de Barthel y FIM	427
1. Calidad metodológica	427
2. Calidad de la evidencia	429
3. Principales conclusiones	430
v. Revisión psicométrica de Fugl-Meyer Assessment	431
b. Validación de escalas de evaluación	434
i. Traducción y adaptación transcultural de cuestionarios	434
ii. Propiedades psicométricas de los cuestionarios	435
1. Validez estructural	435
2. Consistencia interna	436
3. Fiabilidad	437
4. Error de medición	437
5. Validez de criterio	438
6. Prueba de hipótesis	438
7. Sensibilidad al cambio	439
iii. Validación de ENVI-R	440
iv. Validación de Fugl-Meyer Assessment	442
c. Limitaciones	444
d. Implicaciones para la práctica clínica, para la investigación y recomendaciones para futuras líneas de investigación	447
8. Conclusiones	450
a. Conclusión general	451
b. Conclusiones específicas	451
9. Competencias adquiridas	453
a. Competencias instrumentales	454
b. Competencias interpersonales	455
c. Competencias sistemáticas	456
10. Prospectiva	458
11. Bibliografía	460
12. Anexos	511
a. Anexo 1: Fugl-meyer Assessment de Miembro Superior	512
b. Anexo 2: Fugl-Meyer Assessment de Miembro Inferior	515

c. Anexo 3: NSDA-R Instructions for the therapist	517
d. Anexo 4: NSDA-R Scale	520
e. Anexo 5: NSDA-R Action error analysis	539
f. Anexo 6: ENVI-R Instrucciones para el terapeuta	542
g. Anexo 7: ENVI-R Escala	545
h. Anexo 8: ENVI-R Análisis de los errores de acción	565
i. Anexo 9: Licencia de uso de la figura 1	568
j. Anexo 10: Licencia de uso de la figura 3	575
k. Anexo 11: Modelo de Hoja Informativa, Consentimiento Informado y Hoja de revocación	580
l. Anexo 12: Autorización para investigar en Centro Asistencial 1	584
m. Anexo 13: Autorización para investigar en Centro Asistencial 2	585
n. Anexo 14: Autorización para investigar en Centro Asistencial 3	586
o. Anexo 15: Autorización para investigar en Centro Asistencial 4	587
p. Anexo 16: Autorización para investigar en Centro Asistencial 5	588
q. Anexo 17: Autorización para investigar en Centro Asistencial 6	589
r. Anexo 18: Autorización para investigar en Centro Asistencial 7	590
s. Anexo 19: Autorización para investigar en Centro Asistencial 8	591
t. Anexo 20: Autorización para investigar en Centro Asistencial 9	592
u. Anexo 21: Autorización para investigar en Centro Asistencial 10	593
v. Anexo 22: Autorización para investigar en Centro Asistencial 11	594

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

1. ÍNDICE DE SIGLAS Y ABREVIATURAS

ADF: Distribución Asintótica Libre

AHRQ: Agency for Healthcare Research and Quality

AIT: Accidente Isquémico Transitorio

AMAT: Arm Motor ability Test

ANOVA: ANalysis Of VAriance

APA: American Psychological Association

ARAT: Action Research Arm Test

AUC: Area Under the Curve

AUC/ROC: Area Under the Curve / Receiver Operating Characteristic

AVD: Actividades de la Vida Diaria

BBT: Box and Block Test

BI: Barthel Index/Índice de Barthel

CAHAI: Chedoke Arm and Hand Activity Inventory

CEBM: Centre for Evidence Based Medicine

CFA: Confirmatory Factor Analysis

CFI: Comparative Fit Index

CIF: Clasificación Internacional del Funcionamiento, de la Discapacidad y de la Salud

CMSA: Chedoke-McMaster Astroke Assessment

COSMIN: COnsensus-based Standards for the selection of health Measurement Instruments

CV: Coeficiente de Variación

DeCS: Descriptores de Ciencias de la Salud

DIF: Differential Item Function

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

ECA: Ensayo Clínico Aleatorizado

EFA: Exploratory Factor Analysis/Análisis Factorial Exploratorio

ENVI-R: Escala Nottingham de Vestido en Ictus-Revisada

EO: Evaluación Ocupacional

ESD: Early Supported Discharge

ESO: European Stroke Organisation

HSA: Hemorragia Sub-Aracnoidea

FAI: Frenchai Activity Index

FIM: Functional Independence Measure/Medida de Independencia Funcional

FMA: Fugl-Meyer Assessment

FMA-LE: Fugl-Meyer Assessment – Lowe Extremity

FMA-MI: Fugl-Meyer Assessment de Miembro Inferior

FMA-MS: Fugl-Meyer Assessment de Miembro Superior

FMA-UE: Fugl-Meyer Assessment – Upper Extremity

GBD: Global Burden of Disease

GRADE: Grading of Recommendations Assessment, Development and Evaluation

ICC: Intraclass Correlation Coefficient

IRT: Item Reponse Theory

ISPOR: International Society for Pharmacoeconomics and Outcomes Research

KMO: Kaiser-Meyer-Olkin

KR: Kuder-Richardson

LoA: Limits of Agreement

MGCFA: Multi-Group Confirmatory Factor Analysis

MIC: Minimal Important Change

ML: Máxima Verosimilitud

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

mRS: modified Rankin Scale/Escala de Rankin modificada

NEADL: Nottingham Extended Activities of Daily Living

NHPT: Nine Hole Peg Test

NICE: National Institute for Health and Clinical Excellence

NIHSS: National Institute of Health Stroke Scale

NSDA: Nottingham Stroke Dressing Assessment

NSDA-R: Nottingham Stroke Dressing Assessment-Revised

OMI: Outcome Measurement Instrument/Instrumento de Medición de Resultados

OMS: Organización mundial de la Salud

PPT: Physical Performance Test

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

PRO: Patient-Reported Outcome

PROM: Patient-Reported Outcome Measure

PROMIS: Patient-Reported Outcomes Measurement Information System

RHB: Rehabilitación

RM: Resonancia Magnética

RMSEA: Root Mean Square Error of Approximation

SAC-MOT: Comité Asesor Científico sobre las Propiedades de Medición

SAS: Servicio Andaluz de salud

SDC: Smallest Detectable Change

SEN: Sociedad Española de Neurología

SEM: Standard Error of Mean

SL: Mínimos Cuadrados

STARD: Standards for Reporting of Diagnosis Accuracy

SRMR: Standardized Root Mean Square Residual

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

TAC: Traducción y Adaptación Transcultural

TC: Tomografía Computarizada

TCT: Test Classic Theory

TLI: Tucker-Lewis Index

TO: Terapia Ocupacional

UBDS: Upper Body Dressing Scale

UI: Unidad de Ictus

WMFT: Wolf Motor Function Test

2. ÍNDICE DE FIGURAS Y TABLAS

2.1 Índice de figuras

Figura 1: Ictus.

Figura 2: Clasificación de los ictus según el tipo.

Figura 3: Prevalencia mundial del ictus 1990-2021.

Figura 4: Fases de evolución tras sufrir un ictus.

2.2 Índice de tablas

Tabla 1: Criterios Terwee y Prinsen.

Tabla 2: Método GRADE modificado.

Tabla 3: Definición de los niveles de calidad de la evidencia de GRADE.

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

3. ESTRUCTURA GENERAL DE LA TESIS

Postulamos esta tesis doctoral como una tesis monográfica, avalada por la publicación de uno de los artículos que conforman el cuerpo de la tesis (de Blas-Zamorano et al., 2025).

En el marco teórico se abordan diferentes puntos cuya importancia radica en la evaluación de personas con ictus y cuyo objetivo será actuar de enlace con el marco empírico. En este punto se busca reforzar las investigaciones realizadas para la tesis doctoral y que el marco teórico actúe como nexo entre la bibliografía actual y lo investigado por el doctorando.

Los objetivos presentados en el cuerpo de esta tesis doctoral pretenden dar forma a las inquietudes personales transformadas en investigación que han dado como resultado la realización de cuatro artículos científicos.

En el marco empírico se encuentran los artículos científicos que constituyen el cuerpo principal de la tesis. Para un mejor entendimiento han sido divididos en dos subgrupos: investigación de variables de resultados, donde se engloban dos de las cuatro investigaciones de esta tesis doctoral. En este caso, las investigaciones estarán categorizadas como revisiones sistemáticas psicométricas. La primera revisión sistemática psicométrica se plantea con el objetivo de identificar las diferentes versiones de la escala Fugl-Meyer Assessment (FMA) y seleccionar la más recomendada empleando la metodología COSMIN. La segunda revisión sistemática psicométrica tiene como objetivo detectar las diferentes versiones del Índice de Barthel (BI) y de la Medida de Independencia Funcional/Functional Independence Measure (FIM) y seleccionar la más recomendada según los criterios COSMIN.

El segundo apartado, investigación clínica, está conformado por otras dos investigaciones realizadas en el ámbito clínico. La primera consiste en la revisión y actualización de la escala Nottingham Stroke Dressing Scale (NSDA), y la posterior traducción, adaptación transcultural y validación de la versión española (Escala Nottingham de Vestido en Ictus-Revisada (ENVI-R)), la cual consiste en una escala de evaluación de la actividad del vestido en personas que han sufrido un ictus. La segunda investigación de este segundo apartado consiste en la validación de la versión española de la escala FMA en población española que había sufrido un ictus.

Esta investigación, al darse en ámbito clínico, avala que la tesis doctoral se presente como una tesis monográfica y al realizar análisis, validaciones y mediciones con propiedades psicométricas de las escalas se refuerza que, de las líneas de investigación propuestas por la Universidad de Málaga, esta tesis doctoral se englobe dentro de la “Línea 3: Clinimetría y Fisioterapia”.

Para finalizar se discutirán los apartados más importantes del marco empírico con la bibliografía existente en el apartado discusiones y se realizará una

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

síntesis de los principales puntos destacables de esta tesis doctoral en el apartado conclusiones.

4. MARCO TEÓRICO

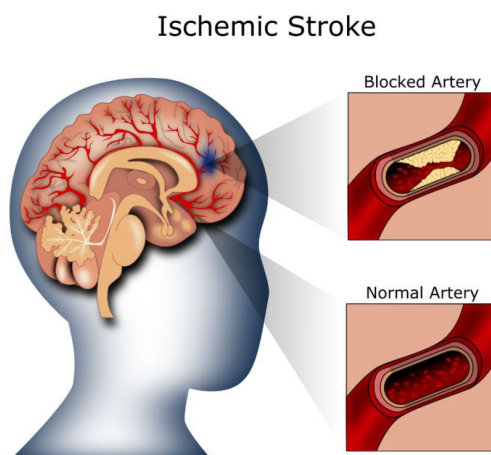
1. Ictus

1.1 Definición

En 1970, la Organización Mundial de la Salud (OMS) describió el ictus como “desarrollo rápido de signos clínicos de alteración focal (o global) de la función cerebral, con una duración de más de 24 horas o que provoca la muerte, sin causa aparente que no sea de origen vascular” (Aho et al., 1980). Posteriormente, esta definición fue revisada por la American Heart Association y la American Stroke Association, las cuales, teniendo en cuenta los avances científicos, lo redefinieron como “déficit neurológico atribuido a una lesión focal aguda del sistema nervioso central (incluyendo lesión cerebral, medular o retiniana) por una causa vascular, incluyendo el infarto cerebral, hemorragia intracerebral y hemorragia subaracnoidea cuyos síntomas persisten más de 24 horas o hasta la muerte” (Sacco et al., 2013).

El ictus se define como una disfunción neurológica aguda originada por una alteración en la perfusión cerebral secundaria a un evento vascular. Comúnmente denominado ictus, del latín *ictus* (“golpe”), hace referencia a su inicio súbito y de presentación violenta. En la literatura anglosajona, este evento es identificado como *stroke* (Careaga, 2020; Feigin et al., 2023).

Figura 1. Ictus.



Fuente: Rajendra Acharya, U., Meiburger, K. M., Faust, O., En Wei Koh, J., Lih Oh, S., Ciaccio, E. J., Subudhi, A., Jahmunah, V., & Sabut, S. (2019). Automatic detection of ischemic stroke using higher order spectra features in brain MRI images. *Cognitive Systems Research*, 58, 134-142. <https://doi.org/10.1016/j.cogsys.2019.05.005>

Las patologías cardiovasculares, principalmente el ictus, representan una de las principales cargas para los sistemas de salud pública a nivel global. En

particular, el ictus constituye la primera causa de discapacidad adquirida en adultos, dado que las secuelas motoras y funcionales resultantes generan limitaciones significativas en las Actividades de la Vida diaria (AVD) y restricciones en la participación social (de Sousa et al., 2018; Feigin et al., 2023).

Entre los déficits funcionales más prevalentes tras un ictus se encuentran la marcha alterada y la disfunción del miembro superior, particularmente en actividades de autocuidado (Medis, 2013). Aproximadamente el 50% de los supervivientes experimentan hemiparesia o hemiplejía, lo que conlleva desequilibrios posturales y asimetría corporal (Dickstein et al., 2004; Tovar et al., 2021). Asimismo, en aproximadamente el 80% de los casos, se reporta afectación del miembro superior, lo que compromete aún más la funcionalidad (Ingram et al., 2023; Rathore et al., 2002; Soleimani et al., 2024).

Estas alteraciones neuromusculares y funcionales tienen un impacto directo sobre la calidad de vida, limitando la independencia y la participación activa en el entorno social y laboral (de Diego Alonso, 2024; Winstein et al., 2016).

1.2 Tipos

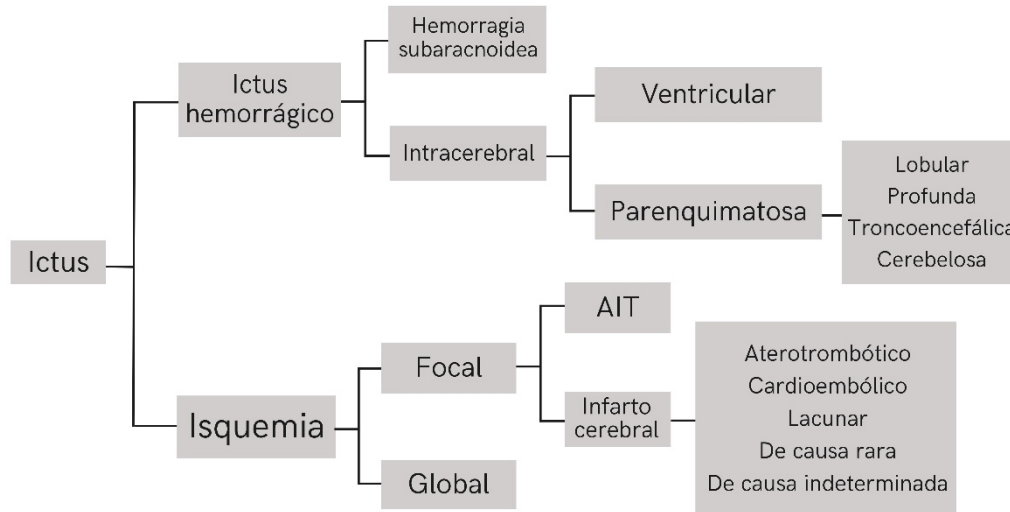
Los ictus se dividen en dos grandes grupos dependiendo de cómo se produce: isquemia cerebral y hemorragia intracraneal.

Los ictus isquémicos suponen entre el 80 y el 85% de todos los ictus. Los hemorrágicos suponen el 15-20% del total (Lui et al., 2025).

Como se muestra en la Figura 2, dentro de la isquemia cerebral hay dos tipos: isquemia cerebral focal, afectando únicamente a una zona del encéfalo y, por otro lado, la isquemia cerebral global, la cual afecta al encéfalo de manera extensa y difusa (Lui et al., 2025).

Dentro de los ictus de tipo hemorrágico, se pueden diferenciar en dos tipos: los hematomas cerebrales y la hemorragia subaracnoidea (Unnithan et al., 2025) (ver Figura 2).

Figura 2. Clasificación de los ictus según el tipo.



Fuente: Elaboración propia a partir de la adaptación de Díez Tejedor, E. (2006). *Guía para el diagnóstico y tratamiento del ictus*. Prous Science Sociedad Española de Neurología.

Se diferencian dos tipos de isquemia cerebral focal en función de la duración del proceso de isquemia: el ataque isquémico transitorio y el infarto cerebral (Díez-Tejedor, 2006; Lui et al., 2025)

El ataque isquémico transitorio se trata de un periodo de isquemia cerebral focal o monocular que dura menos de 24 horas; sin embargo, el infarto cerebral genera un déficit neurológico que se mantiene durante más de 24 horas, produciendo necrosis del tejido tisular.

Dependiendo de la causa, hay varios tipos de infarto cerebral: cardioembólico, aterotrombótico, lacunar, de causa inhabitual y de origen indeterminado. Para conocer la causa se realizan pruebas de exploración física, neuroimagen (Tomografía Computarizada (TC) o Resonancia Magnética (RM)), estudios neurovasculares, cardiológicos y de hemostasia.

En la isquemia cerebral global se da un descenso del flujo sanguíneo en todo el encéfalo, afectando a ambos hemisferios de manera difusa; pudiendo o no, dañarse el tronco del encéfalo y el cerebelo. Esto puede provocar clínicamente síndromes cerebrales focales, déficits cognitivos, estado vegetativo persistente e incluso la muerte (Díez-Tejedor et al., 2001; Lui et al., 2025).

Los hematomas y las hemorragias cerebrales son derrames de sangre hacia el parénquima cerebral. La causa más frecuente de este suceso es la hipertensión arterial; aunque no únicamente, ya que puede deberse también a la angiopatía

amiloide, malformaciones vasculares, fármacos, tóxicos y tumores, entre otros (Díez-Tejedor, 2006; Unnithan et al., 2025).

Dependiendo de la localización se puede clasificar la hemorragia parenquimatosa en: lobar, profunda (ganglios basales, subtalámica), troncoencefálica y cerebelosa (Unnithan et al., 2025; Ustrell et al., 2007).

Según la localización y el volumen de la extravasación de sangre, puede causar diferentes síntomas, aunque el más frecuente es la disminución de consciencia y la cefalea intensa. Los síntomas de una hemorragia cerebral son difíciles de diferenciar de los síntomas de los ictus isquémicos.

Para diagnosticar tanto hematomas como hemorragias cerebrales se emplea el TC craneal.

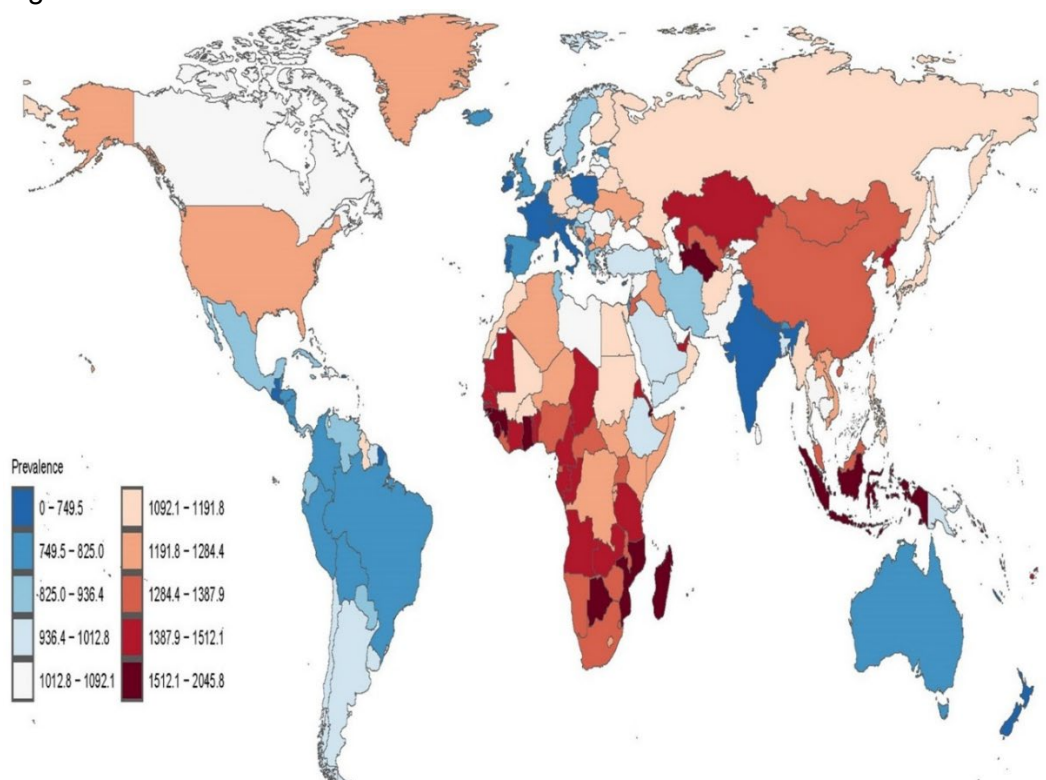
La hemorragia subaracnoidea (HSA) es el sangrado directamente al espacio subaracnoideo. La etiología más frecuente es la rotura de un aneurisma (dilatación anormal de las paredes de una arteria o una vena). Normalmente provoca síntomas como cefalea intensa, signos meníngeos y alteración del nivel de consciencia (SEN, 2019).

Al igual que con la hemorragia parenquimatosa, la hemorragia subaracnoidea también se diagnostica a través del TC craneal.

1.3 Epidemiología: incidencia y prevalencia

En el año 2019, la prevalencia global del ictus alcanzó los 101 millones de casos, con un total de 6,55 millones de muertes atribuidas a esta patología. La incidencia global fue de 12,2 millones de nuevos casos, de los cuales el 62,4% correspondieron a ictus isquémicos, el 27,9% a hemorragias intracerebrales, y el 9,7% a hemorragias subaracnoideas (GBD, 2021) (ver Figura 3).

Figura 3. Prevalencia mundial del ictus 1990-2021.



Fuente: He, Q., Wang, W., Zhang, Y., Xiong, Y., Tao, C., Ma, L., Ma, J., You, C., & Wang, C. (2024). Global, Regional, and National Burden of Stroke, 1990–2021: A Systematic Analysis for Global Burden of Disease 2021. *Stroke*, 55(12), 2815-2824. <https://doi.org/10.1161/strokeaha.124.048033>

A pesar de que los años de vida ajustados por discapacidad debidos al ictus fueron mayores en hombres (76,9 millones) que en mujeres (66,4 millones), la prevalencia e incidencia específicas fueron ligeramente superiores en la población femenina. No se observaron diferencias significativas entre sexos en cuanto a la mortalidad por ictus (GBD, 2021).

En el ámbito nacional, según el informe de la Sociedad Española de Neurología (SEN) (SEN, 2019), la prevalencia estimada de ictus fue del 6,4% en personas mayores de 70 años, con una mayor frecuencia en varones (7,3%) en comparación con mujeres (5,6%), y con una distribución más elevada en zonas urbanas (8,7%) frente a áreas rurales (3,8%) (SEN, 2019). Por su parte, el estudio Iberictus estimó una incidencia de 187,4 casos por cada 100.000 habitantes (Díaz-Guzman et al., 2012).

Es relevante señalar que el número total de casos de ictus a nivel mundial aumentó un 70% entre 1990 y 2019 (GBD, 2021). Se proyecta además un incremento adicional del 35% en la incidencia entre 2015 y 2035, en gran parte debido al envejecimiento progresivo de la población (SEN, 2019). De acuerdo

con las proyecciones de la OMS, una de cada cuatro personas mayores de 25 años sufrirá un ictus a lo largo de su vida (GBD, 2021; GBD et al., 2018).

Según apunta la SEN (2019), en España durante el año 2017 surgieron 71.780 casos nuevos; y se produjeron 26.937 defunciones por ictus. Con una estimación de una prevalencia de 1,7%, lo que supone un total de 661.512 casos.

En el contexto europeo, se estimó que, en el año 2025, cerca de 1,5 millones de personas sufrirían un ictus, y se prevé que esta cifra siga en aumento durante las próximas tres décadas, impulsada por un incremento del 27% en la supervivencia post-ictus, lo que conllevará una mayor carga asistencial y social (Wafa et al., 2020). Se estima que la cantidad de ictus aumentará un 50% en los próximos años, pasando de 6,6 millones de casos en 2020 a 9,7 millones en 2050 (Feigin et al., 2023).

Actualmente, el ictus se posiciona como la segunda causa de muerte en países desarrollados, solo por detrás de la cardiopatía isquémica (GBD, 2021; WHO, 2016). Afecta principalmente a personas mayores de 65 años (GBD, 2021; Virani et al., 2020). En España, en 2018, fue la segunda causa de muerte global, la tercera entre hombres y la primera entre mujeres (SEN, 2019). La tasa de mortalidad por ictus ajustada por edad fue de 25,1 por cada 100.000 habitantes, con cifras más altas en hombres (28,7) que en mujeres (21,9) y un total de aproximadamente 27.000 fallecimientos anuales por esta causa en el país (Bernhardt et al., 2017; SEN, 2019).

1.4 Factores de riesgo

En el año 2019, el estudio de The Global Burden of Disease (GBD) identificó un total de 19 factores de riesgo responsables del 87% de la discapacidad atribuible al ictus. Entre ellos, los cinco con mayor impacto específico sobre la mortalidad y la discapacidad asociadas a esta patología fueron: presión arterial sistólica elevada (55,5%), índice de masa corporal elevado (24,3%), hiperglucemia en ayunas (20,2%), exposición a contaminación ambiental por partículas (20,1%) y consumo de tabaco (17,6%) (GBD, 2021).

Según autores como Berenguer et al. (2016); Cabrera et al. (2014), o Lira (2015), los factores de riesgo pueden clasificarse como modificables, potencialmente modificables y no modificables:

a) Modificables: hipertensión arterial, cardiopatías (fibrilación auricular, endocarditis infecciosa, estenosis mitral, infarto de miocardio reciente), tabaquismo, anemia de células falciformes, accidentes isquémicos transitorios (AIT) previos y estenosis carotídea asintomática.

b) Potencialmente modificables: diabetes mellitus, hipertrofia ventricular, cardiopatías (miocardiopatía, discinesia de la pared ventricular, endocarditis no bacteriana, calcificación del anillo mitral, prolapso mitral, contraste ecocardiográfico espontáneo, estenosis aórtica, foramen oval permanente, aneurisma del septo atrial), uso de anticonceptivos orales, consumo excesivo de alcohol, consumo de drogas, sedentarismo, obesidad, factores dietéticos, hematocrito elevado, hiperinsulinemia/resistencia a la insulina y desencadenantes agudos (estrés, migraña, estados de hipercoagulabilidad e inflamación, enfermedad subclínica, engrosamiento íntima-media).

c) No modificables: edad, sexo, factores hereditarios, raza/etnia y localización geográfica.

La hipertensión arterial es el factor de riesgo con mayor prevalencia y con gran potencial de riesgo para el ictus, tanto isquémicos como hemorrágicos (Lira, 2015; Ministerio de Sanidad y Política Social, 2024).

Factores de riesgo en supervivientes de ictus

La recurrencia del ictus representa una preocupación clínica y epidemiológica significativa, con un riesgo de segundo evento del 30% en los supervivientes (Rand et al., 2010). La incidencia acumulada se estima en uno de cada seis casos durante el primer año post-ictus, y en uno de cada tres al transcurrir los primeros tres años (Bergstrom et al., 2017; Kolmos et al., 2021). Afortunadamente, se proyecta que esta tendencia ascendente en la incidencia de recurrencia del ictus será mitigada a lo largo de este siglo gracias a los avances en la prevención secundaria (Bergstrom et al., 2017; Kolmos et al., 2021) y al creciente impulso de la investigación en el ámbito del ictus.

La gestión de factores de riesgo modificables, como el tabaquismo, la obesidad, la diabetes tipo 2, la hipertensión, la hipercolesterolemia, la inactividad física y los patrones dietéticos, es fundamental (Bailey et al., 2019). Estos factores no solo influyen en la prevención primaria del ictus, sino que también desempeñan un papel crucial en la prevención secundaria (Bergstrom et al., 2017; Kolmos et al., 2021), a través de la adopción de estilos de vida saludables (Goldstein et al., 2011; Chiuve et al., 2008). Es relevante señalar que investigaciones recientes sugieren una posible heterogeneidad en la prevalencia e impacto de estos factores de riesgo modificables entre poblaciones rurales y urbanas (Howard, 2021; Li et al., 2019; Nakibuuka et al., 2015; Sridharan et al., 2009).

En este contexto, la actividad física emerge como una intervención particularmente eficaz, documentándose una reducción del 25% en el riesgo de ictus (de Diego Alonso, 2024; Lee et al., 2003). Además, la actividad física ejerce un efecto beneficioso integral sobre la totalidad de los factores de riesgo

modificables asociados a la recurrencia del ictus (Billinger et al., 2014; Boyne et al., 2017).

1.5 Consecuencias

Los supervivientes de un ictus experimentan diversas secuelas, incluidas limitaciones sensoriomotoras, cognitivas y emocionales entre otras, que afectan significativamente su autonomía e independencia (Sacco et al., 2013; Virani et al., 2020).

Tras dos semanas de haber sufrido un ictus, más del 60% de los supervivientes es dependiente para llevar a cabo AVDs, entre el 70% y el 80% presenta hemiparesia, entre el 70-80% presenta problemas de deambulación, el 60-70% déficit visual, el 55% disartria, el 40% sufre depresión, alrededor del 20-30% presenta afasia, entre el 15% y el 35% disfagia (Camoës-Barbosa et al., 2012; Espárrago Llorca et al., 2015; French et al., 2010; Hilari, 2011; Morone et al., 2015). Y, aproximadamente dos tercios de estos individuos presentan dificultades persistentes en la deambulación independiente incluso 3 meses después del ictus (Kennedy et al., 2021).

1.5.1 Alteraciones motoras

El ictus ocasiona una serie de déficits motores. A continuación, se detallan las habilidades motoras comprometidas y las manifestaciones clínicas más frecuentes.

Alteraciones de la fuerza (paresia/debilidad): La paresia o debilidad se manifiesta como la incapacidad para modular o reclutar adecuadamente las unidades motoras, constituyendo una alteración característica de la motoneurona superior (Gracies, 2005a, 2005b). El fenotipo de este signo es altamente variable, pudiendo presentarse de forma unilateral o bilateral, con diferente grado de afectación o distribución regional (Andrews & Bohannon, 2000). La debilidad es el signo más prevalente post-ictus y, comúnmente, se localiza en el hemicuerpo contralateral a la lesión cerebral, resultando en hemiparesia o, en casos más severos, hemiplejía (Broeks et al., 1999; Page et al., 2001; Thielbar et al., 2017).

Alteraciones del tono muscular: Un signo patológico frecuente es la espasticidad o hiperresistencia neural, definida como un trastorno motor caracterizado por un incremento, dependiente de la velocidad, de los reflejos miotáticos tónicos, acompañado de una exacerbación de los espasmos tendinosos, como consecuencia de la hiperexcitabilidad del reflejo miotático (Lance & Gail, 1965; García-Díez, 2004). La espasticidad se asocia con un tono muscular elevado en la musculatura afectada (hipertonía). También puede presentarse una

disminución del tono muscular, denominada hipotonía. Ambas alteraciones del tono son una característica común en la población post-ictus.

Alteraciones del movimiento: La afectación de estructuras subcorticales tras un ictus puede generar una sintomatología motora muy diversa. La localización de la lesión es un determinante crítico del perfil sintomático. En relación con las alteraciones del movimiento, pueden observarse trastornos hipercinéticos o hipocinéticos; adicionalmente, puede manifestarse una pérdida de la capacidad para producir movimientos aislados y selectivos, resultando en la aparición de movimientos masivos o patrones de movimiento en masa, conocidos como sinergias anómalas. Asimismo, la afectación de la coordinación motora puede conducir a movimientos imprecisos o atáxicos.

Las alteraciones motoras, con la debilidad en el miembro superior destacando como la más común, son las que mayor impacto tienen en el desempeño ocupacional de las personas post-ictus (Canning et al., 2004; Gurari et al., 2018; Wagner et al., 2007). Estos déficits comprometen significativamente la participación de la persona (Carey et al., 2018) y su calidad de vida, dado que la limitada habilidad motora impide la ejecución de numerosas actividades cotidianas.

1.5.2 Alteraciones sensitivas

Aproximadamente la mitad de los individuos que sufren un ictus experimentan una disminución en la percepción de las sensaciones corporales, un fenómeno conocido como alteraciones sensitivas o somatosensoriales (Bernal Jiménez, 2024).

Las modalidades sensoriales que pueden verse comprometidas tras un ictus pueden ser: tacto ligero, localización táctil, discriminación entre dos puntos, nocicepción, sensibilidad térmica, estereognosis, propiocepción y cinestesia.

La alteración en la discriminación de dos puntos se correlaciona con una recuperación deficiente de la destreza de los miembros superiores (Au-Yeung & Hui-Chan, 2009).

La estereognosis es la modalidad sensorial con mayor valor predictivo para la destreza manual, dado que su preservación se correlaciona positivamente con mejoras en el test de Jebsen & Taylor y, por ende, en la destreza manual (Cuesta-García, 2016).

La alteración propioceptiva se asocia con una peor recuperación motora de los miembros superiores (Desrosiers et al., 2002; Markus, 2022).

Las lesiones cerebrales resultantes de un ictus pueden llevar a una interpretación anómala de los estímulos sensitivos, dando lugar a diversas manifestaciones clínicas, como la hipoestesia, hiperestesia, hipoalgesia,

hiperalgesia, parestesia, disestesia, anestesia, analgesia y alodinia (Ropper et al., 2017).

Estas alteraciones sensitivas son prevalentes tras un ictus, destacando las dificultades en la discriminación de texturas y el reconocimiento de objetos mediante el tacto (Carey, 1995; Carey, 2012), así como las modificaciones en la percepción del dolor (Klit et al., 2011), entre otras. Existe una sólida correlación entre las alteraciones sensoriales y la recuperación motora después de una lesión del sistema nervioso central (Polonio et al., 2010), prediciendo un resultado funcional negativo post-ictus, impactando adversamente la movilidad, la ejecución de las AVD y el desempeño ocupacional general (Pollock et al., 2014; Tyson et al., 2008).

1.5.3 Alteraciones cognitivas

Las funciones cognitivas más frecuente afectadas tras el ictus son la atención, memoria, funciones ejecutivas, lenguaje y praxias (Cramer, 2008; Peña-Casanova, 2014; Redolar Ripoll, 2023). Las alteraciones más comunes incluyen:

- Inatención: Se refiere a una deficiencia generalizada en la orientación o procesamiento de estímulos. La alteración atencional más frecuente es la heminegligencia, que se define como la dificultad para orientarse, actuar o responder a estímulos o acciones que ocurren en el hemisferio contralateral a la lesión cortical, no atribuible a déficits sensoriales o motores primarios (Allegri, 2000).
- Amnesia: Pérdida de memoria, definida como la incapacidad parcial o total para recordar experiencias pasadas (amnesia retrógrada) o la imposibilidad de formar nuevos recuerdos (amnesia anterógrada). En ocasiones, las personas con ictus no presentan amnesia per se, pero sí dificultades para recuperar información almacenada y aplicarla en las AVD.
- Síndrome disejecutivo: Además de las alteraciones en las funciones ejecutivas propiamente dichas, se asocian síntomas específicos, como la anosognosia, perseveración, incapacidad para tomar decisiones, etc. (Bernal Jimenez, 2024)
- Alteraciones del lenguaje: Implican dificultades tanto en la expresión como en la comprensión del lenguaje, es decir, afasia. También se da la anomia (pérdida de la capacidad de nombrar personas o cosas), alexia (incapacidad para leer) y agafia (incapacidad para escribir). Los problemas de comunicación pueden estar provocados por secuelas físicas, como en la disartria (lentitud o incluso parálisis de los músculos encargados del habla) o las disfonías (problemas para emitir la voz) (Murciego et al. 2019).

- Apraxia: Déficit en la ejecución de un movimiento aprendido que no puede explicarse por incoordinación, pérdida sensorial, falta de comprensión o inatención. Los dos tipos principales de apraxia son (Foundas & Duncan, 2019):
 - Apraxia ideatoria: Incapacidad para ejecutar una secuencia de gestos propositivos dirigidos a un fin, debido a la pérdida del plan de acción subyacente. El individuo es incapaz de utilizar objetos reales de manera apropiada.
 - Apraxia ideomotora: Caracterizada por la incapacidad para realizar un acto motor previamente aprendido, pudiendo manifestarse errores en la generación del plan de acción, la secuenciación y organización de los elementos apropiados, o en la ejecución del plan. La persona es incapaz de realizar un acto motor tras una orden verbal o por imitación.

Las alteraciones de las habilidades cognitivas y su persistencia en las etapas crónicas del ictus implican un pronóstico desfavorable para la recuperación, tanto de las propias habilidades cognitivas como en relación con su influencia sobre el desempeño ocupacional y la participación del sujeto (Lugtmeijer et al., 2021; Revet et al., 2021; Spaccavento et al., 2019). Estas disfunciones cognitivas impactan negativamente la funcionalidad motora del miembro superior, reduciendo su eficacia en el movimiento y el rendimiento en las actividades, lo que a su vez afecta profundamente la ejecución de las AVD y el desempeño ocupacional de la persona superviviente del ictus.

1.5.4 Alteraciones perceptivas

La información sensorial, una vez captada por los receptores periféricos, es transmitida al sistema nervioso central para su interpretación y procesamiento, lo que permite la atribución de significado a los estímulos y así, posteriormente, poder interactuar con el entorno (Bernal Jimenez et al., 2024).

El ictus causa alteraciones sobre estos sistemas perceptivos, causando limitaciones en la percepción del color, percepción de la profundidad, percepción figura-fondo, en el esquema corporal, en la orientación espacial y en la discriminación de derecha-izquierda, entre otras.

Uno de los principales déficits perceptivos son las agnosias, que se conceptualizan como alteraciones en el reconocimiento e identificación de estímulos, es decir, la incapacidad para comprender la información sensorial. Estas alteraciones pueden manifestarse independientemente de la presencia de déficits sensoriales o cognitivos. Según la modalidad sensorial comprometida, se distinguen agnosias visuales, auditivas y táctiles. Incluso dentro de una misma

modalidad sensorial, la afectación puede ser específica para categorías de estímulos concretas, como colores, formas o rostros (Polonio et al., 2010).

1.5.5 Alteraciones emocionales

Tras un ictus, la pérdida de la función motora y su subsiguiente impacto en el desempeño ocupacional y la calidad de vida de la persona a menudo precipitan la aparición de alteraciones emocionales, entre las que se incluyen ansiedad, depresión, labilidad emocional, desinhibición, apatía, entre otras (de Noreña et al., 2010).

De las diversas alteraciones emocionales que pueden surgir post-ictus, la depresión y la ansiedad son particularmente destacadas por su prevalencia e impacto (Delgado Losada, 2019; Rafsten et al., 2018; Shi et al., 2017).

- **Ansiedad:** Se define como una respuesta emocional o un patrón de respuestas que abarca componentes cognitivos (anhedonia, tensión, aprensión), alteraciones fisiológicas caracterizadas por una elevada activación del sistema nervioso autónomo, y alteraciones motoras que suelen manifestarse como comportamientos desajustados y escasamente adaptativos. En el contexto post-ictus, esta ansiedad es frecuentemente inducida por la incertidumbre de la evolución clínica, el riesgo de recurrencia, y las características del afrontamiento familiar. La presencia de ansiedad puede influir negativamente en el proceso de recuperación del ictus.
- **Depresión:** Caracterizada por un estado de ánimo de profunda tristeza o desinterés generalizado por la mayoría de las actividades, acompañado de una marcada disminución de la actividad física y psicológica. Este trastorno es comúnmente una consecuencia de la incapacidad para ejecutar AVD y actividades significativas, y de la adaptación al nuevo estado de salud. La depresión es la complicación psiquiátrica más común tras un ictus (Siotto et al., 2021; Yi et al., 2020).

El ictus ejerce un elevado impacto emocional, tanto en el individuo afectado como en su entorno familiar, comprometiendo negativamente el funcionamiento personal y familiar, ambos pilares fundamentales en el proceso de rehabilitación. La presencia de ansiedad en personass con ictus empeora la ejecución de las AVD y el funcionamiento social del individuo (Leno Díaz et al., 2016).

Adicionalmente, la depresión constituye otra alteración que deteriora la evolución general del ictus, influyendo negativamente en la ejecución de las AVD y en la calidad de vida de las personas (Herrera et al., 2020). Cabe destacar que estas alteraciones emocionales también han sido asociadas con una mayor tasa de mortalidad en esta población (Cai et al., 2019; Khan et al., 2023).

1.5.6 Discapacidad e impacto

El conjunto de las secuelas anteriores, causan dificultades en las personas que han sufrido un ictus para desempeñar las AVD de manera eficaz, tanto básicas (alimentación, vestido, ducha, aseo, deambulación, etc.) como instrumentales (uso del teléfono, uso del transporte público, lavado de ropa, cocinar, manejo del dinero, etc.), por lo tanto, hacen que la persona que ha sufrido el ictus, sea dependiente de los demás para hacer tareas cotidianas del día a día (Castillo et al., 2013; Hoyas et al., 2015; Ingram et al., 2023; Li et al., 2023; López, 2010; Soleimani et al., 2024; Zhao et al., 2023). A los seis meses del ictus, más del 60% de las personas post-ictus presentarán una mano no funcional para las AVDs, y entre el 20% y el 25% no podrán deambular sin asistencia (Arias, 2009).

Todas estas limitaciones afectan considerablemente la capacidad para realizar AVD y restringen la participación (Norlander et al., 2016) en áreas como el autocuidado, el trabajo, el ocio y las actividades sociales (Ezekiel et al., 2019; Lee et al., 2023; Martín-Sáez et al., 2021; Norlander et al., 2016; Singam et al., 2015; Svensson et al., 2019). El retorno a una participación plena en estas esferas puede requerir más de 6 meses, extendiéndose en algunos casos a varios años (Della Vecchia et al., 2021).

Consecuentemente, el ictus se posiciona como la tercera causa de discapacidad a nivel global (Feigin et al., 2022), con una prevalencia de 143 millones de personas, de las cuales 512.380 residen en España. En 2019, la incidencia anual de supervivientes de ictus con algún grado de discapacidad en España alcanzó las 61.102 personas (GBD, 2021).

Aproximadamente el 50% de los supervivientes a un ictus experimentan secuelas discapacitantes. Además, se prevee un incremento del 35% en la incidencia de ictus en España para 2035 (SEN, 2019).

Las alteraciones causadas inducidas por el ictus conllevan una discapacidad de gravedad variable. La etiología de la discapacidad post-ictus es multifactorial: el 86% de los afectados presenta alteraciones motoras, el 39% disfunciones del lenguaje y el 34% déficits cognitivos que impiden el aprendizaje. La naturaleza multidimensional de estas alteraciones conduce a una discapacidad que impacta predominantemente la movilidad, la autonomía en la vida doméstica y el autocuidado (Noé-Sebastián et al., 2017).

Estas alteraciones discapacitantes repercuten en el desempeño ocupacional, la participación comunitaria y la calidad de vida (AOTA, 2020).

Las secuelas del ictus ejercen un impacto considerable en el desempeño ocupacional y en el desarrollo de roles ocupacionales (Buzzelli et al., 2023). La presencia de déficits motores, en particular, predice la participación de los supervivientes en la comunidad de manera inversamente proporcional: a mayor

severidad de las alteraciones motoras, menor nivel de participación (Ribeiro De Souza et al., 2023).

Los déficits tras el ictus ejercen una influencia negativa en la calidad de vida, destacando las alteraciones motoras, ya que estas impiden el desempeño de las AVD y la participación en actividades significativas (Bhimani & Anderson, 2014).

La disminución de la calidad de vida en estas personas trasciende las implicaciones directas de los déficits y la discapacidad (Bernal Jiménez, 2024; Leno et al., 2016). Las alteraciones emocionales, como el temor a un nuevo episodio cerebrovascular y la ansiedad ante la nueva situación funcional, exacerbaban aún más el deterioro de la calidad de vida (Leno et al., 2016). Por lo tanto, el logro de este estado de bienestar debe constituir un objetivo primordial en el proceso rehabilitador (Bernal Jiménez, 2024).

1.5.7 Gasto socioeconómico y sanitario

El ictus representa una patología con una elevada complejidad económica, fundamentalmente debido a las secuelas funcionales que ocasiona a los supervivientes (Rochmah et al., 2021). Estas deficiencias perjudican la calidad de vida de las personas que han sufrido un ictus (Dawson et al., 2023), además de tener un enorme impacto social y económico, con un alto coste humano y material, tanto intrahospitalario como extrahospitalario (Dobkin, 1995; Evers et al., 2004; Roschmah et al., 2021). Recuperar la independencia tras un ictus depende en gran medida de la recuperación motora, en particular de la recuperación de la actividad de las extremidades superiores (Shanks et al., 2025). La gestión de estas secuelas demanda un tratamiento prolongado, el cual, consume aproximadamente el 34% del gasto sanitario total a nivel mundial (Rochmah et al., 2021). En España, el coste total asociado al ictus fue de 2.908 millones de euros en 2015. Específicamente, en 2017, 623 millones de euros se destinaron al coste de 93.022 hospitalizaciones por ictus. No obstante, la mayor parte de los costes reales se materializa post-hospitalización, por los cuidados informales, cuyo gasto recae principalmente sobre las familias. En consonancia con el informe estadístico de 2015 de The European Heart Network sobre enfermedades cardiovasculares, los cuidados informales y las pérdidas de productividad laboral (debidas a mortalidad o discapacidad) suponían 35€ al día y por ciudadano (Al-Senani et al., 2020; Ben Shoham et al., 2019; de Diego Alonso, 2024; GBD, 2021; Pacheco-Barrios et al., 2022; Purroy et al., 2019; Wafa et al., 2020; Yao et al., 2023).

Se prevé un aumento significativo en el número global de supervivientes de ictus entre 2040 y 2050 (Al-Senani et al., 2020; Ben Shoham et al., 2019; de Diego Alonso, 2024; GBD, 2021; Pacheco-Barrios et al., 2022; Purroy et al., 2019; Wafa et al., 2020; Yao et al., 2023). Este incremento implicará una mayor prevalencia de discapacidad, lo que a su vez elevará los costes sanitarios y

socioeconómicos asociados (GBD, 2021; Rochmah et al., 2021; Strilcivuc et al., 2021). Dentro de este marco de gastos, la Organización Mundial de la Salud (OMS) ha subrayado la necesidad de contemplar la neurorrehabilitación, tanto en fases agudas como a largo plazo (Cieza et al., 2021).

1.6 Fases de evolución clínica y pronóstico

1.6.1 Fases tras el ictus

Tras sufrir un ictus, las semanas y los meses siguientes son claves para la recuperación de la persona de las secuelas adquiridas. Esto se debe a la habilidad que posee el sistema nervioso para adaptarse y responder a las variaciones del entorno, gracias a la plasticidad cerebral. La plasticidad cerebral se define como la capacidad del sistema nervioso para reorganizar sus estructuras, funciones y conexiones neuronales en respuesta a diversos estímulos y experiencias (Purves et al., 2020).

Para maximizar el potencial neuroplástico en la rehabilitación de la persona, esta se divide en las siguientes fases (Durá-Mata et al., 2011; Murié-Ferández et al., 2010) (ver Figura 4):

Fase Aguda (Primeras 24 horas post-ictus)

Durante la fase aguda, las prioridades se centran en el soporte vital y la implementación de medidas preventivas tempranas. Estas últimas incluyen:

- Correcta alineación postural: Fundamental para prevenir deformidades y facilitar la circulación.
- Movilizaciones pasivas: Aseguran el rango de movimiento articular y previenen la rigidez.
- Cambios posturales regulares: Esenciales para mitigar el riesgo de complicaciones cutáneas, como las úlceras por presión.

Fase Subaguda (24 horas a 6 meses post-ictus)

La fase subaguda marca la transición hacia un tratamiento rehabilitador más activo, que exige la participación de la persona en la medida de sus posibilidades. El enfoque ideal en este periodo, siempre que el estado de dicha persona lo permita, es un programa de rehabilitación intensiva.

La duración mínima de estas sesiones debe ser de 3 horas diarias, administradas durante 5 días a la semana. El diseño de la intervención es estrictamente individualizado, adaptándose a los déficits específicos de la persona y a su tolerancia al ejercicio. La duración total del tratamiento se

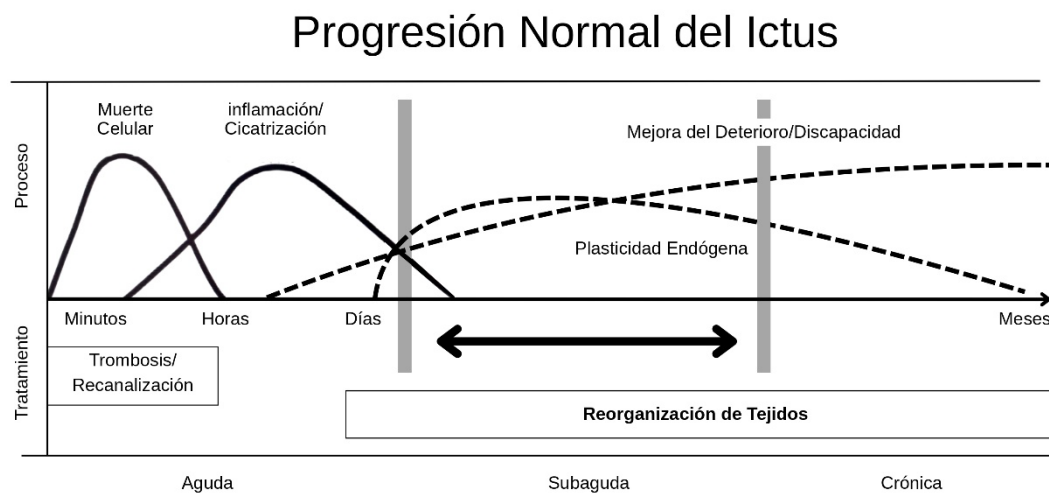
determinará por la comorbilidad existente, la severidad de los déficits neurológicos y la consecución de los objetivos terapéuticos establecidos.

Para personas con déficits leves, se puede optar por un programa de rehabilitación convencional. Este también es un tratamiento multidisciplinar (terapia ocupacional, fisioterapia, logopedia y/o neuropsicología), pero con una intervención diaria de menos de 3 horas y una frecuencia inferior a 5 días a la semana. Los principios de diseño y finalización de la intervención son los mismos que para el programa intensivo.

Fase Crónica (Más de 6 meses post-ictus)

La fase crónica se caracteriza por una estabilización de las mejoras, donde la progresión funcional suele ser más limitada. En este periodo, el objetivo principal es facilitar la integración de la persona en una vida activa acorde con su capacidad funcional residual. Es frecuente que esta fase coexista con la continuación de programas de rehabilitación convencional para mantener las ganancias y prevenir el deterioro.

Figura 4. Fases de evolución tras sufrir un ictus.



Fuente: Elaboración propia a partir de la traducción de Kwakkel et al. (2023).

1.6.2 Pronóstico tras el ictus

La planificación de los objetivos en la rehabilitación post-ictus requiere una estimación precisa del pronóstico funcional de la persona que ha sufrido el ictus

(Murié Fernández et al., 2010). La predicción temprana de la función a medio y largo plazo en individuos con ictus es fundamental para establecer una comunicación efectiva con la persona post-ictus y sus allegados, y para el diseño de objetivos de rehabilitación realistas (Murié Fernández et al., 2010). A pesar de que se han identificado más de 150 variables con valor pronóstico potencial, hasta la fecha no se dispone de un indicador que permita determinar la evolución con total certeza (Ferrer, 2016). Por consiguiente, solo es posible formular estimaciones con un grado variable de precisión. En este contexto, los porcentajes y las estimaciones derivadas de estos factores pronósticos deben interpretarse con cautela, ya que son orientativos. No obstante, resultan herramientas valiosas para informar a la persona y a sus familiares sobre las expectativas de la evolución clínica (Murié Fernández et al., 2010).

1.6.2.1 Factores pronósticos

Los factores pronósticos principales comprenden elementos agrupados en (Ferrer, 2016):

- No relacionados con la gravedad clínica del ictus: edad, etiología, hiperglucemia, hipertermia, entorno de la persona y apoyo social, situación funcional previa al ictus, depresión post-ictus
- Relacionadas con la gravedad clínica del ictus: severidad clínica inicial, retraso en la evolución de la recuperación, incontinencia urinaria, control de tronco en sedestación, déficit motor, percepción auditiva visual, nivel de discapacidad inicial, déficits cognitivos, localización y tamaño.

Factores pronósticos no relacionados con la gravedad del ictus:

La edad avanzada se asocia inversamente con la recuperación funcional post-ictus. Específicamente, una edad superior a 85 años (Ergeletzis et al., 2002) es considerada un punto de inflexión en el potencial de recuperación. No obstante, la literatura enfatiza que una respuesta reducida al tratamiento rehabilitador no implica ausencia de beneficio, por lo que la edad avanzada no debe constituir un criterio restrictivo para el acceso a la rehabilitación (Lieberman et al., 2005). Otros autores argumentan que la edad per se no es el factor pronóstico adverso, sino que su asociación con un peor pronóstico funcional se debe a la presencia de deterioro funcional previo, circunstancias socioeconómicas adversas y una mayor severidad del ictus en personas de edad avanzada (Macciochi et al., 1998).

La literatura no proporciona una evidencia robusta que diferencie la evolución clínica del ictus en función de sus mecanismos etiopatogénicos, pero si hay

diferencia en función de la severidad inicial del evento, el cual es un modulador del pronóstico. En este sentido, los ictus isquémicos relacionados con fibrilación auricular están asociados con una mayor mortalidad, tasas de recurrencia elevadas y un desenlace funcional menos favorable en comparación con personas sin esta arritmia (Lin et al., 1996). Esta diferencia en evolución funcional en personas con fibrilación auricular se debe a la mayor gravedad inicial del ictus (Jorgensen et al., 1996). En el caso de los ictus hemorrágicos, si bien se manifiestan con mayor gravedad clínica, tienden a una recuperación funcional mayor, aunque de forma más lenta, en comparación con los de origen isquémico por infarto (Ween et al., 1996).

La hiperglucemia al ingreso hospitalario se correlaciona con un incremento del riesgo de deterioro neurológico precoz tras el ictus (Naveiro et al., 1998).

La hipertermia en el momento de la admisión, incluso cuando es ligera, constituye un factor pronóstico desfavorable en la evolución funcional (Azzimondi et al., 1995).

La localización y el tamaño del ictus representan factores de mal pronóstico (Miyai et al., 1997).

El entorno familiar y el apoyo social son factores pronósticos cruciales en la recuperación post-ictus. Dada la dependencia temporal o permanente que los supervivientes pueden experimentar, la integración en un ambiente de apoyo es fundamental. Este vínculo no solo favorece la adherencia al tratamiento rehabilitador, sino que también optimiza el aprovechamiento de las intervenciones. Una red de apoyo social se correlaciona positivamente con el pronóstico funcional de la persona (Glass et al., 1993).

La situación funcional previa al ictus constituye un componente indispensable. Un deterioro funcional físico preexistente, ya sea secundario a ictus recurrentes o a comorbilidades discapacitantes, se ha identificado como un predictor independiente de la capacidad funcional post-ictus y de la probabilidad de institucionalización, especialmente en la población geriátrica (Pohjasvaara et al., 1997). La explicación de este pronóstico desfavorable en personas con antecedentes de ictus se debe al deterioro neurológico acumulado, a menudo subclínico, que estos eventos previos ocasionan (Kwakkel et al., 1996).

Depresión post-ictus: el estado de ánimo deprimido es considerado como un factor de mal pronóstico funcional (Siotto et al., 2021; Yi et al., 2020).

Factores pronósticos relacionados con la gravedad clínica del ictus:

El grado de severidad clínica inicial del ictus es un determinante fundamental de la evolución funcional en la fase subaguda.

La magnitud de la gravedad en la presentación clínica del ictus constituye el principal factor pronóstico funcional. Se observa que una menor puntuación inicial en la escala National Institute of Health Stroke Scale (NIHSS) se correlaciona directamente con un mayor grado de independencia funcional alcanzado al final del proceso (Ferrer, 2016).

Retraso en la evolución: La ausencia de mejoría funcional en los primeros días post-ictus, y de manera más consolidada al tercer mes, se considera un indicador de mal pronóstico funcional a largo plazo (Ferrer, 2016).

La incontinencia urinaria prolongada ha sido reconocida como un indicador de pronóstico funcional desfavorable, particularmente cuando se evalúa a los 12 meses de evolución (Barer et al., 1989).

La ausencia de recuperación del control de tronco en sedestación en el momento del alta hospitalaria no solo compromete la continuidad de la rehabilitación en el ámbito ambulatorio, sino que también se erige como una señal de alerta sobre una evolución desfavorable de la persona (Schorage et al., 2023).

La recuperación del déficit motor en personas con ictus está fundamentalmente ligada a la gravedad de la hemiplejía o hemiparesia inicial (Rochmah et al., 2021). Se ha documentado que solo el 6% de los supervivientes del ictus que presentan una parálisis inicial grave alcanzan una recuperación completa de la movilidad a los seis meses de evolución (Heinsius et al., 1998). En aquellas personas que exhiben un déficit motor grave y persistente a las tres semanas del ictus, se observa que la mayoría mantiene un grado de parálisis grave o moderada a los seis meses (Rochmah et al., 2021; Wade et al., 1987). Si se identifica una prensión voluntaria en la mano afectada, se puede anticipar razonablemente el retorno de una función rudimentaria al menos cinco meses después (Heller et al., 1987; Katrak et al., 1998). En esta línea, las personas que no logran recuperar una fuerza de prensión cuantificable con un dinamómetro en la mano afectada antes de los 24 días desde el ictus no alcanzan una fuerza de prensión completa a los tres meses (Sunderland et al., 1989).

En conclusión, el grado de recuperación del déficit motor actúa como una variable predictiva directa del nivel del objetivo funcional que se puede esperar alcanzar.

La percepción sensitiva y visual en personas con hemiparesia reviste una gran importancia en el proceso de reaprendizaje de habilidades perdidas tras el ictus (Ferrer, 2016).

En las personas con ictus, se diferencian en tres categorías según el patrón de déficit (Ferrer, 2016; Sánchez et al., 2000): 1) exclusivamente motor, 2) motor y sensitivo, y 3) sensitivo-motor con hemianopsia. La diferenciación de las personas post-ictus en estos grupos demostró que cada categoría alcanzaba niveles de funcionalidad diferenciados y con distintas temporalidades de recuperación, observándose los mejores resultados en el primer grupo y los peores en el tercero (Sánchez et al., 2000).

Particularmente, se ha identificado que las personas que tras un ictus presentan alteración en la sensibilidad profunda presentan un riesgo significativamente mayor (de 9 a 13 veces) de obtener un bajo resultado funcional en comparación con aquellos sin déficit sensitivo, independientemente de la presencia de otras deficiencias clínicas asociadas (Ferrer, 2016).

El nivel de discapacidad funcional inicial se establece como un predictor significativo del pronóstico a largo plazo en personas post-ictus. Específicamente, los individuos que, en las tres primeras semanas tras el ictus, mantienen una puntuación en el Índice de Barthel inferior a 20, presentan un riesgo cuatro veces mayor de tener una dependencia severa en comparación con aquellos con una puntuación inicial superior a 20 (Ferrer, 2016).

Los déficits cognitivos constituyen un factor pronóstico desfavorable en la recuperación post-ictus (Censori et al., 1996), ya que altera las interacciones interpersonales entre la persona, sus familiares y cuidadores; a la vez que reducen la capacidad del individuo para afrontar y compensar las deficiencias físicas. En casos de mayor severidad, los déficits cognitivos pueden interferir sustancialmente con la implementación y progresión de los programas de rehabilitación (Censori et al., 1996).

De entre las diversas variables clínicas, las de mayor relevancia pronóstica en la discapacidad de la persona, incluyen específicamente el grado de paresia, los trastornos sensitivos y la inestabilidad en el control de tronco (Viosca et al., 2005).

1.6.2.2 Protocolos de predicción de recuperación

Para orientar los objetivos, el plan de intervención, la toma de decisiones entre los profesionales que rodean a la persona que ha sufrido un ictus, y con el fin de maximizar la recuperación; se han desarrollado diversos protocolos de predicción de la recuperación clínica de la persona (Stinear et al., 2019):

Protocolos de predicción de recuperación de miembro superior

Destacan, principalmente los protocolos PUPPI (de Havenon et al., 2022), PREP-2 (Stinear et al., 2017) y EPOS-UL (Veerbeek et al., 2022). Estos algoritmos pronostican la probabilidad de alcanzar un resultado de actividad del miembro superior a los 3 meses de la evaluación inicial, la cual debe ser en la fase aguda tras sufrir el ictus. En estos protocolos predictores se tienen en cuenta factores como la edad, los MEP (potenciales evocados), puntuación en la escala NIHSS, puntuación en la escala FMA, y la evaluación SAFE (abducción de hombro y extensión de dedos).

Protocolos de predicción de recuperación de miembro inferior

Los más recomendados por la European Stroke Organisation (ESO) (Kwakkel et al., 2023) son los protocolos EPOS-Walking (Veerbeek et al., 2022), TWIST (Smith et al., 2022) y Kwah (Langerak et al., 2021). Estos algoritmos pronostican la probabilidad de recuperación de la marcha entre los 3 y 6 meses de la evaluación inicial de la fase aguda post-ictus. En los protocolos predictores de recuperación de miembro inferior y marcha, se tienen en cuenta factores como la edad, puntuación en escalas como TCT, BBS, NIHSS, y la fuerza de extensión de rodilla.

Otros protocolos de predicción de recuperación

Existen herramientas que permiten predecir las trayectorias de recuperación en cualquier momento durante el primer año tras un ictus, para la capacidad de actividad de la extremidad superior (Selles et al., 2021) y la independencia en las AVD (Douiri et al., 2017).

La evaluación estandarizada de la función motora al inicio es un elemento esencial de todas las herramientas de predicción disponibles y debe repetirse durante la recuperación para comprender el progreso hacia los resultados previstos. Por lo tanto, se requieren programas de evaluación estandarizados para una planificación y un seguimiento de la rehabilitación consistentemente eficaces (Kwakkel et al., 2023).

2. Cuestionarios de evaluación en Terapia Ocupacional

2.1 Necesidad de evaluación en Terapia Ocupacional

La evolución teórica de la Terapia Ocupacional (TO) ha propiciado diversas concepciones acerca de su proceso intrínseco, mostrando diversos puntos de vista entre diferentes autores. Rogers y Holm (1991) conceptualizan el “Proceso de TO” como una secuencia que comprende dos etapas fundamentales:

- La evaluación, entendida como la fase inicial cuyo objetivo es la identificación y definición de los problemas funcionales del individuo.
- La intervención y re-evaluación, constituyendo la segunda etapa, orientada a la resolución de los problemas previamente identificados en la fase evaluativa.

Posteriormente se ha propuesto una organización alternativa del Proceso de TO (Moruno et al., 2012), delineando sus fases del siguiente modo:

- Evaluación Ocupacional (EO): Se erige como el pilar fundamental del proceso terapéutico, destinada a determinar las limitaciones, necesidades, fortalezas, intereses y metas del usuario. La exhaustividad de esta evaluación es crucial para garantizar una intervención adecuada, dado que en ella se establecen los objetivos y resultados deseables, así como la selección de los métodos de intervención idóneos (Creek, 2008).
- Planificación de la intervención: Fase en la que se organiza la información recabada para establecer un conjunto de objetivos estructurados y mutuamente relacionados. Esto posibilita la selección de los abordajes, metodologías de intervención y técnicas pertinentes para la consecución de los resultados esperados (Moruno et al., 2012).
- Intervención: Definida como la implementación efectiva y eficiente del plan de intervención (Talavera et al., 2022).
- Valoración de resultados: Caracterizada como la investigación sistemática del mérito, valor o éxito de la intervención ejecutada (Creek, 2008).

La etapa de EO se subdivide en tres momentos clave (Romero-Ayuso et al., 2003): la Evaluación Ocupacional Inicial, la Evaluación Ocupacional Continuada y la Valoración de los Resultados o Valoración Ocupacional Final.

La EO Inicial se concibe como un proceso de investigación que determina las áreas primarias del cliente susceptibles de intervención y la idoneidad del usuario para la misma (Creek, 2008). Dentro de la Evaluación Ocupacional Inicial se integran los procesos de derivación, cribado y evaluación comprensiva (Moruno et al., 2012).

- La derivación representa el punto de inicio de la intervención, actuando como el nexo entre el usuario y el profesional. En entornos institucionales, puede ajustarse a formatos o protocolos preestablecidos (Talavera et al., 2022).
- El cribado o screening implica la recopilación de información preliminar por parte del profesional para discernir la necesidad de una futura evaluación o intervención desde TO (Romero-Ayuso et al., 2003).
- La evaluación comprensiva constituye la fase en la que el profesional procura identificar las aspiraciones y necesidades ocupacionales del individuo, así como las barreras, dificultades y facilitadores que influyen en su desempeño (Moruno et al., 2012).

La EO Continuada se desarrolla de manera recurrente a lo largo de todo el proceso terapéutico, con el propósito de recopilar y analizar información referente a la participación del usuario en los programas y espacios terapéuticos acordados (Moruno et al., 2012).

Por último, la Valoración Ocupacional Final se define como la estimación o apreciación global y final del proceso terapéutico (Moruno et al., 2012).

Durante la EO, los profesionales obtienen información de manera sistemática para interpretar la naturaleza, condición y evolución del estado ocupacional del sujeto evaluado (Moruno et al., 2012). Es imperativo que esta fase recoja información tanto de las necesidades como de las fortalezas del individuo, con el fin de generar una imagen clínica equilibrada de su estado ocupacional (Alotaibi et al., 2009). El objetivo principal de la EO se establece en el conocimiento del desempeño ocupacional de la persona (García-Peña et al., 2002; Mathiowetz, 1993; Talavera et al., 2022), considerando su capacidad para elegir, organizar y desempeñar, de forma satisfactoria, ocupaciones significativas (culturalmente definidas y apropiadas para la edad) en los ámbitos del autocuidado, el disfrute de la vida y la contribución al entramado socioeconómico de la comunidad (Kielhofner, 2006).

La EO posee unas características intrínsecas: es un procedimiento singular, circular y se ejecuta de forma sistemática y exhaustiva. Su naturaleza circular se deriva de su manifestación en diferentes momentos a lo largo del Proceso de TO (Creek, 2008). Es singular al proponer una evaluación individualizada del usuario, abordando el "objeto" de su quehacer y las capacidades requeridas, la "forma" que adquiere dicho quehacer, y el "sentido" que este posee para el usuario. La exhaustividad se manifiesta en la adherencia a una serie de pasos que confieren al proceso una estructura definida (Romero-Ayuso et al., 2003).

Otros estudios (Brea, 2008; Trombly, 1993; Weinstock-Zlotnick et al., 2004; Youngstrom, 2002) evidencian que el procedimiento de la EO puede abordarse desde dos enfoques distintos:

- El enfoque *bottom-up* (de abajo hacia arriba): Estructura los pasos a seguir en la EO centrándose en los "componentes funcionales" (ej., rango articular, fuerza). Este enfoque postula que la optimización de estos componentes conduce a un desempeño ocupacional exitoso. No obstante, esta perspectiva ha sido cuestionada por autores como Trombly (1993), quienes argumentan que puede generar insatisfacción en el individuo al desvincularse de la esencia de lo significativo en la TO.
- El enfoque *top-down* (de arriba hacia abajo): Como contraposición al anterior, propone esta secuencia de pasos en la EO, que comienza con la investigación del desempeño ocupacional del individuo, su historia de vida, valores y creencias, a la vez que se consensúa con el usuario el propósito de la TO.

Asimismo, la práctica del terapeuta ocupacional no solo se ve influenciada por el procedimiento evaluativo, sino también por otros factores como el razonamiento clínico. Reed y Sanderson (1999) enfatizan la imperatividad del razonamiento clínico del terapeuta ocupacional para la comprensión del individuo y sus necesidades en el quehacer diario. Autores señalan que el razonamiento clínico es un proceso cognitivo idiosincrásico de cada profesional, supeditado a su experiencia profesional y moldeado por los Marcos de Referencia que sustentan su práctica (Rogers y Holm, 1991; Talavera et al., 2022). Consecuentemente, el razonamiento clínico moldea las estrategias y habilidades mediante las cuales el terapeuta ocupacional dirige e implementa su evaluación e intervención, aplicando su conocimiento y optimizando la práctica a través de la experiencia (Talavera et al., 2022).

Finalmente, los terapeutas ocupacionales emplean fuentes de información individualizadas, adaptadas a las necesidades y situaciones específicas de cada persona a lo largo de la EO (Pollock, 1993; Talavera et al., 2022).

Las fuentes de información a partir de las cuales los profesionales de TO recaban datos sobre los usuarios se consultan siguiendo una secuencia predefinida, priorizando en primera instancia las fuentes de información indirectas y, posteriormente, las fuentes de información directas (Romero-Ayuso et al., 2003).

- Las fuentes de información indirectas engloban informes (médicos, sociales, familiares, etc.), reuniones de equipo, comités de admisión, sesiones clínicas, autoinformes, cuestionarios, autorregistros, entrevistas familiares y registros de huella (Romero-Ayuso et al., 2003).
- Las fuentes de información directas comprenden las entrevistas de TO, la observación directa y la aplicación de herramientas de evaluación (Romero-Ayuso et al., 2003).

Las herramientas de evaluación son instrumentos diseñados para recopilar información sobre la interacción dinámica entre el cliente, el contexto, el entorno y la actividad, cuantificando y examinando los factores que pueden incidir en el desempeño ocupacional (AOTA, 2020; Talavera et al., 2022).

Es necesario que el terapeuta ocupacional domine la utilidad y las condiciones de aplicación de cada herramienta de evaluación para una planificación y ejecución adecuadas del tratamiento. La omisión de este conocimiento sesgaría la práctica basada en la evidencia de la profesión (Romero-Ayuso et al., 2003). Algunos autores (Rogers y Holm, 1991) sugieren que los profesionales exhiben una predisposición en la metodología y características de la recogida de datos. Esta predisposición puede explicarse por diversos factores, como el lugar de práctica, donde la regulación interna de la institución puede determinar el enfoque de la evaluación y, consecuentemente, las competencias profesionales del terapeuta ocupacional.

Además, diversos artículos indican que la selección de las herramientas de evaluación puede estar influenciada por el "paradigma" en el que se enmarca el terapeuta ocupacional (Romero-Ayuso et al., 2003; Talavera et al., 2022). La evolución de los paradigmas de TO ha generado una preocupación por emplear herramientas que trasciendan la mera evaluación de funciones y estructuras corporales, adoptando una visión holística del individuo que integre factores medioambientales y sociales (Reed y Sanderson, 1999; Talavera et al., 2022). Otras investigaciones abordan factores adicionales que influyen en esta selección, tales como los recursos físicos y económicos, el conocimiento del terapeuta ocupacional sobre la herramienta, su disponibilidad, la rapidez en la puntuación y aplicación, la facilidad para expresar los resultados, y la aceptación por parte del usuario, entre otros (Piernik-Yoder et al., 2012).

2.2 Definición de cuestionario de evaluación

Los cuestionarios son herramientas diseñadas para la recolección, universalización y cuantificación de información. Su finalidad principal es homogeneizar los datos y estandarizar el procedimiento de la entrevista (Badía et al., 1998; Martín-Arribas, 2004).

El empleo de cuestionarios es cada vez más prevalente en la investigación y la práctica clínica. Constituyen, junto con las entrevistas, el método de recolección de datos más utilizado en investigación debido a su eficiencia económica, su capacidad para alcanzar a una amplia muestra de individuos y su aptitud para el análisis cuantitativo. No obstante, su implementación no está exenta de limitaciones inherentes que pueden comprometer el desarrollo del estudio (Alcaraz et al., 2006; Argimon et al., 2004; Carvajal et al., 2011; Martín-Arribas, 2004).

Comúnmente, al referirse a un cuestionario, se alude a las escalas de evaluación. Estos instrumentos posibilitan una acumulación progresiva de sus ítems, generando puntuaciones globales al concluir la evaluación. Esta característica acumulativa es la que los distingue de los inventarios de síntomas, los cuestionarios de recogida de datos, los formularios o las entrevistas estandarizadas. Consecuentemente, en la mayoría de los casos, la referencia a "cuestionarios" alude a medidas de resultado (García de Yébenes et al., 2009).

Al igual que las entrevistas, los cuestionarios fundamentan su información en percepciones subjetivas, basándose en la validez otorgada a la información verbal concerniente a sensaciones, comportamientos o actitudes reportadas por el encuestado, o datos que son difíciles de contrastar y convertir en un sistema de medida cuantificable. Esta cualidad intrínseca dificulta el establecimiento de criterios de calidad rigurosos para este tipo de instrumentos (Martín-Arribas, 2004).

2.3 Creación de cuestionarios

La elaboración rigurosa de un cuestionario como instrumento de medición en el ámbito científico-académico requiere una serie de etapas metodológicas. A continuación, se detallan los pasos esenciales para asegurar la validez y fiabilidad del instrumento.

Definición del constructo

La fase inicial y fundamental en el desarrollo de un cuestionario reside en la definición precisa del constructo o del fenómeno que se pretende cuantificar. Este proceso exige una conceptualización inequívoca del objeto de medición, ya sean conductas, actitudes o conocimientos. Para ello, es indispensable llevar a cabo una exhaustiva revisión bibliográfica, complementada, si es pertinente, con la consulta a expertos en la materia. La articulación de esta definición debe ser clara y concisa, fundamentada en las teorías existentes que la sustentan. Un mismo problema puede ser interpretado desde múltiples perspectivas teóricas, lo que puede derivar en diferentes definiciones del constructo. La validez de constructo del cuestionario se determinará por su capacidad para reflejar las teorías relevantes que explican el fenómeno evaluado (Argimon et al., 2004).

Finalidad del cuestionario

La finalidad del cuestionario abarca la determinación de su contenido, la caracterización de la población objetivo, y la especificación de su modo de administración y formato de presentación (Alconero-Camarero et al., 2016). El propósito del cuestionario influye directamente en el diseño de los ítems, en los

aspectos relativos a la recogida de datos y en la estructura general del instrumento.

Al diseñar el cuestionario, es imperativo considerar las características de la población a la que será administrado (por ejemplo, patologías, edad, intereses, nivel cultural). Si se opta por un cuestionario previamente validado, se debe verificar en qué poblaciones específicas ha sido validado (Alconero-Camarero et al., 2016).

Asimismo, el tema de estudio determinará el método de administración: autoadministrado, telefónico, mediante entrevista personal, observacional, etc. La elección del método condicionará la redacción de las preguntas y el formato del cuestionario (Alcaraz et al., 2006). Es esencial ponderar las ventajas y desventajas de cada método de recogida de información. Por ejemplo, las entrevistas personales o encuestas telefónicas suelen presentar tasas de respuesta superiores a las encuestas postales, aunque conllevan un mayor coste. Además, la extensión del cuestionario en encuestas telefónicas puede inducir fatiga en el participante. En ambos casos, se requiere la capacitación del entrevistador, y en los cuestionarios autoadministrados, la inclusión de instrucciones claras y concisas (Alcaraz et al., 2006).

Elaboración de los ítems

Los cuestionarios se componen de un conjunto de ítems. Un ítem se define como la unidad fundamental de información dentro de un cuestionario de evaluación, caracterizándose usualmente por una pregunta y una opción de respuesta cerrada (Martín-Arribas, 2004).

Número de ítems

Se estima que el número mínimo adecuado de ítems para un constructo es de seis, si bien este rango puede oscilar entre 10 y 90, con el fin de abarcar de manera proporcionada todas las dimensiones predefinidas del constructo. Se recomienda diseñar el doble de ítems de los que se considerarán para la versión definitiva del instrumento (Alconero-Camarero et al., 2016).

Contenido

En función de su contenido, los cuestionarios pueden clasificarse como unidimensionales o multidimensionales. En los unidimensionales, más del 80% de los ítems evalúan una única dimensión. Por el contrario, los cuestionarios multidimensionales miden dos o más dimensiones (Martín-Arribas, 2004).

Definición y disposición de los ítems

La definición de cada ítem debe ser mutuamente excluyente y exhaustiva. Al formular la pregunta, es crucial considerar la aceptabilidad y la comprensión por parte del participante, ajustando el lenguaje y el formato de las opciones de respuesta a su nivel sociocultural. Se deben seguir las siguientes directrices para la redacción: evitar palabras redundantes; emplear preguntas breves y de fácil comprensión; abstenerse de utilizar la interrogación "por qué"; evitar la formulación de preguntas en negativo; prescindir de preguntas que demanden esfuerzos de memoria o cálculos complejos; y no formular preguntas cuyas opciones de respuesta incluyan una opción tan atractiva que sea difícil de rechazar (Alconero-Camarero et al., 2016).

Una vez redactados, los ítems deben ser ordenados. Si el cuestionario no sugiere un orden específico, se puede optar por una disposición aleatoria.

Identificación de sesgos

Es fundamental identificar y mitigar los posibles sesgos durante la fase de cumplimentación del cuestionario, lo cual contribuye a perfeccionar la validez del instrumento. Algunos de los sesgos más comunes incluyen: el error de tendencia central (elección de opciones centrales); la deseabilidad social (respuestas acordes con lo socialmente aceptable); el sesgo de aprendizaje o proximidad (respuestas análogas a las anteriores); y el error lógico (asumir puntuaciones idénticas para ítems relacionados) (Choi et al., 2005). El formato, la redacción y el orden de las preguntas y categorías de respuesta pueden inducir sesgos en la recogida de información.

Clasificación de las respuestas

Según el tipo o número de opciones, las respuestas pueden ser: dicotómicas (dos opciones), policotómicas (tres o más opciones) o analógicas (variables en función del sistema analógico empleado, como escalas lineales, numéricas o verbales) (Alconero-Camarero et al., 2016).

Puntuación de los ítems

Es imprescindible establecer el sistema de puntuación a emplear, el cual puede ser simple o ponderado. En los ítems simples, la puntuación directa se obtiene mediante la suma de las respuestas correctas o de los valores asignados a cada opción. En los ítems ponderados, el valor de cada opción de respuesta no es uniforme, o no se concede el mismo valor a cada acierto (Martín-Arribas, 2004).

Fase de validación

Una vez definida la información, y especificadas y ordenadas las preguntas, se procede a la fase de validación, que incluye la realización de una prueba piloto y la evaluación de las propiedades psicométricas del cuestionario (Argimon et al., 2004; Carvajal et al., 2011; Gómez-Benito et al., 2005; Grau, 1995).

Prueba piloto o pre-test

Se administra un borrador del cuestionario a una muestra reducida de 5-10 participantes, preferentemente con características similares a la población de estudio. Esta prueba piloto permite identificar los tipos de preguntas más adecuados, verificar la idoneidad de la categorización de las respuestas, la lógica del ordenamiento interno, la aceptabilidad de la duración, la comprensión y adecuación del enunciado, la extensión de las preguntas y la presencia de intolerancia psicológica o rechazo hacia algún ítem (Terwee et al., 2018; Wild et al., 2005). Los métodos empleados para la prueba piloto incluyen entrevistas informales, encuestas de comprensión, valoración del cuestionario por parte de los participantes y grupos focales de la población diana (Martín-Arribas, 2004).

Evaluación de las propiedades psicométricas

Un cuestionario diseñado para medir un aspecto de la salud y permitir comparaciones intra e inter-individuales debe poseer validez y fiabilidad (Ara et al., 2006; Polit et al., 2000; Tovar, 2007).

La validez determina el grado en que un instrumento mide aquello para lo cual fue diseñado. Es un componente crítico tanto en el diseño como en la confirmación de la utilidad de la medida. Puede estimarse mediante diversas aproximaciones, incluyendo la validez lógica o aparente, la validez de contenido, la validez de criterio y la validez de constructo (Carvajal et al., 2011).

En el contexto de los cuestionarios diagnósticos se introducen dos conceptos adicionales: la sensibilidad (capacidad del cuestionario para detectar cambios en sujetos o atributos evaluados tras una intervención) y la especificidad (capacidad para clasificar correctamente a un sujeto sano como tal). Una mayor sensibilidad y especificidad del cuestionario implican un menor número de falsos positivos y falsos negativos. La evaluación de la validez de un instrumento implica la aplicación de métodos estadísticos (Argimon et al., 2004).

2.4 Evaluación del vestido

La autonomía en las AVD es un pilar fundamental para la independencia y la calidad de vida de las personas (Mlinac et al., 2016). Dentro de las AVD, el

vestido ocupa un lugar central, siendo una actividad básica que todas las personas llevan a cabo a diario (Mlinac et al., 2016). La capacidad de elegir, preparar y ponerse la ropa de forma autónoma contribuye a la autoestima, la imagen corporal y la participación social. Desde TO, la evaluación del vestido busca identificar las limitaciones y desarrollar estrategias para optimizar la participación de la persona en esta actividad (AOTA, 2020; Mlinac et al., 2016).

La evaluación del vestido en TO se basa en un enfoque holístico que considera no solo los componentes físicos, sino también los cognitivos, sensoriales, perceptuales y psicosociales que influyen en la ejecución de la tarea. El terapeuta ocupacional comienza con una entrevista exhaustiva para comprender las rutinas de vestido del individuo, sus preferencias, el tipo de ropa que utiliza habitualmente y cualquier dificultad percibida. Se indaga sobre el contexto en el que se realiza la actividad (dormitorio, baño), la disponibilidad de asistencia y dispositivos de apoyos y el nivel de ayuda que recibe (AOTA, 2020; Pedretti, 2001).

Posteriormente, se realiza una observación directa de la actividad de vestido. Esta observación se recomienda llevarla a cabo en un entorno ecológico, permitiendo al terapeuta analizar la secuencia de acciones, la calidad del movimiento, la coordinación, el equilibrio y la presencia de compensaciones, entre otros aspectos. Se evalúa la capacidad de la persona para planificar y organizar, manipular la ropa, alcanzar y posicionar, coordinar bimanualmente, mantener el equilibrio y la estabilidad, así como su sensibilidad y percepción, y su capacidad de resolución de problemas (Mlinac et al., 2016; Dirette y Gutman, 2020).

Además de la observación, se pueden utilizar herramientas de evaluación estandarizadas para cuantificar el nivel de independencia, siendo las más comúnmente empleadas la Medida de Independencia Funcional (FIM) o el Índice de Barthel (BI), que incluyen ítems relacionados con el vestido. Aunque existen otras escalas de evaluación de la independencia que incluyen un subdominio de vestido, como son la Escala de Rankin modificada (mRS), National Institutes of Health Stroke Scale (NIHSS), Frenchay Activity Index (FAI), Nottingham Extended Activities of Daily Living (NEADL), etc. (Baldelli et al., 2004; Colomer et al., 2023; Jhong et al., 2013; Langhammer et al., 2017; Miki et al., 2016; Sarker et al., 2012).

Sin embargo, no son escalas específicas del vestido, que puedan evaluar de manera exhaustiva los diferentes factores implicados en esta AVD, las diferentes prendas y dificultades que influyen en el correcto desempeño de esta actividad. Actualmente se dispone de tres escalas específicas del vestido: Nottingham Stroke Dressing Assessment (NSDA) (Fletcher-Smith et al., 2010; Fletcher-Smith, 2011; Walker, 1992; Walker et al., 1990; Walker et al., 1991; Walker et al., 2003; Walker et al., 2004; Walker et al., 2012), Upper Body Dressing Scale

(UBDS) (Endo et al., 2015; Ng et al., 2023; Suzuki et al., 2008), y Jacket Test (Chen et al., 2017), que es una subescala de la escala Physical Performance Test (PPT).

2.4.1 Dificultades comunes en la actividad del vestido

Las dificultades en la actividad del vestido pueden ser el resultado de múltiples condiciones y patologías. Algunas de las más comunes incluyen limitaciones físicas como la restricción del rango de movimiento, debilidad muscular, problemas de coordinación y destreza fina, alteraciones del equilibrio y dolor (Pedretti, 2001; Shanks et al., 2025; Walker et al., 2003).

También se presentan déficits cognitivos como la apraxia del vestido, desorientación espacial, déficits de memoria, problemas de atención como la heminegligencia, etc.; los cuales impiden la correcta secuenciación y ejecución de la tarea. Las alteraciones sensoriales y perceptuales, como déficits visuales y alteraciones de la sensibilidad táctil o propioceptiva, también dificultan el vestido (Dirette y Gutman, 2020; Walker et al., 2003).

Finalmente, los factores psicosociales como la depresión, ansiedad, el deterioro de la imagen corporal y el aislamiento social pueden afectar la motivación y la iniciativa para vestirse, impactando directamente en la autonomía de la persona (Trombly y Radomski, 2002).

El vestido tras un ictus

El vestido es una de las AVDs más comúnmente afectada tras el ictus. Se estima que alrededor del 54% de los supervivientes del ictus presentan dificultades para vestirse de forma independiente a los 6 meses de haberlo sufrido; y a los 2 años, el 36% continúa con dificultades en el vestido (Fletcher-Smith et al., 2010; Walker et al., 2012). Esto se debe a la gran variedad de secuelas que provoca el ictus, y a la gravedad de las mismas.

Las secuelas más comunes de un ictus que afectan el vestido incluyen hemiparesia, que es la debilidad o parálisis de un lado del cuerpo, dificultando la manipulación bimanual de la ropa (Ingram et al., 2023; Dawson et al., 2023). La espasticidad, caracterizada por un aumento del tono muscular, también puede generar patrones de flexión o extensión anormales que complican el movimiento de las extremidades para colocarse la ropa (Pedretti, 2001; Walker et al., 2003).

Las alteraciones del equilibrio y la marcha aumentan el riesgo de caídas durante los cambios posturales al vestirse, pudiendo requerir apoyos o la necesidad de vestirse en sedestación (Walker et al., 2003).

La fatiga, una secuela común del ictus, puede incrementar el tiempo y la energía necesarios para vestirse, lo que puede llevar a la persona a posponer la actividad (Walker et al., 2003).

Los déficits sensoriales, como la pérdida de sensibilidad, dificultan la manipulación eficaz de las prendas (Walker et al., 2003). Una puntuación baja en las evaluaciones perceptivas se correlaciona con mayor dependencia para el vestido (Walker et al., 2003). Las alteraciones perceptivas, como las agnosias visuales o táctiles, problemas de campo visual, problemas de tridimensionalidad, de figura-fondo, de orientación espacial, etc., pueden afectar la identificación y colocación de la ropa (Dirette y Gutman, 2020; Sunderland et al., 2006; Walker et al., 2004), siendo las personas que presentan dificultades perceptivas tras el ictus, más dependientes en el vestido (Walker y Lincoln, 1991).

Los déficits cognitivos más comunes tras un ictus que interfieren en la ejecución de la tarea del vestido son: problemas de atención, memoria, planificación, secuenciación, habilidades espaciales, impulsividad, perseveración, etc. (Walker et al., 2003; Walker et al., 2012). Los problemas cognitivos afectan en mayor medida a las personas que se visten empleando una sola mano; y mucho menos a las personas que se visten de manera bimanual (Walker et al., 2004). Además, la heminegligencia y la falta de conciencia del lado afectado, puede llevar a que la persona no se vista la extremidad parética (Walker et al., 2003).

La apraxia afecta al control y uso de objetos cotidianos, como la ropa. La apraxia del vestido, frecuente tras un ictus, puede generar dificultades para secuenciar los pasos, poner la ropa al revés u orientarla espacialmente de forma incorrecta (Walker et al., 2003). El 70% de las personas que presentan apraxia tras un ictus, muestra dependencia para vestirse; siendo los errores más comunes la estrategia desorganizada, colocar la mano afecta en otro orificio diferente al de la manga, no cubrir el hombro parético, y no colocar la manga del brazo parético (Walker et al., 2004).

Existen diferencias tanto en la afectación como en la recuperación del vestido según el hemisferio afectado (Sunderland et al., 2006; Walker et al., 2012). Para aquellos con ictus en el hemisferio derecho, el vestido se ve afectado por problemas visuoespaciales o una atención sostenida deficiente. Las personas con ictus en el hemisferio izquierdo y apraxia ideomotora no pudieron aprender el procedimiento correcto para compensar la hemiparesia al vestirse. El tratamiento resulta más eficaz en personas con ictus en el hemisferio derecho; mientras que las personas con ictus en el hemisferio izquierdo no mejoran tanto en el vestido (Sunderland et al., 2006).

Las personas con ictus en el hemisferio derecho en ocasiones presentan problemas para seleccionar la manga correcta, autocomprobar su lado izquierdo o cubrir el hombro parético, sugiriendo déficits en la percepción visuoespacial o

negligencia. Las personas con ictus en el hemisferio izquierdo tienden a vestir primero el brazo no parético o mostrando una estrategia de vestido desorganizada, lo que sugiere un control de la acción deficiente debido a la apraxia (Walker et al., 2004).

Walker y Lincoln (1991) hallaron que los problemas más prevalentes a la hora de vestirse en personas que habían sobrevivido a un ictus fueron ponerse el zapato en el pie afecto, subirse los pantalones y atarse los cordones. Siendo los errores más frecuentes la desorganización sin estrategias, la heminegligencia, el error de selección y la perseveración (Fletcher-Smith et al., 2010).

Walker y Lincoln (1991) determinaron que la recuperación motora y de las habilidades perceptivas son determinantes y muy importantes para vestirse tras un ictus.

Desde TO, la intervención tras un ictus en la actividad del vestido se centra en la rehabilitación de los componentes subyacentes, como la fuerza y el rango de movimiento; acompañado de un entrenamiento funcional (Walker et al., 2012). Aunque también se realiza entrenamiento en estrategias compensatorias, se promueve la adaptación de la ropa con prendas holgadas y cierres de velcro, y el uso de productos de apoyo (Pedretti, 2001; Walker et al., 2004; Walker et al., 2012).

También se realiza la modificación del entorno, organizando el armario y asegurando una buena iluminación. Se entrena en la secuencia del vestido, descomponiendo la tarea en pasos más pequeños. Además, se aplican intervenciones cognitivas para la apraxia del vestido y la heminegligencia, y se brinda educación a la persona y cuidadores sobre las dificultades, estrategias y productos de apoyo (Trombly y Radomski, 2002; Walker et al., 2003).

En muchas ocasiones, las personas que han sobrevivido a un ictus son capaces de vestirse a pesar de las limitaciones, con un desempeño lento y torpe, pero logran vestirse por sí solos disponiendo del tiempo suficiente (Walker et al., 2004).

En resumen, la evaluación del vestido desde TO es un proceso minucioso que aborda las múltiples dimensiones que influyen en esta AVD. Tras un ictus, las secuelas neurológicas pueden comprometer severamente la capacidad de vestirse, requiriendo una evaluación exhaustiva que permita orientar la intervención especializada de forma que fomente la máxima independencia posible y mejore la calidad de vida de estas personas (Dirette y Gutman, 2020; Fletcher-Smith et al., 2010).

2.4.2 Nottingham Stroke Dressing Assessment

La Nottingham Stroke Dressing Assessment (NSDA) fue originalmente creada y publicada en inglés por Marion Walker en 1990 (Walker et al., 1990). Esta herramienta fue diseñada como una medida de evaluación muy completa para la AVD del vestido, valorando un gran número de prendas, descomponiendo en pasos la colocación de cada prenda y otorgando una puntuación ajustada a la independencia o necesidad de ayuda para vestirse en cada paso. A su vez, analiza si el origen de los errores es motor, cognitivo, perceptivo, emocional, o la combinación de ellos, lo cual no se tiene en cuenta en otras escalas. Finalmente, refleja una puntuación total de la independencia de la persona que ha sufrido un ictus para vestirse de forma autónoma todo el cuerpo, no solo una parte del cuerpo (Fletcher-Smith et al., 2010; Fletcher-Smith, 2011; Walker, 1992; Walker et al., 1990; Walker et al., 1991; Walker et al., 2003; Walker et al., 2004; Walker et al., 2012)

Esta escala consta de 3 documentos:

- El primer documento, que recibe el nombre de “General Therapist’s Instructions”, recoge las instrucciones para la administración de la escala, recomendaciones para el terapeuta, así como la explicación de su sistema de puntuación, tanto para los ítems, como para la puntuación final.
- El segundo documento es de evaluación cuantitativa. Tiene 2 versiones: femenina (“NSDA-Female”) y masculina (“NSDA-Male”). Consiste en una tabla de puntuación y un apartado de observaciones. En ella aparecen las diferentes prendas de vestir y cierres que más se emplean. Para cada prenda hay un desglose de los diferentes pasos que se realizan desde que se coge la prenda hasta que la persona consigue colocarla. Teniendo en cuenta las diferentes estrategias de vestido que se pueden emplear para la colocación de algunas prendas, como el sujetador o la falda.
- El tercer documento es de evaluación cualitativa. Es una tabla de análisis de los errores más comúnmente cometidos durante el vestido, por diferentes factores, tanto motores, como cognitivos y perceptivos, y cómo influyen a la hora de vestirse en la persona a evaluar.

El segundo documento se divide en los siguientes apartados para hombre:

- Cierres: cremallera, botón grande, botón pequeño, velcro, botón de presión, nudo, hebilla, corchete pantalón.
- Camiseta de tirantes
- Camiseta de manga corta
- Camisa/chaqueta
- Camiseta/jersey de manga larga
- Calzoncillos
- Pantalones

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

- Calcetines
- Zapatos

Y en los siguientes apartados en la escala de mujer:

- Cierres: cremallera, botón grande, botón pequeño, velcro, corchete, botón de presión, nudo, hebilla, corchete de pantalón
- Sujetador
- Camiseta de tirantes
- Camiseta de manga corta
- Vestido
- Camisa/chaqueta
- Camiseta/jersey de manga larga
- Falda
- Bragas
- Pantalones/leggings
- Medias/mallas/calcetines
- Zapatos

Siendo la escala de mujer más extensa, con 46 ítems en total; mientras que la escala de hombre posee 29 ítems a evaluar.

Los cierres solo cuentan con un paso a evaluar, todas las demás prendas cuentan con tres pasos a evaluar en la colocación de dicha prenda, excepto calzoncillos, pantalones y calcetines, que cuentan con dos pasos a evaluar.

En la escala de mujer, el sujetador cuenta con 3 estrategia de vestido diferente; y la falda cuenta con otras 2 estrategias de vestido.

Cada paso o cierre de prenda del segundo documento se puntúa en una escala ordinal de 3 puntos (0=necesita asistencia física para realizarlo, 1=necesita asistencia no física para realizarlo, 2=lo hace de manera independiente) (Walker, 1992).

La única propiedad psicométrica estudiada para la escala NSDA es la fiabilidad, la cual fue analizada empleando el Índice de Correlación Intraclass (Intraclass Correlation coefficient-ICC), presentando una puntuación de 0.99 (Fletcher-Smith et al., 2010; Fletcher-Smith, 2011)

2.5 Escalas de evaluación de la función sensoriomotora en miembro superior

La evaluación de la función sensoriomotora del miembro superior es un pilar fundamental en la rehabilitación neurológica; permitiendo no solo cuantificar el déficit presente, sino también guiar la intervención terapéutica, permitiendo monitorizar la recuperación y ajustar las estrategias de tratamiento (Kwakkel et al., 2023; Langhorne et al., 2009). Un enfoque integral de la evaluación abarca tanto los componentes motores como los sensoriales, dada su intrínseca interdependencia en la ejecución de movimientos en las AVDs y la participación social (Dawson et al., 2023; Doyle et al., 2010; Soleimani et al., 2024).

La afección del miembro superior es una de las secuelas más comunes y discapacitantes tras un ictus, viéndose afectada la función del miembro superior en el 80% de los supervivientes de un ictus (de Rooij et al., 2016; Rathore et al., 2002), e impactando significativamente la autonomía del individuo (Dawson et al., 2023). La hemiparesia o hemiplejía, caracterizadas por la debilidad o parálisis de un lado del cuerpo, respectivamente, se deben a la interrupción de las vías corticoespinales descendentes que controlan el movimiento voluntario (Planton et al., 2012; Ingram et al., 2023; Zhao et al., 2023). Más allá de la debilidad, el ictus puede manifestarse con alteraciones en el tono muscular, como la espasticidad, que dificulta el movimiento y puede causar deformidades (Gracies, 2005; Li et al., 2023; Zhao et al., 2023). La incoordinación, ataxia y temblor son otras posibles disfunciones motoras que afectan la precisión y fluidez de los movimientos (Kilbride et al., 2019).

Paralelamente, las alteraciones sensoriales, como la disminución de la propiocepción, el tacto, la discriminación de dos puntos y la estereognosia, son prevalentes y contribuyen a la disfunción global del miembro superior (Carey, 2013).

Estas deficiencias pueden comprometer la capacidad de la persona superviviente del ictus para manipular objetos, realizar tareas bimanuales y mantener el equilibrio, lo que repercute en su calidad de vida (Shanks et al., 2024; Veerbeek et al., 2011).

2.5.1 Escalas de evaluación

Para evaluar la función sensoriomotora del miembro superior tras un ictus, actualmente hay disponibles diversas escalas; entre las más comunes se encuentran: Fugl-Meyer Assessment (FMA), Action Research Arm Test (ARAT), Box and Block Test (BBT), Nine Hole Peg Test (NHPT), Chedoke Arm and Hand Activity Inventory (CAHAI), Chedoke-McMaster Stroke Assessment (CMSA), Arm Motor Ability Test (AMAT), Wolf Motor Function Test (WMFT), entre muchas otras (Bushnell et al., 2015; Gunduz et al., 2023; Kwakkel et al., 2023;

Santisteban et al., 2016). Cada una de ellas ha sido diseñada para medir aspectos específicos de la recuperación y la funcionalidad (Bushnell et al., 2015; Gunduz et al., 2023; Kwakkel et al., 2023). La selección de la escala adecuada depende del objetivo de la evaluación, la fase de recuperación de la persona y el entorno clínico (Prinsen et al., 2018; Santisteban et al., 2016).

Una de las herramientas más utilizadas y validadas es la Fugl-Meyer Assessment (Gunduz et al., 2023; Kwakkel et al., 2023). Esta escala es una medida exhaustiva de la recuperación motora después del ictus, evaluando el movimiento voluntario, la sinergia, el rango de movimiento y el dolor (Fugl-Meyer et al., 1975). La FMA es ampliamente reconocida por su fiabilidad y validez, siendo un estándar de oro en la investigación de la recuperación motora (Bushnell et al., 2015; De Blas et al., 2025; Gladstone et al., 2002; Gunduz et al., 2023; Santisteban et al., 2016; Yu et al., 2024).

Otra escala relevante es la Action Research Arm Test (ARAT), que evalúa la capacidad de la persona para realizar actividades funcionales del miembro superior. Se compone de 19 ítems que miden la prensión, el agarre, el transporte de objetos y las actividades gruesas (Yozbatiran et al., 2008). ARAT es particularmente útil para documentar las mejoras en la destreza manual y la funcionalidad en tareas de la vida diaria (Dandekar y Ganvir, 2019; van der Lee et al., 2001a; van der Lee et al., 2001b). Cuenta con una versión traducida, adaptada y validada al castellano (Fernández-Solana et al., 2022). Su enfoque en la capacidad funcional la convierte en una herramienta complementaria a la FMA-UE, que se centra más en los patrones de movimiento.

La escala Wolf Motor Function Test (WMFT) es una prueba cronometrada que evalúa la velocidad y la precisión de los movimientos del miembro superior en tareas funcionales (Wolf et al., 2001). Incluye 15 ítems que van desde movimientos simples hasta tareas más complejas que requieren coordinación bimanual. La WMFT es sensible a los cambios pequeños en la función motora y es frecuentemente utilizada en estudios de investigación para evaluar la efectividad de las intervenciones terapéuticas (Morris et al., 2006).

En resumen, la evaluación precisa de la función sensoriomotora del miembro superior después de un ictus es un componente crítico de la atención neurorrehabilitadora (Kwakkel et al., 2023). Existen diversas escalas para evaluar este constructo, que proporcionan herramientas validadas para cuantificar los déficits, monitorear la recuperación y guiar las intervenciones terapéuticas (Kwakkel et al., 2023; Langhorne et al., 2009). La combinación de estas herramientas permite una comprensión holística de la afectación del miembro superior, facilitando la toma de decisiones clínicas informadas y la planificación de estrategias de rehabilitación individualizadas para optimizar la recuperación funcional de las personas (Dawson et al., 2023; Lang et al., 2008).

2.5.2 Fugl-Meyer Assessment

La escala Fugl-Meyer Assessment (FMA) es una herramienta ampliamente reconocida y utilizada en el ámbito de la neurorrehabilitación para cuantificar la recuperación motora y sensorial en personas que han sufrido un ictus (Prinsen et al., 2018; Santisteban et al., 2016; Yu et al., 2023). Su origen se remonta a la necesidad de disponer de una medida estandarizada y fiable para evaluar la función motora post-ictus, en un momento en que las evaluaciones carecían de la rigurosidad necesaria para la investigación y la práctica clínica (Bushnell et al., 2015; Fugl-Meyer et al., 1975; Gunduz et al., 2023; Santisteban et al., 2016).

El desarrollo de FMA fue llevado a cabo por Axel Fugl-Meyer, Lennart Jääskö e Iréne Leyman en Suecia a principios de la década de 1970. Basaron su trabajo en la observación detallada de los patrones de recuperación motora en personas con hemiplejía, inspirándose en las etapas de recuperación descritas por Brunnstrom (Brunnstrom, 1966). Su objetivo era crear una escala de evaluación exhaustiva que cubriera los diferentes aspectos de la función motora y sensorial. La escala fue publicada por primera vez en 1975, sentando las bases para su adopción global como estándar de oro (Fugl-Meyer et al., 1975; Santisteban et al., 2016; Yu et al., 2023).

FMA es una escala multidimensional que evalúa la función motora, el equilibrio, la sensación y la función articular/dolor; es decir, evalúa miembro superior (mediante la subescala FMA de extremidad superior (FMA-UE) y miembro inferior (mediante la subescala FMA de extremidad inferior (FMA-LE)).

La subescala FMA-UE, se divide en las secciones Extremidad Superior (la cual, a su vez, se subdivide en: actividad refleja, movimiento voluntario dentro de sinergias, movimiento voluntario mezclando sinergias, movimiento voluntario con poca o ninguna sinergia, y actividad refleja normal), Muñeca, Mano (incluyendo agarres), Coordinación/Velocidad, Sensación, Movimiento Articular Pasivo y Dolor (Amano et al., 2018; De Weerd et al., 1985; Hsieh et al., 2009; Kim et al., 2012; Kim et al., 2016; Lin et al., 2009; Rabadi y Rabadi, 2006; Sullivan et al., 2011; Tauchi et al., 2022; Van der Lee y Beckman, 2001; Wei et al., 2011; Woodbury et al., 2007).

Por otro lado, la subescala FMA-LE se divide en las secciones Extremidad Inferior (que se subdivide en: actividad refleja, movimiento voluntario dentro de sinergias, movimiento voluntario mezclado con sinergias, movimiento voluntario con poca o ninguna sinergia, y actividad refleja normal), Coordinación/Velocidad, Sensación, Movimiento Articular Pasivo y Dolor (Hochleitner et al., 2022; Kim et al., 2012; Sullivan et al., 2011).

Estas subsecciones permiten una evaluación detallada de la recuperación motora, desde los movimientos más básicos hasta los más complejos y selectivos (Gladstone et al., 2002). La evaluación se realiza mediante la observación directa del desempeño de la persona en una serie de tareas

específicas, con ítems graduados para reflejar diferentes niveles de dificultad y control motor (Fug-Meyer et al., 1975; Gladstone et al., 2002).

La puntuación de FMA es un proceso sistemático donde cada ítem se califica en una escala ordinal de 3 niveles, desde 0 puntos hasta 2 puntos: 0 para "no puede realizar", 1 para "realiza parcialmente o con dificultad", y 2 para "realiza completamente o sin dificultad" (Fugl-Meyer et al., 1975).

La puntuación máxima para la sección motora del miembro superior (compuesta por las secciones de extremidad superior, muñeca, mano y coordinación/velocidad) es de 66 puntos; mientras que para la sección motora de miembro inferior (compuesta por las secciones de extremidad inferior y coordinación/velocidad) la puntuación máxima es de 34 puntos; donde una puntuación más alta indica un mejor funcionamiento motor (Sanford et al., 2014). Las puntuaciones de cada subsección se suman para obtener una puntuación total. Esta metodología de puntuación permite una cuantificación objetiva de la recuperación motora y facilita la monitorización de los cambios a lo largo del tiempo, siendo fundamental para la investigación y la toma de decisiones clínicas (Santisteban et al., 2016; Yu et al., 2023).

FMA es internacionalmente conocida, esto hace que tenga versiones traducidas y validadas en diferentes idiomas, como colombiano (Barbosa et al., 2019), danés (Busk et al., 2022; Lundquist et al., 2017), italiano (Cecchi et al., 2021; Hochleitner et al., 2023), japonés (Amano et al., 2012), coreano (Kim et al., 2021), urdu (Ikram et al., 2022), rumano (Onose et al., 2023) y portugués-brasileño (Michaelsen et al., 2011).

También se han creado versiones reducidas de FMA, como la versión de 37 ítems (Chen et al., 2014; Hsieh et al., 2007; Lin et al., 2021), de 12 ítems (Chen et al., 2014; Fu et al., 2012; Hsieh et al., 2007; Hsueh et al., 2008; Lin et al., 2021) y de 6 ítems (Amano et al., 2020).

Todas las propiedades psicométricas de FMA han sido estudiadas: validez de contenido (Malounin et al., 1994); validez estructural (presentando un solo factor (Chen et al., 2014; Lin et al., 2021)); consistencia interna (alfa de cronbach entre 0,94 y 0,98 (Lin et al., 2004)); error de medición (Hsueh et al., 2008; Lin et al., 2021; Sanford et al., 1993), fiabilidad (Hsueh et al., 2008; Lin et al., 2004; Sanford et al., 1993; See et al., 2013), validez de criterio (desde 0,31 hasta 0,99 (Chen et al., 2014; Hsueh et al., 2008; Lin et al., 2004; Lin et al., 2021)), prueba de hipótesis (Hsueh et al., 2008; Lin et al., 2004; Lin et al., 2021; Malounin et al., 1994; See et al., 2013) y sensibilidad al cambio (Chen et al., 2014; Hsueh et al., 2008; Lin et al., 2004; Lin et al., 2021; Malounin et al., 1994; See et al., 2013).

Todo ello causa que FMA sea considerada la "gold standard" para evaluar la función del miembro superior tras sufrir un ictus (Bushnell et al., 2015; Gunduz et al., 2023; Santisteban et al., 2016; Yu et al., 2023).

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de
Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un
Ictus

3. Análisis psicométrico y revisión de los cuestionarios de evaluación

3.1 Normativa COSMIN

3.1.1 Lista de verificación COSMIN

La lista de verificación COSMIN (COnsensus-based Standards for the selection of health Measurement INstruments) constituye una herramienta metodológica fundamental para la evaluación de la calidad de estudios sobre las propiedades de medición de cuestionarios de evaluación en el ámbito de la salud (Mokkink et al., 2010a; Mokkink et al., 2010b). Su propósito principal es facilitar la selección de instrumentos de alta calidad para la investigación y la práctica clínica.

Dentro de la estructura de la lista de verificación COSMIN, destacan dos conceptos esenciales: el Resultado Informado por el Paciente (Patient-Reported Outcome - PRO) y la Medida de Resultado Informada por el Paciente (Patient-Reported Outcome Measure - PROM). Un PRO se define como cualquier aspecto del estado de salud de una persona que es evaluada directamente por el propio individuo, sin interpretación externa de su respuesta ((U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research et al., 2006). Los PROs se cuantifican con mayor frecuencia mediante cuestionarios autoadministrados, los cuales son comúnmente denominados PROMs. No obstante, la calidad metodológica de estos cuestionarios muestra una gran variabilidad, lo que a menudo dificulta determinar si la herramienta seleccionada posee la fiabilidad y validez óptimas (Griffiths et al., 2015).

Un aspecto crucial en la evaluación de los instrumentos de medición en salud es la incorporación de la lista de verificación COSMIN. Esta herramienta incluye una tabla para cada propiedad de medición que especifica estándares metodológicos relativos al diseño del estudio y a la selección de métodos estadísticos, lo que permite evaluar la calidad metodológica de las investigaciones (Mokkink et al., 2018a; Mokkink et al., 2019).

La aplicación de la metodología COSMIN en las revisiones sistemáticas eleva la capacidad de estas últimas para identificar el cuestionario más idóneo para la cuantificación de un constructo específico en una población de estudio determinada. Esta relevancia se acentúa en un contexto donde el número de revisiones sistemáticas ha experimentado un crecimiento exponencial, pasando de aproximadamente una por año a principios de la década de 1990 a más de 100 anuales en la actualidad (Mokkink et al., 2016). No obstante, este incremento ha conllevado un deterioro concomitante en la calidad metodológica de algunas de estas revisiones, situación que ha sido corroborada por una

revisión reciente que evidenció un considerable margen de mejora en este ámbito (Terwee et al., 2016).

Estas revisiones, que incorporan la lista de verificación COSMIN, ofrecen una panorámica exhaustiva de las propiedades de medición de los cuestionarios evaluados. Esto permite respaldar la formulación de recomendaciones basadas en la evidencia para la selección del instrumento más adecuado, atendiendo a su propósito específico (investigación o práctica clínica) y a su aplicación (discriminatoria, evaluativa o predictiva) (Prinsen et al., 2018). Adicionalmente, facilitan la identificación de brechas en el conocimiento sobre las propiedades psicométricas de estos instrumentos, lo cual es fundamental para el diseño de futuras investigaciones en este ámbito (Prinsen et al., 2018).

Ante la ausencia de una guía extensiva y formalmente publicada para revisiones sistemáticas de cuestionarios de evaluación, el comité directivo de COSMIN desarrolló un protocolo específico. Este protocolo, diseñado para este tipo de revisiones, está disponible públicamente en la página web de COSMIN desde el año 2011 (<http://www.cosmin.nl>) (COSMIN, 2011).

La guía COSMIN detalla en diez puntos la metodología para la realización de revisiones sistemáticas de cuestionarios existentes. Esta metodología se aplica a instrumentos de medición para los cuales ya se dispone de información sobre sus propiedades psicométricas, su idoneidad para fines evaluativos y su selección para un propósito específico (COSMIN, 2011).

Para obtener información detallada que respalde la ejecución de una revisión sistemática, se pueden consultar las guías "COSMIN methodology for systematic reviews of PROMs—user manual" y "COSMIN methodology for assessing the content validity of PROMs—user manual", ambas disponibles en el sitio web de COSMIN (COSMIN, 2011; Mokkink et al., 2018b; Terwee et al., 2018).

El manual al que se hace referencia fue elaborado siguiendo una metodología rigurosa y consolidada, que integra directrices de diversas fuentes reconocidas en el ámbito de las revisiones sistemáticas y la evaluación de instrumentos. Específicamente, su desarrollo se fundamentó en los siguientes marcos y estándares:

- Manual Cochrane para revisiones sistemáticas de intervenciones (Higgins et al., 2024).
- Manual Cochrane para revisiones de precisión de pruebas diagnósticas (Deeks et al., 2023).
- Declaración PRISMA (Moher et al., 2010).
- Lista de verificación STARD (Bossuyt et al., 2003).
- Normas del Instituto de Medicina para revisiones sistemáticas de la investigación de la eficacia comparativa (Institute of Medicine, 2011).

- Criterios actualizados para las buenas propiedades de medición de Terwee y Prinsen (Mokkink et al., 2018b; Terwee et al., 2018).
- Metodología GRADE (Grading of Recommendations Assessment, Development and Evaluation) para la formulación y calificación de recomendaciones en la práctica clínica (Mokkink et al., 2018b; Terwee et al., 2018; GRADE Handbook, 2013).

3.1.2 Guía COSMIN para revisiones sistemáticas

Para la realización de una revisión sistemática orientada a la medición de un constructo de interés en una población de estudio específica, se recomienda seguir el procedimiento estructurado en diez puntos. Este procedimiento se organiza, a su vez, en tres secciones principales:

3.1.2.1 Búsqueda de la literatura

Este apartado consta de cuatro puntos (del 1 al 4), los cuales establecen procedimientos estandarizados que cumplen con las normas actuales para la realización de revisiones sistemáticas (Higgins et al., 2024; Deeks et al., 2023).

Punto 1: Formulación del objetivo de la revisión

El objetivo de una revisión sistemática, cuyo fin es la medición de un constructo de interés en una población de estudio específica mediante cuestionarios de evaluación, debe enfocarse en la calidad de dichos instrumentos. Para ello, debe integrar cuatro elementos clave (Prinsen et al., 2018):

- El constructo (qué medir): La variable o concepto que se pretende cuantificar.
- La población: El grupo de individuos al que se aplica el instrumento.
- El tipo de instrumento: La naturaleza del cuestionario o escala utilizada.
- Las propiedades de medición de interés: Las propiedades psicométricas a evaluar (Prinsen et al., 2018).

Punto 2: Establecimiento de criterios de elegibilidad

Los criterios de elegibilidad para la revisión sistemática deben ser coherentes con los cuatro elementos clave del objetivo de la revisión previamente formulado. Esto implica que (Prinsen et al., 2018):

- El cuestionario seleccionado debe ser directamente relevante para el constructo de interés.

- La muestra del estudio ha de ser representativa de la población de interés.
- El estudio debe hacer referencia al cuestionario.
- El propósito del estudio debe ser la evaluación de una o más propiedades psicométricas, el desarrollo de un cuestionario (para validar su contenido) o la evaluación de la interpretabilidad de los cuestionarios de interés (Prinsen et al., 2018).

Se recomienda la exclusión de estudios que empleen el cuestionario únicamente como un instrumento de medición de resultados (por ejemplo, en ensayos controlados aleatorios o como validación de otro instrumento), dado que estos proporcionan evidencia indirecta sobre las propiedades de medición del propio cuestionario. Finalmente, es aconsejable incluir exclusivamente artículos a texto completo. Los resúmenes suelen contener información limitada sobre el diseño del estudio, lo que puede dificultar significativamente la evaluación de la calidad metodológica y los resultados de las propiedades de medición en las siguientes fases de la revisión (Puntos 5 a 7) (Prinsen et al., 2018).

Punto 3: Realización de la búsqueda bibliográfica

Para la ejecución de una revisión sistemática, es necesario llevar a cabo una búsqueda bibliográfica exhaustiva. Conforme a la metodología Cochrane (Deeks et al., 2023; Higgins et al., 2024) y por consenso generalizado, MEDLINE y EMBASE constituyen las bases de datos mínimas requeridas para realizar las búsquedas. Adicionalmente, se recomienda explorar otras bases de datos especializadas, cuya selección dependerá del constructo y la población de interés específicos (Prinsen et al., 2016).

Una estrategia de búsqueda eficaz se caracteriza por una colección completa de términos, que incluye tanto términos como descriptores de ciencias de la salud (DeCS), o palabras de texto libre. Estos términos deben corresponder a los cuatro elementos clave del objetivo de la revisión: el constructo, la población, el tipo de instrumento y las propiedades de medición (Prinsen et al., 2016).

En lo que respecta a los términos de búsqueda relacionados con las propiedades de medición, se aconseja la utilización de un filtro de búsqueda validado y de alta sensibilidad. Estos filtros están disponibles para PubMed y EMBASE en el sitio web de COSMIN (COSMIN, 2011).

Finalmente, y en línea con la metodología Cochrane, se recomienda que las búsquedas en las bases de datos se realicen desde su fecha de inicio hasta la actualidad. La aplicación de restricciones de idioma dependerá de los criterios de inclusión definidos en el Punto 2. No obstante, se desaconseja emplear dichas restricciones en la estrategia de búsqueda, incluso si no se dispone de recursos para la traducción de los artículos. Esta aproximación permite que los revisores,

al menos, puedan informar sobre la existencia de estos estudios (Deeks et al., 2023; Higgins et al., 2024).

Punto 4: Selección de resúmenes y artículos a texto completo

La selección de resúmenes y artículos a texto completo debe ser realizada por dos revisores de manera independiente. Si un estudio es considerado potencialmente relevante por al menos uno de los revisores, basándose en el resumen, o en caso de incertidumbre, se procederá a la recuperación y revisión del artículo a texto completo. Las discrepancias entre los dos revisores deben ser discutidas, y en ausencia de consenso, se recomienda la consulta a un tercer revisor.

Adicionalmente, se aconseja verificar todas las referencias de los artículos finalmente incluidos para identificar estudios adicionales que puedan ser relevantes. La detección de un número significativo de artículos nuevos durante este proceso podría indicar que la estrategia de búsqueda inicial no fue lo suficientemente exhaustiva, lo que justificaría su mejora o repetición (Deeks et al., 2023; Higgins et al., 2024).

3.1.2.2 Evaluación de las propiedades de medición

Este apartado, que comprende los puntos 5 a 7, se centra en la evaluación sistemática de las propiedades de medición de los cuestionarios incluidos en la revisión. Dicho proceso se estructura en tres fases consecutivas:

1. Evaluación de la calidad metodológica del estudio: la calidad metodológica de cada estudio, en relación con una propiedad de medición específica, se evalúa empleando la lista de verificación COSMIN Risk of Bias. Cada estudio se clasifica en una de las siguientes categorías de calidad: "muy buena", "adecuada", "dudosa" o "inadecuada" (Mokkink et al., 2018a).
2. Calificación de la calidad de los resultados: en la segunda fase, la calidad de los resultados obtenidos por cada estudio para una propiedad de medición se califica utilizando los criterios actualizados de Terwee y Prinsen (Mokkink et al., 2018b; Terwee et al., 2018). Los resultados pueden ser valorados como "suficiente" (+), "insuficiente" (-) o "indeterminado" (?) para todas las propiedades psicométricas. Sin embargo, para la validez de contenido, la calificación difiere, siendo "suficiente" (+), "insuficiente" (-), "inconsistente" ($\pm$) o "indeterminado" (?).
3. Resumen y calificación de la evidencia: Finalmente, la evidencia recopilada se resume y su calidad global se califica mediante la aplicación de la metodología GRADE (GRADE Handbook, 2013; Mokkink et al., 2018b; Terwee et al., 2018).

Los hallazgos de todos los estudios disponibles relativos a una propiedad de medición específica son agrupados de manera cuantitativa o resumidos cualitativamente. Posteriormente, estos resultados se comparan con los criterios establecidos para las buenas propiedades de medición, lo que permite determinar si la propiedad del cuestionario es "suficiente" (+), "insuficiente" (-), "inconsistente" ($\pm$) o "indeterminada" (?) (Mokkink et al., 2018b; Terwee et al., 2018).

Estas calificaciones generales asignadas a cada propiedad de medición se complementan con una calificación de la calidad de la evidencia. El método GRADE, diseñado para revisiones sistemáticas de estudios de intervención, establece cuatro niveles para clasificar la calidad de la evidencia: "alta", "moderada", "baja" o "muy baja". Esta clasificación se basa en la evaluación de cinco factores clave: riesgo de sesgo, ausencia de evidencia directa, inconsistencia, imprecisión y sesgo de publicación (GRADE Handbook, 2013; Mokkink et al., 2018b; Terwee et al., 2018).

Para las revisiones sistemáticas centradas en la evaluación de instrumentos, se emplea una metodología GRADE modificada. Esta adaptación se utiliza específicamente para degradar la calidad de la evidencia cuando existen preocupaciones sobre la fiabilidad de los resultados. Para este propósito, se han adoptado cuatro de los cinco factores originales de GRADE, integrándolos en la metodología COSMIN:

- Riesgo de sesgo: Hace referencia a la calidad metodológica de los estudios individuales.
- Inconsistencia: Se refiere a la incongruencia inexplicable de los resultados entre los estudios.
- Imprecisión: Relacionada con el tamaño total de la muestra de los estudios disponibles.
- Ausencia de evidencia directa: Indica la existencia de evidencia proveniente de poblaciones diferentes a la población de interés en la revisión (GRADE Handbook, 2013; Mokkink et al., 2018b; Terwee et al., 2018).

Punto 5: Evaluación de la validez de contenido

La validez de contenido hace referencia al grado en que el contenido de un cuestionario representa de forma adecuada el constructo que pretende medir (Mokkink et al., 2010b). Esta propiedad se considera la más importante, ya que es fundamental asegurar que los ítems del cuestionario sean relevantes, exhaustivos y comprensibles en relación con el constructo de interés y la población de estudio específica (Terwee et al., 2018).

Si se identifica alta evidencia que indica una validez de contenido insuficiente en un cuestionario, este instrumento será excluido de las fases posteriores de la revisión sistemática (Puntos 6, 7 y 8). En tal caso, se procederá directamente a formular una recomendación explícita sobre su uso en el Punto 9 (Prinsen et al., 2018).

Punto 6: Evaluación de la estructura interna: validez estructural, consistencia interna y validez transcultural/invariancia de medición

La estructura interna de un cuestionario hace referencia a las interrelaciones entre sus distintos ítems, un aspecto fundamental para determinar cómo estos pueden agruparse en escalas o subescalas (Mokkink et al., 2010b). En este punto, se lleva a cabo una evaluación exhaustiva que comprende la validez estructural (incluyendo la unidimensionalidad), la consistencia interna y la validez transcultural/invariancia de medición. Estas tres propiedades se focalizan en la calidad de los ítems individuales y en las relaciones subyacentes entre ellos (Prinsen et al., 2018).

Es fundamental que todas las propiedades psicométricas que componen la estructura interna de un cuestionario sean evaluadas inmediatamente después de la validación del contenido. Dado que la evidencia de la validez estructural (o unidimensionalidad) de una escala o subescala es un requisito previo para la interpretación de los análisis de consistencia interna (ej., Alfa de Cronbach); por lo que se recomienda un orden específico para su evaluación: primero, la validez estructural, seguida de la consistencia interna, y finalmente, la validez transcultural/invariancia de medición (Prinsen et al., 2018).

El punto 6 es relevante únicamente para los cuestionarios basados en un modelo reflexivo. Este modelo postula que todos los ítems de una escala o subescala son manifestaciones de un constructo subyacente y, por lo tanto, se espera que estén correlacionados. En contraste, si los ítems de una escala o subescala no se asumen como correlacionados (es decir, se sigue un modelo formativo), los análisis correspondientes a este punto no son fundamentales y pueden ser omitidos. En aquellos casos donde se desconozca el tipo de modelo, los revisores deberán determinar, basándose en el contenido del instrumento, si este se adhiere a un modelo reflexivo o formativo (Fayers et al., 1997).

Punto 7: Evaluación de propiedades de medición adicionales

Este apartado se centra en la evaluación de las propiedades de medición restantes, que incluyen la fiabilidad, el error de medición, la validez de criterio, la prueba de hipótesis (validez de constructo) y la sensibilidad al cambio. La valoración de estas propiedades debe seguir los tres pasos metodológicos previamente descritos (3.1.2.2 Evaluación de las propiedades de medición) (Prinsen et al., 2018).

A diferencia de la validez de contenido y la estructura interna, la evaluación de estas propiedades de medición proporciona información sobre la calidad global de la escala o subescala en su conjunto, en lugar de centrarse en el nivel de ítems individuales (Prinsen et al., 2018).

3.1.2.3 Selección del cuestionario

Este apartado, que comprende los puntos 8 a 10, aborda la evaluación de la interpretabilidad y viabilidad de los cuestionarios. Asimismo, se encarga de la formulación de recomendaciones y de la presentación de la información pertinente sobre la revisión sistemática.

Punto 8: Descripción de la interpretabilidad y la viabilidad de los cuestionarios

La interpretabilidad es la capacidad de asignar un significado cualitativo (connotaciones clínicas o de comprensión común) a las puntuaciones cuantitativas de un cuestionario o a sus cambios. Por su parte, la viabilidad se refiere a la facilidad de aplicación del cuestionario en su entorno previsto, considerando las limitaciones temporales o económicas. Este concepto engloba aspectos como el coste del instrumento, su duración y la simplicidad o tipo de administración (Prinsen et al., 2016). La viabilidad es relevante tanto para las personas que completan el cuestionario (en el caso de autoadministración) como para los investigadores o clínicos que lo administran.

Cabe destacar que la interpretabilidad y la viabilidad no son consideradas propiedades de medición en sentido estricto, ya que no reflejan la calidad psicométrica intrínseca de un cuestionario. No obstante, se reconocen como elementos indispensables para su selección. En situaciones donde dos cuestionarios presenten una calidad psicométrica comparable, se recomienda priorizar aquellos que demuestren mayor viabilidad, es decir, que sean factibles en términos de tiempo de administración y coste (Prinsen et al., 2016).

Punto 9: Formulación de recomendaciones para cuestionarios

Las recomendaciones sobre el cuestionario más idóneo para la aplicación en evaluaciones se formulan considerando el constructo de interés y la población de estudio específica. Para garantizar una recomendación basada en la evidencia y totalmente transparente, se propone clasificar los cuestionarios evaluados en tres categorías distintivas (De Vet et al., 2011):

- a) Cuestionarios con potencial de recomendación óptima: Incluye instrumentos que demuestran evidencia de validez de contenido suficiente (independientemente del nivel de calidad metodológica) y, al menos, evidencia de consistencia interna suficiente, incluso si esta es limitada.

- b) Cuestionarios con potencial condicional de recomendación: Abarca aquellos instrumentos que no se clasifican ni en la categoría "a" ni en la "c", lo que sugiere la necesidad de estudios de validación adicionales para consolidar su recomendación.
- c) Cuestionarios no recomendables: Comprende los instrumentos para los cuales existe evidencia de alta calidad que indica una propiedad de medición insuficiente.

Es importante justificar la asignación de cada cuestionario a su categoría correspondiente y, si aplica, proporcionar directrices para futuras investigaciones de validación. Asimismo, se aconseja ofrecer una recomendación sobre el cuestionario más adecuado. Esta recomendación no solo debe basarse en la evaluación de las propiedades de medición, sino que también puede incorporar consideraciones de interpretabilidad y viabilidad (Prinsen et al., 2016).

Punto 10: Información detallada de la revisión sistemática

Conforme a las directrices establecidas en la Declaración PRISMA (Moher et al., 2010), se recomienda la inclusión de la siguiente información esencial en el informe de la revisión sistemática:

1. Resultados de la búsqueda y selección: Presentación de los resultados de la búsqueda bibliográfica y el proceso de selección de estudios e instrumentos mediante el diagrama de flujo PRISMA. Esto debe incluir el número final de artículos y de instrumentos incorporados en la revisión.
2. Características de los instrumentos: Detalle de las propiedades de los instrumentos incluidos, tales como su denominación, los constructos que evalúan, la población de desarrollo original, los contextos de uso previstos, el idioma, el número de subescalas, la cantidad de ítems, las opciones de respuesta, el período de recuperación y aspectos relevantes de interpretabilidad y viabilidad.
3. Características de las poblaciones de estudio: Descripción de las poblaciones de estudio abarcando su ubicación geográfica, idioma, patología, población objetivo, tamaño muestral, edad, sexo, entorno de aplicación y país.
4. Calidad metodológica individual: Evaluación de la calidad metodológica de cada estudio por propiedad de medición e instrumento.
5. Tabla resumen de hallazgos: Presentación de una tabla resumen por cada propiedad de medición, que incluya los resultados agrupados o cualitativamente sintetizados, su calificación general (suficiente [+], insuficiente [-], inconsistente [±] o indeterminada [?]) y la calificación de la calidad de la evidencia (alta, moderada, baja o muy baja).

Estas tablas, estructuradas individualmente para cada propiedad de medición, serán herramientas fundamentales para la formulación de recomendaciones y la selección del instrumento más adecuado para un propósito o contexto concretos.

3.2 Criterios de Terwee y Prinsen

Los criterios de Terwee y Prinsen fueron desarrollados con el objetivo de evaluar la calidad de los resultados en estudios que abordan las propiedades de medición de cuestionarios. Estos criterios se aplican a los siguientes dominios: validez de contenido, validez estructural, consistencia interna, fiabilidad, error de medición, prueba de hipótesis (validez de constructo), validez transcultural, validez de criterio y sensibilidad al cambio. Además, su creación tuvo como propósito identificar déficits y brechas en el conocimiento de estas propiedades psicométricas, con el fin de crear futuros estudios de validación (Mokkink et al., 2018b; Prinsen et al., 2016; Prinsen et al., 2018).

Inicialmente, las propiedades sujetas a evaluación por estos criterios incluían la validez de contenido, consistencia interna, validez de criterio, validez de constructo, reproducibilidad (concordancia y fiabilidad), sensibilidad, efecto suelo-techo e interpretabilidad. La evaluación se realizaba mediante cuatro ítems que determinaban su alineación con estándares de calidad metodológica, utilizando una escala de cuatro categorías: "suficiente" (+), "indeterminado" (?), "insuficiente" (-) y "no se ha hallado información" (0) (Terwee et al., 2007).

Actualmente, el sistema de calificación se ha simplificado a tres categorías: "suficiente" (+), "indeterminado" (?) e "insuficiente" (-) para la mayoría de las propiedades psicométricas (Mokkink et al., 2018b; Prinsen et al., 2016; Prinsen et al., 2018). Una excepción es la validez de contenido, que mantiene una evaluación diferenciada basada en cuatro criterios: "suficiente" (+), "inconsistente" ($\pm$), "insuficiente" (-) e "indeterminado" (?) (Terwee et al., 2018) (ver Tabla 1).

Tabla 1. Criterios Terwee y Prinsen.

Propiedad de medición	Puntuación	Criterios
Validez estructural	+	<u>CTT</u> CFA: CFI o TLI o una medición similar >0.95 ó RMSEA <0.06 ó SRMR <0.08 <u>IRT/Rasch</u> No violación de la unidimensionalidad: CFI o TLI o una medición similar >0.95 ó RMSEA <0.06 ó SRMR <0.08 Y No violación de la independencia local: correlaciones residuales entre los ítems después de controlar el factor dominante <0.20 ó Q3's <0.37 Y No violación de la monotonía: gráficos de aspecto adecuado ó escalabilidad de los ítems >0.30 Y Ajuste adecuado del modelo IRT: $\chi^2 >0.001$ Rasch: media cuadrada de infit y aoutfit ≥ 0.5 y ≤ 1.5 o valor Z estandarizado >-2 y <2
	?	CTT: no está disponible toda la información para “+” IRT/Rasch: ajuste del modelo no respaldado
	-	No se cumplen los criterios para “+”
Consistencia interna	+	Al menos baja evidencia para validez estructural suficiente y alfa de Cronbach ≥ 0.70 para cafa escala o subescala unidimensional
	?	No se cumple el criterio de “al menos baja evidencia para validez estructural suficiente”
	-	Al menos baja evidencia para validez estructural suficiente y alfa de Cronbach ≤ 0.70 para cafa escala o subescala unidimensional
Fiabilidad	+	ICC o Kappa ponderado ≥ 0.70

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

	?	ICC o Kappa ponderado no reportado
	-	ICC o Kappa ponderado ≤ 0.70
Error de medición	+	SDC o LoA $< MIC$
	?	MIC no definido
	-	SDC o LoA $> MIC$
Prueba de hipótesis para validez de constructo	+	El resultado concuerda con la hipótesis
	?	No se define la hipótesis
	-	El resultado no concuerda con la hipótesis
Validez tras cultural/invarianza de medición	+	No se encontraron diferencias importantes entre las características de los grupos (como edad, género, idioma) en el análisis factorial múltiple ó no hay DIF importante entre las características del grupo (McFadden's $R^2 < 0.02$
	?	Sin análisis factorial de grupos múltiples o no hay análisis DIF
	-	Diferencias importantes entre las características de los grupos o DIF encontrado
Validez de criterio	+	Correlación con la gold standard ≥ 0.70 o AUC ≥ 0.70
	?	No se encuentra toda la información para “+”
	-	Correlación con la gold standard < 0.70 o AUC < 0.70
Sensibilidad	+	Los resultados concuerdan con la hipótesis o AUC ≥ 0.70
	?	No se define la hipótesis
	-	Los resultados no concuerdan con la hipótesis o AUC < 0.70

Fuente: Elaboración propia a partir de la traducción de Prinsen et al. (2018).

3.2.1 Propiedades psicométricas: calidad de los resultados

3.2.1.1 Validez de contenido

La validez de contenido aborda la medida en que los conceptos de interés son representados exhaustivamente por los ítems de un cuestionario (Mokkink et al., 2010b). Para asegurar la calidad de un instrumento, los autores deben proporcionar una descripción clara de los siguientes aspectos relativos a su desarrollo:

- **Objetivo de medición del cuestionario:** El objetivo puede ser discriminativo, evaluativo o predictivo. Esta distinción es crucial, ya que la validez de los ítems puede variar según el propósito de medición (Terwee et al., 2007).
- **Población objetivo:** Una descripción precisa de la población para la cual se desarrolló el cuestionario es fundamental. Esto permite determinar el alcance y la aplicabilidad del instrumento en otras poblaciones (Terwee et al., 2007).
- **Conceptos propuestos para la medición:** Es esencial que los autores definan un marco conceptual claro sobre el constructo general que se pretende medir. Los conceptos relevantes pueden categorizarse como síntomas, funcionamiento (físico, psicológico o social), percepciones generales de salud o calidad de vida. Es importante que estos distintos niveles de resultados se diferencien y se midan mediante subescalas independientes (Terwee et al., 2007).
- **Selección y reducción de ítems:** Los métodos empleados para la selección y reducción de ítems, así como la realización de un estudio piloto para examinar la legibilidad y comprensión, deben ser justificados e informados. Los ítems del cuestionario deben reflejar áreas consideradas importantes por la población objetivo, por lo que la participación de esta población durante la fase de selección de ítems es crucial (Terwee et al., 2007).
- **La interpretabilidad de los ítems es crucial.** Un cuestionario no debería requerir habilidades de lectura superiores a las de un estudiante de primaria para minimizar los valores perdidos y las respuestas inconsistentes. Por lo tanto, los ítems deben ser concisos, simples, y libres de terminología compleja o preguntas dobles. Además, el periodo de tiempo al que se refieren las preguntas debe estar claramente definido y justificado (Terwee et al., 2007).

Para evaluar la validez de contenido de un cuestionario, se sigue un proceso de tres pasos:

- **Calificación del resultado de estudios individuales:** Inicialmente, cada resultado de un estudio específico sobre el desarrollo de un

cuestionario y su validez de contenido se califica empleando los diez criterios establecidos para una buena validez de contenido. En esta fase, los revisores evalúan el contenido del cuestionario directamente (Terwee et al., 2018).

- En la segunda fase del proceso de evaluación, se procede a la síntesis cualitativa de los resultados de todos los estudios disponibles. El objetivo es determinar si, en su conjunto, la relevancia, exhaustividad, comprensibilidad y la validez general del contenido del cuestionario pueden calificarse como "suficiente" (+), "insuficiente" (-), "inconsistente" ($\pm$) o "indeterminada" (?), considerando la totalidad de la evidencia existente. Cabe señalar que, en esta etapa, el foco de la evaluación se traslada del estudio individual al cuestionario en sí; a diferencia del paso previo, donde la atención se centraba en la calidad metodológica de los estudios particulares. Estas calificaciones globales se registran en la penúltima columna de la tabla resumen y se asignan como "suficiente" (+), "insuficiente" (-) o "inconsistente" ($\pm$). Una calificación global de "indeterminada" (?) no es admisible en esta fase, ya que la valoración del revisor siempre resultará en una de las categorías definidas: "suficiente" (+), "insuficiente" (-) o "inconsistente" ($\pm$) (Terwee et al., 2018).
- En la tercera y última fase, las calificaciones globales establecidas en el paso previo se complementan con una calificación de la calidad de la evidencia. Para ello, se emplea el método GRADE modificado, que asigna niveles de calidad como "alto", "moderado", "bajo" o "muy bajo". Este proceso es crucial para corroborar la fiabilidad de las calificaciones generales (Terwee et al., 2018). La metodología GRADE modificada tiene como propósito degradar la evidencia cuando existen dudas sobre su calidad. El punto de partida es siempre la suposición de una calificación global de alta calidad. Posteriormente, esta calidad de la evidencia se degrada en uno o dos niveles por cada factor adverso, resultando en clasificaciones de "moderada", "baja" o "muy baja". Esta degradación ocurre cuando:
 - La calidad metodológica de los estudios es dudosa o inadecuada (indicando un riesgo de sesgo).
 - Existe inconsistencia inexplicable en los resultados.
 - Se dispone únicamente de resultados indirectos.

En el método GRADE modificado para la evaluación de la validez de contenido, no se incorporan los factores de imprecisión ni sesgo de publicación. Esta exclusión se justifica porque la imprecisión resulta menos relevante en este contexto, dado que el desarrollo del cuestionario y los estudios de validez de contenido se inscriben predominantemente en el ámbito de la investigación cualitativa. Por otro lado, la evaluación del sesgo de publicación se torna

compleja debido a la ausencia de registros sistemáticos para el desarrollo de nuevas herramientas y a la escasez de investigaciones específicas sobre validez de contenido (Terwee et al., 2018).

Según los criterios de Terwee y Prinsen, para que un cuestionario alcance una calificación de validez de contenido "suficiente", es imprescindible que la revisión sistemática reporte la siguiente información:

- Características del cuestionario: Incluyendo el constructo medido, la población objetivo, el contexto de uso previsto, el número de escalas e ítems, y el período de recuperación, entre otros.
- Calidad del desarrollo del cuestionario: Detalle del proceso de creación del instrumento.
- Calidad y características de los estudios de validez de contenido disponibles: Información sobre la población del estudio, tamaño de la muestra, diseño del estudio, y la participación de pacientes y/o profesionales.
- Calidad asociada de la evidencia: Evaluación de la solidez de la evidencia.
- Calificación de la relevancia, exhaustividad, comprensibilidad y validez general de contenido por cada cuestionario.

En ausencia de una descripción clara y completa de estos elementos, la validez de contenido del cuestionario se considerará inconsistente (Terwee et al., 2018).

3.2.1.2 Validez estructural

La validez estructural define el grado en que las puntuaciones de un instrumento reflejan adecuadamente una dimensión del constructo que se pretende medir. Su evaluación se realiza típicamente mediante análisis factorial o análisis basados en la Teoría de Respuesta al Ítem (Item Response Theory - IRT) o el modelo Rasch (Mokkink et al., 2010b).

En el contexto de una revisión sistemática, es fundamental diferenciar entre estudios que emplean el análisis factorial para evaluar la validez estructural general de la herramienta y aquellos que lo utilizan para determinar la unidimensionalidad de cada subescala de forma independiente.

Para evaluar la validez estructural, el Análisis Factorial Exploratorio (Exploratory Factor Analysis - EFA) se aplica a la totalidad de los ítems del instrumento. Este análisis busca validar el número hipotético de subescalas de la herramienta y la agrupación de ítems dentro de estas subescalas (referido como estudios de validez estructural). Por otro lado, la unidimensionalidad por subescala se evalúa mediante análisis factoriales múltiples aplicados a los ítems de cada subescala de forma separada. El propósito es determinar si cada subescala, por sí misma,

mide un único constructo (conocidos como estudios de unidimensionalidad) (Prinsen et al., 2018).

Para que la calificación de esta propiedad psicométrica sea considerada suficiente, deben cumplirse los siguientes criterios, basados en la Teoría Clásica de los Tests (Test Classic Theory - TCT) o la Teoría de Respuesta al Ítem (IRT) / Modelo de Rasch:

- Teoría Clásica de los Tests (TCT) - Análisis Factorial Confirmatorio (Confirmatory Factor Analysis - CFA):
 - Índices de ajuste comparativo (CFI (Comparative Fit Index) o TLI (Tucker-Lewis Index)) o medidas comparables > 0.95 .
 - Error cuadrático medio de aproximación (RMSEA (Root Mean Square Error of Approximation)) < 0.06 .
 - Raíz cuadrada media residual estandarizada (SRMR (Standardized Root Mean Square Residual)) < 0.08 (US Department of Health, 2006).
- Teoría de Respuesta al Ítem (IRT) / Modelo de Rasch:
 - Unidimensionalidad: Ausencia de violación de la unidimensionalidad, lo que se evalúa también mediante índices de ajuste como CFI o TLI > 0.95 , o RMSEA < 0.06 , o SRMR < 0.08 .
 - Independencia Local: Ausencia de violación de la independencia local, indicada por correlaciones residuales entre los ítems (después de controlar el factor dominante) < 0.20 o Q3's < 0.37 (Mokkink et al., 2018b; Prinsen et al., 2018).

3.2.1.3 Consistencia interna

La consistencia interna mide el grado en que los ítems de una escala o subescala están intercorrelacionados, indicando que miden un mismo concepto subyacente (Mokkink et al., 2010b). Esta propiedad psicométrica es crucial para cuestionarios diseñados para evaluar un único constructo a través de múltiples ítems.

Una escala con una alta consistencia interna (homogénea o unidimensional) se logra mediante una definición precisa del constructo, la selección de ítems idóneos, y la aplicación de EFA y CFA (De Vet et al., 2005).

El método estadístico más empleado para determinar la consistencia interna es el Alfa de Cronbach (α). Un valor bajo de Alfa de Cronbach sugiere una escasa correlación entre los ítems de una escala, lo que invalida la suma de sus puntuaciones. Por el contrario, un Alfa de Cronbach excesivamente alto (valores

cercanos a 1) indica correlaciones elevadas entre los ítems, lo que puede implicar redundancia o duplicación de contenido (Streiner, 2003).

Según los criterios de Terwee y Prinsen, una calificación de "suficiente" para la consistencia interna requiere al menos una evidencia de bajo nivel de validez estructural suficiente y un Alfa de Cronbach ≥ 0.70 para cada escala o subescala unidimensional (Mokkink et al., 2018b; Prinsen et al., 2018).

3.2.1.4 Fiabilidad

La fiabilidad se define como la estabilidad de un instrumento al realizar mediciones repetidas (Mokkink et al., 2010b). Desde una perspectiva estadística, el Coeficiente de Correlación Intraclass (Intraclass Correlation Coefficient - ICC) es el índice más apropiado y utilizado para cuantificar la concordancia entre mediciones sucesivas de una variable. El ICC representa la proporción de la variación observada en la población, dividida por la variación real entre individuos, en relación con la variación total (que incluye la variación inter-individual y el error de medición intra-individual), y se expresa en un rango de 0 a 1 (Henson, 2001; Streiner et al., 2008).

El intervalo temporal entre administraciones repetidas del instrumento debe ser lo suficientemente prolongado para evitar efectos de recuerdo, pero a su vez lo suficientemente corto para asegurar que no se hayan producido cambios clínicos significativos en el constructo medido. Generalmente, un periodo de una a dos semanas es idóneo. No obstante, pueden existir justificaciones para variaciones de este periodo, como la administración de una intervención en el intervalo, cambios clínicos en un número considerable de personas, influencias del entorno de administración del instrumento, o instrucciones de cumplimentación poco claras o insuficientes (Mokkink et al., 2010c; Mokkink et al., 2018b; Terwee et al., 2007).

Por consiguiente, los criterios de Terwee y Prinsen no evalúan la idoneidad del periodo de tiempo en sí, sino que requieren que este periodo esté claramente descrito y justificado. Otorgan una calificación de "suficiente" cuando el ICC o el Kappa ponderado es ≥ 0.70 (Mokkink et al., 2018b; Prinsen et al., 2018).

3.2.1.5 Error de medición

El error de medición se define como la variabilidad sistemática y aleatoria en la puntuación de una persona que no es atribuible a cambios reales en el fenómeno que se está midiendo (Mokkink et al., 2010b). Para evaluar adecuadamente un buen error de medición, se requiere información sobre el cambio detectable más pequeño (Smallest Detectable Change - SDC) o los límites de acuerdo (Limits of Agreement - LoA), además del cambio mínimo importante (Minimum Important Change - MIC).

El MIC se calcula de forma óptima a partir de múltiples estudios y mediante el uso de anclajes diversos (Crosby et al., 2003; de Vet et al., 2006; Revicki et al., 2008; Yost et al., 2011). En ausencia de información suficiente para determinar si el SDC o el LoA son menores que el MIC, se recomienda simplemente reportar la información disponible sobre el SDC o el LoA, sin realizar una calificación de la calidad de la evidencia.

Según los criterios de Terwee y Prinsen, la calificación de esta propiedad se considera "suficiente" cuando el SDC o el LoA son menores que el MIC (estos datos pueden obtenerse de estudios diferentes) (Mokkink et al., 2018b; Prinsen et al., 2018).

3.2.1.6 Prueba de hipótesis para la validez de constructo

La validez de constructo se refiere al grado en que las puntuaciones de una escala se correlacionan con otras medidas de manera consistente con las hipótesis teóricamente derivadas sobre los conceptos que se están evaluando (Mokkink et al., 2010b). La evaluación de esta validez debe realizarse mediante la comprobación de hipótesis predefinidas, las cuales deben ser tan específicas como sea posible. La ausencia de hipótesis explícitas aumenta significativamente el riesgo de sesgo, ya que podría inducir a la formulación retrospectiva de explicaciones erróneas para correlaciones bajas, en lugar de concluir una posible falta de validez del cuestionario (Terwee et al., 2007).

Según los criterios de Terwee y Prinsen, se otorga una calificación de "suficiente" a la validez de constructo si las hipótesis se especifican previamente y si al menos el 75% de los resultados empíricos son congruentes con dichas hipótesis (Mokkink et al., 2018b; Prinsen et al., 2018).

3.2.1.7 Validez transcultural / Invarianza de medición

Este concepto se refiere al grado en que el rendimiento de los ítems de una escala que ha sido traducida o culturalmente adaptada refleja de manera adecuada su versión original (Mokkink et al., 2010b). Es necesario evaluar esta propiedad cuando un instrumento se utiliza o se va a utilizar, en poblaciones con diferencias culturales (incluyendo etnia, idioma, género o grupos de edad) (Mokkink et al., 2018b). Esta evaluación se realiza cuantificando el Funcionamiento Diferencial de Ítems (Differential Item Function - DIF) mediante análisis de regresión logística o Análisis Factorial Confirmatorio Multigrupo (Multi-Group Confirmatory Factor Analysis - MGCF), siempre que la estructura factorial y las cargas factoriales sean equivalentes entre los grupos.

La invarianza de medición y la ausencia de DIF indican si los evaluados de distintos grupos, que poseen el mismo nivel del rasgo latente (ajustado por las diferencias de grupo), responden de forma similar a un ítem concreto (Embretson

et al., 2013). Terwee y Prinsen asignan una puntuación de "suficiente" a esta propiedad psicométrica si no se identifican diferencias significativas entre los factores de grupo (como edad, género, idioma) en el análisis factorial multigrupo, o si no se hallan DIF significativos para los factores de grupo (indicado por un R^2 de McFadden's < 0.02) (Mokkink et al., 2018b; Prinsen et al., 2018).

3.2.1.8 Validez de criterio

La validez de criterio se define como el grado en que las puntuaciones obtenidas de un instrumento específico se correlacionan con las de un criterio de referencia ("gold standard") (Mokkink et al., 2010b). De acuerdo con los criterios de Terwee y Prinsen, esta propiedad recibe una calificación de "suficiente" si la correlación con el gold standard es ≥ 0.70 o si el Área Bajo la Curva (Area Under the Curve - AUC) es ≥ 0.70 (Mokkink et al., 2018b; Prinsen et al., 2018).

3.2.1.9 Sensibilidad al cambio

La capacidad de detección de cambio (responsiveness) de un instrumento se define como su habilidad para identificar modificaciones mínimas y clínicamente significativas en el constructo medido a lo largo del tiempo (Mokkink et al., 2010b). Esta propiedad representa una medida de validez longitudinal y, por lo tanto, su evaluación se realiza mediante la verificación de hipótesis predefinidas. Este enfoque confirma la capacidad del instrumento para registrar cambios cuando estos realmente han ocurrido (Terwee et al., 2033).

Adicionalmente, el instrumento debe ser capaz de diferenciar entre cambios clínicamente importantes y el error de medición. Para ello, la capacidad de respuesta debe evaluarse relacionando el SDC con el MIC. Alternativamente, puede emplearse el AUC de las características operativas del receptor (Area Under the Curve / Receiver Operating Characteristic - AUC/ROC) (Deyo et al., 1986). Este último método permite distinguir entre personas que han experimentado un cambio y aquellos que no, según un criterio externo.

Conforme a los criterios de Terwee y Prinsen, una capacidad de respuesta se califica como suficiente si se obtiene un AUC ≥ 0.70 y al menos el 75% de los resultados concuerdan con las hipótesis planteadas (Mokkink et al., 2018b; Prinsen et al., 2018).

3.3. Método GRADE

El acrónimo GRADE (Grading of Recommendations, Assessment, Development and Evaluations) señala un marco metodológico transparente diseñado para el desarrollo y la presentación de resúmenes de evidencia científica, así como para la articulación sistemática de recomendaciones en la práctica clínica (Guyatt et al., 2008b; Guyat et al., 2008c; Siemieniuk et al., 2019). Constituye un enfoque estandarizado empleado por revisores sistemáticos y desarrolladores de guías clínicas con el fin de evaluar la calidad de la evidencia disponible y fundamentar las decisiones relativas a la recomendación de intervenciones (Guyatt et al., 2008c).

GRADE se distingue de otras herramientas de evaluación por tres características fundamentales:

- Diferenciación entre calidad de la evidencia y fuerza de la recomendación: Este marco establece una clara distinción conceptual entre la calidad de la evidencia que sustenta una conclusión y la fuerza con la que se emite una recomendación.
- Evaluación de resultado por resultado: La calidad de la evidencia se valora de manera individual para cada resultado de interés, permitiendo una descomposición al detalle en la evaluación que otras metodologías no ofrecen.
- Posibilidad de aumento de calidad en estudios observacionales: A diferencia de enfoques más restrictivos, GRADE permite "elevar la calidad de la evidencia" de estudios observacionales si estos cumplen con criterios metodológicos específicos que mitigan el riesgo de sesgo.

El marco GRADE se ha consolidado como la metodología predominante para la evaluación de estudios controlados y la formulación de recomendaciones en revisiones sistemáticas y guías de práctica clínica (Balshem et al., 2011; Guyatt et al., 2008c; Guyatt et al., 2011a; Guyatt et al., 2011b; Guyatt et al., 2011c; Guyatt et al., 2011d; Guyatt et al., 2011e; Guyatt et al., 2011f; Guyatt et al., 2011g; Guyatt et al., 2011h; Langer et al., 2012). La comprensión y aplicación efectiva de este método suele requerir la revisión de una serie de publicaciones especializadas y, en ocasiones, la participación en talleres dedicados a adquirir el conocimiento suficiente que permita una evaluación crítica. Cabe señalar la ausencia de explicaciones concisas sobre GRADE dirigidas a una audiencia de investigadores generalistas, así como la falta de textos que lo aborden de manera comprensiva (Balshem et al., 2011; Guyatt et al., 2011a; Guyatt et al., 2011b; Guyatt et al., 2011c; Guyatt et al., 2011d; Guyatt et al., 2011e; Guyatt et al., 2011f; Guyatt et al., 2011g; Guyatt et al., 2011h; Langer et al., 2012).

Es importante reconocer la naturaleza inherentemente subjetiva del método GRADE. Su implementación no puede ser puramente mecánica; cada decisión

incorpora una considerable dosis de juicio subjetivo. En consecuencia, es razonable que dos evaluadores que analizan el mismo conjunto de evidencias puedan llegar a conclusiones diferentes. No obstante, a pesar de esta subjetividad, GRADE proporciona un marco reproducible y transparente para la calificación de la certeza de la evidencia (Mustafa et al., 2013).

3.3.1 Evaluación de la calidad de la evidencia y elaboración de la recomendación

El marco metodológico GRADE puede expresarse mediante los siguientes cinco puntos:

1. Niveles de certeza de la evidencia

GRADE categoriza la certeza, o calidad, de la evidencia en cuatro niveles jerárquicos: "muy bajo", "bajo", "moderado" y "alto". La asignación inicial de la certeza se realiza a priori: los ensayos clínicos aleatorizados (ECAs) reciben una calificación inicial de "alto", mientras que los estudios observacionales se designan con una calificación inicial de "bajo". Esta diferenciación inicial obedece a la menor susceptibilidad de los ECAs al sesgo, en comparación con los estudios observacionales (Howick, 2011; Odgaard-Jensen et al., 2011).

2. Modificación del nivel de certeza: disminución o aumento

La segunda fase implica una clasificación inicial para determinar la posible "disminución" o "aumento" del nivel de certeza. Es frecuente que tanto los ECAs como los estudios observacionales experimenten una disminución en su nivel de certeza debido a la presencia de sesgos identificables. De manera recíproca, los estudios observacionales pueden ver su nivel de certeza incrementado cuando diversas investigaciones de alta calidad presentan resultados consistentemente congruentes (Ioannidis, 2022).

- Factores que disminuyen la calidad de la evidencia en GRADE

El marco GRADE identifica diversos dominios que pueden reducir la calidad de la evidencia: riesgo de sesgo, inconsistencia, evidencia indirecta, imprecisión y sesgo de publicación. Para cada uno de dichos factores, los evaluadores poseen la opción de disminuir el nivel de certeza en uno o dos grados (GRADE Handbook, 2013) (ver Tabla 2).

- Riesgo de sesgo (limitaciones en diseño y ejecución)

El riesgo de sesgo se produce cuando los resultados de un estudio no reflejan fielmente la realidad debido a limitaciones en su diseño y/o ejecución, o cuando

la calidad de estudios adicionales sobre validez de contenido es cuestionable (Guyatt et al., 2011h).

En la práctica, la magnitud del impacto de los sesgos potenciales en los resultados es difícil de cuantificar. Dado que GRADE se emplea para puntuar el conjunto de la evidencia en relación con un resultado específico (en lugar de estudios individuales), los evaluadores deben discriminar si el riesgo de sesgo en los estudios individuales es lo suficientemente sustancial como para disminuir la confianza en el efecto de tratamiento. Algunos factores que pueden introducir sesgo incluyen: la ausencia de una secuencia de asignación claramente aleatorizada, la falta de cegamiento, la deficiente ocultación de la asignación, la priorización de tratamientos sobre las evaluaciones de resultados, la interrupción prematura de la investigación o tasas elevadas de pérdidas durante el seguimiento (Guyatt et al., 2011h).

- Inconsistencia de los resultados

La inconsistencia de los resultados surge ante una variabilidad significativa e inexplicable entre los hallazgos de diferentes ensayos. La calidad de la evidencia se considera superior cuando múltiples estudios presentan efectos consistentes. Al evaluar si la certeza debe ser degradada debido a la inconsistencia, los autores deben analizar la similitud de las estimaciones puntuales y la superposición de sus intervalos de confianza, además de considerar los criterios estadísticos de heterogeneidad (Guyatt et al., 2011d).

- Ausencia de evidencia directa/evidencia indirecta

La ausencia de evidencia directa se manifiesta cuando los estudios no comparan directamente las intervenciones de interés en la población objetivo o no informan sobre los resultados que son críticos para la toma de decisiones. Este factor surge ante diferencias en la población (afectando la aplicabilidad), en las intervenciones (afectando también la aplicabilidad), en las medidas de resultado (utilización de resultados sustitutos), o cuando se establecen comparaciones indirectas (es decir, en ausencia de comparaciones directas entre dos o más intervenciones de interés) (Guyatt et al., 2011c).

- Imprecisión

La imprecisión se presenta cuando la amplitud de los intervalos de confianza afecta la calidad de los datos. Para calificar la imprecisión, GRADE se centra en el intervalo de confianza del 95% alrededor de la mejor estimación del efecto absoluto y el número de sujetos o eventos incluidos (Guyatt et al., 2011b). Es importante destacar que la imprecisión tiene una relevancia menor para la validez de contenido, dado que el desarrollo de nuevas herramientas y los

estudios de validez de contenido se enmarcan predominantemente dentro de la investigación cualitativa (Terwee et al., 2018).

- Sesgo de publicación

El sesgo de publicación representa posiblemente el dominio más desafiante dentro del marco GRADE, ya que exige inferencias sobre evidencia no publicada. Consecuentemente, su evaluación se complica debido a la escasez de estudios sobre el desarrollo de nuevas herramientas y la limitada investigación disponible sobre la validez de contenido (Terwee et al., 2018). Este tipo de sesgo es más prevalente en contextos donde existen conjuntos de ensayos de tamaño reducido, con resultados positivos y financiados por la industria (Guyatt et al., 2011e).

Tabla 2.: Método GRADE modificado

Calidad de la evidencia	Disminuir si
Alta	<u>Riesgo de Sesgo</u>
Moderada	-1 Serio
Baja	-2 Muy serio
Muy baja	-3 Extremadamente serio
	<u>Inconsistencia</u>
	-1 Serio
	-2 Muy serio
	<u>Imprecisión</u>
	-1 n total entre 50 y 100
	-2 n total menor de 50
	<u>Evidencia indirecta</u>
	-1 Serio
	-2 Muy serio

Fuente: Elaboración propia a partir de la traducción de GRADE Handbook (2013).

- Factores que aumentan la calidad de la evidencia en GRADE

Así como hay elementos que pueden disminuir la calidad de la evidencia, también hay factores que la incrementan dentro del marco GRADE. Estos incluyen (Guyatt et al., 2011g):

- Fuerza de asociación: Este factor se presenta cuando la magnitud del efecto observado es tan considerable que los sesgos inherentes a los estudios observacionales no pueden, por sí solos, justificar el resultado.
- Gradiente dosis-respuesta: Se refiere a la observación de que el resultado varía proporcionalmente con el grado o nivel de exposición a una intervención.
- Factores de confusión que reducen el tamaño del efecto: En la mayoría de los casos, la presencia de factores de confusión no detectados o mal ajustados tiende a disminuir la confusión residual, en lugar de aumentar la magnitud del efecto observado (Guyatt et al., 2011g).

3. Asignación de la calificación final de certeza de la evidencia

En la tercera etapa, se procede a asignar una puntuación final a la calidad de la evidencia para todos los resultados considerados de gran importancia. Esta calificación se establece en uno de los cuatro niveles: "alta", "moderada", "baja" o "muy baja" (Guyatt et al., 2008b). Este proceso implica un juicio reflexivo que integra aumentos y disminuciones en la calidad de la evidencia. Como norma general, se ajusta la categoría en un nivel por cada factor. Por ejemplo, la calidad de la evidencia para un resultado de un ECA, que al inicio se califica como "alta", podría descender a "moderada" si existe un alto riesgo de sesgo en los estudios incluidos; y podría degradarse aún más a "baja" si se observa una heterogeneidad inexplicable significativa entre los ensayos, y llegar a "muy baja" si un número escaso de eventos resulta en intervalos de confianza que abarcan tanto beneficios como daños importantes (Guyatt et al., 2011b; Guyatt et al., 2011c; Guyatt et al., 2011d; Guyatt et al., 2011e; Guyatt et al., 2011h) (ver Tabla 3).

Tabla 3. Definición de los niveles de calidad de la evidencia de GRADE.

Nivel de calidad	Definición
Alta	Estamos muy seguros de que la verdadera propiedad de medición se encuentra cerca de la estimación de la propiedad de medición.
Moderada	Tenemos una confianza moderada en la estimación de la propiedad de medición: es probable que la propiedad de medición real esté cerca de la estimación de la propiedad de medición, pero existe la posibilidad de que sea ligeramente diferente.

Baja	Nuestra confianza en la estimación de la propiedad de medición es limitada: la propiedad de medición real puede ser bastante diferente de la estimación de la propiedad de medición
Muy baja	Tenemos muy poca confianza en la estimación de la propiedad de medición: es probable que la verdadera propiedad de medición sea muy diferente de la estimación de la propiedad de medición.

Fuente: Elaboración propia a partir de la traducción de Prinsen et al. (2018).

4. Consideración de otros factores en la formulación de recomendaciones

El cuarto punto subraya que la formulación de recomendaciones no debe basarse exclusivamente en la alta calidad de la evidencia. Una evidencia de alta calidad no siempre se traduce en una recomendación firme. Las recomendaciones deben integrar otros factores, más allá de la calidad global de la evidencia, como el equilibrio entre los efectos deseables y no deseables, los valores y preferencias de los usuarios (particularmente cuando la relación riesgo/beneficio no es clara), y, finalmente, los recursos y costes asociados (Balslem et al., 2011; Guyatt et al., 2008a).

5. Formulación de recomendaciones: fuerza y dirección

Finalmente, el quinto punto se centra en la formulación de recomendaciones, clasificándolas como fuertes o débiles, y a favor o en contra de un tratamiento. Esta decisión se basa en la ponderación de los cuatro factores previamente detallados: el equilibrio entre efectos deseables y no deseables, la calidad global de la evidencia, los valores y preferencias de los usuarios en relación con el balance riesgo/beneficio, y los recursos/costes implicados (Balslem et al., 2011; Guyatt et al., 2008a).

Una recomendación fuerte sugiere que la mayoría de los individuos optarían por la intervención recomendada. En contraste, una recomendación débil implica que es probable que exista duda en las decisiones que las personas podrían tomar (Guyatt et al., 2008a).

La fuerza de estas recomendaciones es crucial para la práctica clínica: una recomendación débil subraya la necesidad de un proceso de toma de decisiones compartido. Al contrario, una recomendación fuerte generalmente indica que no es indispensable presentar ambas opciones de tratamiento al usuario (Guyatt et al., 2008a).

Las recomendaciones son más propensas a ser clasificadas como débiles, en lugar de fuertes. Esto ocurre cuando la certeza de la evidencia es baja y hay un equilibrio entre las consecuencias deseables e indeseables de una intervención, se observa una variación sustancial o incertidumbre en los valores y preferencias de los usuarios, o cuando las intervenciones demandan recursos considerables (Alonso-Coello et al., 2016a; Alonso-Coello et al., 2016b; Guyatt et al., 2008a; Guyatt et al., 2008c).

En cuanto a la redacción de estas recomendaciones, GRADE propone una terminología específica: para las recomendaciones fuertes, se emplea la expresión "se recomienda"; mientras que, para las recomendaciones débiles, se utiliza el término "se sugiere" (Guyatt et al., 2008a).

4. Validación de escalas de evaluación

4.1. Traducción y adaptación transcultural de cuestionarios

4.1.1 Guía ISPOR

La Guía ISPOR (International Society for Pharmacoeconomics and Outcomes Research) representa un referente fundamental en la investigación de resultados en salud, especialmente en lo que respecta a la traducción y adaptación transcultural de escalas de evaluación. Su implementación es crucial para garantizar la validez y fiabilidad de estas herramientas en distintos contextos culturales e idiomáticos. Al proporcionar un marco metodológico riguroso, la Guía ISPOR minimiza los sesgos y las inconsistencias que podrían surgir al transferir un instrumento de medida de una lengua y cultura a otra (Acquadro et al., 2008).

El origen de esta guía para las traducciones y adaptaciones transculturales se remonta al Segundo Congreso Europeo Anual de ISPOR de 1998, donde se crearon diversos grupos de trabajo e investigación; uno de ellos fue el grupo de Traducción y Adaptación Cultural (TAC), que se reunió por primera vez en el Tercer Congreso Europeo Anual de ISPOR, en 1999. En dicha reunión pusieron en común la necesidad de una metodología común que permitiese la traducción y adaptación cultural de PROMs, debido a que (Wild et al., 2005):

- Cada grupo de investigación empleaba metodologías diferentes para realizar la misma tarea, dificultando lograr coherencia y comparación.
- Se empleaba gran variedad de terminología para referirse a los mismos aspectos de traducción, impidiendo la claridad.
- Los autores que desarrollaron un instrumento establecieron sus propias pautas de traducción de un instrumento concreto.
- Las escalas mal traducidas sesgaban la validez de los datos de la investigación (Wild et al., 2005).

Las primeras reuniones del grupo TAC se orientaron hacia la evaluación de las normas de traducción existentes, detectando los siguientes problemas (Wild et al., 2005):

- Falta de coherencia terminológica.
- Falta de coherencia metodológica.
- El alcance de la futura normativa.
- Orientación prescriptiva o descriptiva.
- Necesario definir el público objetivo para determinar cómo presentar los resultados (Wild et al., 2005).

El documento generado determinó que cada proceso de traducción (de idioma a idioma) debía ser independiente del proceso de adaptación al mismo idioma, por ejemplo, la traducción de idioma a idioma puede ser de inglés a español, y la adaptación cultural dependerá si es un español de España, o un español de Argentina (Wild et al., 2005).

Tras ello, el grupo TAC decidió publicar una guía de buenas practicas para la traducción y adaptación cultural, estableciendo las normas básicas, con un lenguaje lo suficientemente claro para que cualquier persona pudiera acceder a ellas y comprenderlas; dicho documento se tituló “Los Principios de Buenas Prácticas para la Traducción y Adaptación Cultural”, y se descompone en las dos partes siguientes (Wild et al., 2005):

Parte 1: Definiciones

Se ofrecen descripciones de los términos más empleados:

- Preparación: Trabajo inicial llevado a cabo de forma previa a comenzar el proceso de traducción.
- Traducción directa: traducción del idioma original al idioma de destino.
- Reconciliación: comparar y fusionar más de una traducción en una sola.
- Retrotraducción: traducción a la nueva versión del idioma original.
- Revisión de la traducción inversa: comparación de las versiones traducidas inversamente, con la versión original y detectar las discrepancias entre ambas.
- Armonización: comparación de traducciones inversas de versiones en varios idiomas y destacar las discrepancias entre la version original y las traducciones.
- Informe cognitivo: prueba del instrumento en un grupo reducido de personas para verificar la redacción, la comprensibilidad y la interpretación del documento traducido.
- Revisión de los resultados del informe cognitivo y finalización: comparación de la interpretación de la traducción por parte del grupo reducido de personas.
- Corrección de pruebas: revisión final de la traducción para corregir errores tipográficos y/o gramaticales.
- Informe final: documento final que documenta el proceso de traducción (Wild et al., 2005).

Parte 2: El proceso de traducción y adaptación cultural

- Paso 1: preparación. Se incluye la explicación de los conceptos, la captación de miembros, contacto con el autor original del instrumento para pedirle permiso e invitarle a participar en el proceso.
- Paso 2: traducción directa por parte de dos o más traductores que deben ser nativos del país que hablase el idioma de destino, y que al menos uno de estos traductores, debe residir en el país del idioma de destino.
- Paso 3: conciliación de las diferentes traducciones directas entre el comité de expertos en el que deben formar parte los traductores directos y los traductores de futuras versiones, permitiendo la coherencia y armonización de las versiones.
- Paso 4: retrotraducción por parte de una persona nativa del idioma original, destacando que, en función de cada situación específica de cada escala, el constructo que evalúe, y la población a la que se dirija, se debe escoger entre un enfoque más literal o conceptual.
- Paso 5: revisión de retrotraducción, destacado como una de las estaciones más importantes en este proceso, llevada a cabo entre varias personas del grupo, y si es posible, solicitando aclaración al autor original de la escala.
- Paso 6: armonización, el cual es un paso esencial para garantizar la validez de la traducción.
- Paso 7: informe cognitivo, aplicando el cuestionario a un número reducido de personas que coincidan con la población diana.
- Paso 8: revisión de resultados del informe cognitivo y finalización, permitiendo garantizar la pertinencia cultural.
- Paso 9: corrección de las versiones finales, atendiendo a errores menores, siendo necesario este último paso a modo de “control de calidad”.
- Paso 10: informe final, reflejando todas las decisiones de traducción y adaptación cultural tomadas a lo largo de este proceso (Wild et al., 2005).

La Guía ISPOR es un pilar fundamental en la investigación de resultados en salud. No es solo un conjunto de directrices para la traducción y adaptación transcultural de cuestionarios de evaluación, sino que es la metodología de referencia actualmente y el primer y crucial paso en el proceso de validación de estas escalas. Al garantizar que un instrumento de evaluación mantenga su equivalencia conceptual, semántica y métrica a través de diferentes culturas e idiomas, la guía ISPOR sienta las bases para estudios rigurosos y comparables. Sin una adaptación transcultural adecuada, los datos obtenidos en nuevas poblaciones podrían carecer de fiabilidad y validez, comprometiendo la interpretabilidad y la aplicabilidad de los resultados. Por lo tanto, adherirse a sus principios es indispensable para cualquier investigador que aspire a generar

evidencia robusta y aplicable globalmente (Acquadro et al., 2008; Wild et al., 2005).

4.2. Propiedades psicométricas de los instrumentos de medida

La psicometría se define como la disciplina científica dedicada al desarrollo y aplicación de métodos, técnicas y teorías para la medición y cuantificación de variables psicológicas inherentes a la cognición humana. Su alcance abarca tanto la elaboración conceptual como la construcción, validación y evaluación de instrumentos de medida, tales como tests, pruebas y otras herramientas estandarizadas, garantizando su validez y fiabilidad. La psicometría incorpora operaciones estadísticas para determinar la adecuación de estos instrumentos en la medición de constructos psicológicos o conductas previamente definidas (Meneses et al., 2013; Tornimbeni et al., 2008).

Los pilares teóricos de la psicometría se estructuran en dos dominios principales: la Teoría de los Tests, que comprende los principios para la construcción, validación y administración de pruebas; y el Escalamiento, que engloba los métodos para la derivación de escalas psicológicas y psicofísicas (Meneses et al., 2013; Tornimbeni et al., 2008).

Dentro de este marco disciplinar, la presente tesis doctoral se fundamenta en la Clinimetría. Este término, introducido por Alvan R. Feinstein en 1982, se refiere a la construcción de instrumentos diseñados para medir o describir signos físicos, síntomas y otros fenómenos clínicos (Feinstein, 1982). Se concibe como una rama esencial de la medicina enfocada en la especificación, identificación y medición de fenómenos clínicos en sujetos, englobando los diversos componentes de la interacción entre el usuario y el profesional sanitario (Fava et al., 2012). Este proceso implica la valoración de un espectro diverso de fenómenos clínicos objetivos, respaldados por evidencia empírica, a través de instrumentos de medición. La información recabada es crucial para establecer diagnósticos precisos, evaluar la progresión patológica, considerar intervenciones terapéuticas y formular pronósticos (Hofman, 2008; Mokkink et al., 2010a).

A pesar de sus especificidades, tanto en clinimetría como en psicometría, las propiedades psicométricas evaluadas en los cuestionarios son análogas. Estas incluyen: validez de contenido, validez estructural, consistencia interna, fiabilidad, error de medición, validez de criterio convergente, validez de criterio, validez transcultural, prueba de hipótesis y sensibilidad al cambio.

4.2.1 Principios y propiedades psicométricas de las escalas

El desarrollo, la construcción y la validación de un cuestionario o escala de medida es un proceso inherentemente complejo que exige una comprensión

teórica profunda del constructo a medir, así como conocimientos avanzados en estadística y el dominio de software especializado. El objetivo fundamental es asegurar que el nuevo instrumento cumpla con criterios adecuados de fiabilidad y validez. Cualquier instrumento de medida debe satisfacer las siguientes características fundamentales (Badia et al., 1998):

1. Viabilidad: Debe ser sencillo de administrar y bien aceptado por usuarios e investigadores, reportando así su viabilidad.
2. Fiabilidad: Ha de demostrar la menor cantidad posible de errores de medición, garantizando la precisión y consistencia de los datos obtenidos.
3. Validez de Contenido: Debe ser pertinente y adecuado para el problema o constructo que se pretende medir.
4. Validez de Constructo: Ha de reflejar fielmente la teoría subyacente al concepto o fenómeno que se busca cuantificar.
5. Sensibilidad al Cambio: Debe ser capaz de detectar variaciones tanto en la respuesta de un mismo individuo a lo largo del tiempo como entre diferentes individuos.

4.2.2 Validez de contenido

La validez de contenido constituye la propiedad psicométrica más importante de un instrumento de medición, definida como la correspondencia entre el contenido de dicho instrumento y el constructo subyacente que se propone cuantificar (Mokkink et al., 2018b; Mokkink et al., 2019). En el ámbito de la salud, la determinación de esta propiedad es inherentemente dependiente del juicio de expertos cualificados, quienes desempeñan un papel crítico tanto en la conceptualización como en la subsiguiente validación de las escalas.

Primera etapa

En la etapa inicial del proceso de validación, el método Delphi emerge como una herramienta metodológica de uso extendido (García-Valdés et al., 2013). Este método se fundamenta en un procedimiento estructurado diseñado para la síntesis sistemática y la recopilación del conocimiento de un panel de expertos. Dicha síntesis se logra mediante una serie secuencial de cuestionarios, complementados con ciclos controlados de retroalimentación de opiniones (Figueroa et al., 2012).

El método Delphi fue concebido originalmente por la RAND Corporation, siendo Helmer y Dalkey quienes, en 1953, implementaron un sistema iterativo que incorporaba un proceso de retroalimentación controlada. Este sistema, denominado "Delphi", en alusión al antiguo oráculo griego, fue diseñado para mitigar las limitaciones inherentes a otros métodos y maximizar los beneficios de los enfoques basados en el juicio de expertos (García-Valdés et al., 2013).

En la literatura científica abundan las publicaciones que abordan el método Delphi; entre ellas, destaca la revisión sistemática de García-Valdés et al. (2013). Esta revisión profundiza en aspectos conceptuales clave, delineando la historia del método, sus ventajas inherentes, aplicaciones diversas, características definitorias y una simplificación de su implementación en el campo de la salud. La metodología Delphi se estructura en tres fases distintivas (García-Valdés et al., 2013):

1. Fase de preparación: Implica la selección rigurosa de los expertos, el diseño y desarrollo del instrumento de consulta, y la deliberación para determinar la modalidad de interacción.
2. Fase de consulta: Abarca los ciclos de consulta con los expertos, el procesamiento estadístico de las respuestas y la realimentación controlada.
3. Fase de aprobación e información de resultados: Concluye con la consolidación de los resultados y su difusión.

El panel Delphi (Gagnier et al., 2021; Mokkink et al., 2010a; Mokkink et al., 2010b; Mokkink et al., 2020) establece que la validez de contenido debe ser evaluada mediante un juicio experto sobre la relevancia y la amplitud de los elementos constitutivos del instrumento.

La relevancia de los ítems debe ser juzgada en función de su pertinencia al constructo que se busca medir, a la población objetivo del estudio y al propósito global del instrumento (Mokkink et al., 2010c; Terwee et al., 2018). En la validación de un instrumento, la especificidad y el detalle del contenido deben alinearse con la población diana. Es imperativo reevaluar la relevancia de cada ítem si el instrumento se aplica a una población diferente de la original (Mokkink et al., 2010c; Terwee et al., 2018). Dada la centralidad del instrumento en esta evaluación, los propios usuarios pueden ser considerados como expertos para juzgar la relevancia de los ítems. Un número elevado de observaciones ausentes para un ítem particular puede indicar una falta de relevancia o una formulación ambigua del mismo para la población en cuestión.

Para evaluar la amplitud de los ítems, se consideran tres dimensiones fundamentales: la cobertura de contenido de los ítems, la descripción de los dominios y la fundamentación teórica. Las dos primeras dimensiones abordan si todos los aspectos relevantes del constructo están adecuadamente representados por los ítems y sus respectivos dominios. La tercera, la fundamentación teórica, se refiere a la disponibilidad de una descripción clara del constructo y la teoría que lo sustenta (Mokkink et al., 2010c; Terwee et al., 2018).

El objetivo último en la evaluación de la validez de contenido de un instrumento es juzgar si este satisface los criterios preestablecidos para una óptima validez de contenido. El estudio Delphi define diez criterios relacionados con tres aspectos esenciales de esta propiedad psicométrica (Terwee et al., 2018):

- Relevancia: Todos los elementos del instrumento deben ser relevantes para el constructo de interés en una población y contexto de uso específicos.
- Exhaustividad: El instrumento no debe omitir ningún aspecto crucial del constructo.
- Comprensibilidad: Los usuarios deben interpretar los ítems de manera consistente con su intención original.

Estos diez criterios se articulan como preguntas guía durante la evaluación de la validez de contenido. Las respuestas a estas preguntas deben ser sintetizadas en una calificación global de relevancia, exhaustividad, comprensibilidad y validez de contenido general para cada puntuación o subpuntuación del instrumento, tomando como referencia la información sobre su desarrollo, la evidencia derivada de estudios previos de validez de contenido y el contenido del instrumento actualizado (Terwee et al., 2018).

Segunda etapa

En una segunda fase, se procede a la formación de un panel de expertos. Este panel se compone de un grupo de especialistas independientes, reconocidos y acreditados en al menos uno de los campos objeto de evaluación en el programa. El propósito de este panel es alcanzar un juicio consensuado y colectivo sobre dicho programa. El juicio emitido puede referirse a la implementación práctica o a los efectos, ya sean parciales o totales, del programa, según lo requiera la evaluación. La formación de un panel de expertos se orienta primordialmente a la evaluación de un conjunto de procedimientos estándar, siguiendo una metodología de trabajo predefinida para la reunión y la obtención del juicio (Gagnier et al., 2021; Mokkink et al., 2010a; Mokkink et al., 2010b; Mokkink et al., 2020; Terwee et al., 2018).

Los criterios de selección para los miembros del panel son:

- Experiencia profesional: Indispensable en el área temática a tratar.
- Independencia: Respecto al programa bajo valoración.
- Capacidad de trabajo en equipo: Esencial para la deliberación consensuada.

La selección de los expertos puede realizarse mediante un método rápido (listado de especialistas reconocidos en un campo específico, con verificación de su independencia) o mediante una selección por etapas (que implica la definición previa del perfil de experto requerido en función del tema de evaluación) (Gagnier et al., 2021; Mokkink et al., 2010a; Mokkink et al., 2010b; Mokkink et al., 2020; Terwee et al., 2018).

Tercera etapa

Finalmente, la última etapa comprende la realización de una prueba piloto. Esta fase se considera intrínseca a la evaluación de la validez de contenido, específicamente en lo que respecta a la comprensibilidad. Se recomienda la participación de entre 5 y 8 individuos con características demográficas y clínicas similares a las de la población objetivo, monitorizando sus respuestas y el tiempo de ejecución (Terwee et al., 2018; Wild et al., 2005). La finalidad primordial de esta prueba es identificar particularidades de la nueva versión del instrumento, evaluar el contexto de aplicación y detectar posibles deficiencias o factores que puedan obstaculizar su implementación eficaz (Terwee et al., 2018).

4.2.3 Validez estructural

La validez estructural constituye una propiedad psicométrica fundamental que certifica la idoneidad de las puntuaciones derivadas de un instrumento para representar una medición fiel del fenómeno de estudio (García de Yébenes et al., 2009; Mokkink et al., 2010a; Mokkink et al., 2018b; Mokkink et al., 2019). Esta propiedad evalúa la congruencia entre el instrumento y la teoría subyacente del constructo o fenómeno que se pretende cuantificar, asegurando que las mediciones obtenidas concuerden con las de otros instrumentos diseñados para evaluar condiciones similares (Batista-Foget et al., 2004; García de Yébenes et al., 2009; Gil et al., 2000; Mokkink et al., 2010a; Mokkink et al., 2018b; Mokkink et al., 2019).

Para la evaluación de la validez estructural, se procede inicialmente con un EFA. El propósito primordial del EFA es desvelar la estructura interna subyacente a un conjunto relativamente amplio de variables. Mediante este análisis, se evalúa la adecuación del modelo factorial propuesto, así como la idoneidad de los ítems y de la muestra investigada. Para tal fin, se emplean dos pruebas estadísticas: el Test de Esfericidad de Bartlett, cuya significancia estadística ($p < 0.05$) es un requisito indispensable; y el Test de Kaiser-Meyer-Olkin (KMO), para el cual se consideran valores aceptables aquellos superiores a 0.5 (Kaiser, 1960).

Posteriormente, para la extracción de factores y la determinación de la validez estructural del instrumento, se utilizan como criterios de referencia un autovalor superior a 1, una varianza explicada superior al 10%, y el análisis del scree plot (gráfico de sedimentación de factores rotados) (Kaiser, 1960; Cattell, 1966). Durante este análisis, se calculan las cargas factoriales de las variables, empleando la estimación de máxima verosimilitud (ML) para evaluar el ajuste del modelo. Este método, implementado a través del software estadístico SPSS, permite asignar un determinado ítem a uno o más factores sin descartar ninguna posibilidad.

Únicamente los ítems con cargas factoriales en el EFA iguales o superiores a 0.30, umbral de corte aceptable (Tabachnick et al., 2013), son retenidos para análisis subsiguientes. Aquellos ítems con valores inferiores a 0.30 se eliminan

por carecer de significancia en el contexto del nuevo cuestionario que se pretende validar, al igual que los elementos que presenten cargas cruzadas significativas.

Finalmente, se lleva a cabo un CFA utilizando el programa de ecuaciones estructurales SPSS AMOS. El CFA tiene como objetivo determinar si el número de factores obtenidos y sus respectivas cargas factoriales concuerdan con las expectativas derivadas de una teoría previa sobre los datos. Los índices de ajuste del modelo empleados incluyen el Chi-cuadrado (χ^2), el Error Cuadrático Medio de Aproximación (RMSEA), que representa el ajuste anticipado con la población total en lugar de solo la muestra, y el Índice de Ajuste Comparativo (CFI), que evalúa la bondad de ajuste del modelo. Para el RMSEA, valores inferiores a 0.06 indican un ajuste aceptable (Fan et al., 2007; Schumacker et al., 2004); mientras que, para el CFI, valores iguales o superiores a 0.80 se consideran satisfactorios, y valores iguales o superiores a 0.95 reflejan un ajuste excelente (Bowne et al., 1992; Hu et al., 1999).

4.2.4 Consistencia interna

La consistencia interna se define como el grado de interrelación entre los ítems de un instrumento de medida o entre aquellos ítems que componen una dimensión específica en escalas multidimensionales (Mokkink et al., 2018b; Mokkink et al., 2019). Su evaluación determina la homogeneidad entre los ítems diseñados para cuantificar una misma dimensión, asegurando que cada componente del dominio evalúa el constructo deseado y no un concepto distinto (Campo-Arias et al., 2008; Carretero-Dios et al., 2005; Oviedo et al., 2005). Es crucial considerar que las escalas multidimensionales están conceptualizadas para medir de forma independiente los diversos dominios que conforman un constructo particular, por lo que el cálculo de la consistencia interna debe realizarse para cada dimensión individualmente.

Los instrumentos que presentan una consistencia interna elevada, indicando que sus ítems miden un constructo único y homogéneo, garantizan una relación lineal entre la suma de las puntuaciones de sus ítems y el constructo medido (Argimon et al., 2004; Campo-Arias et al., 2008). Estadísticamente, la consistencia interna puede estimarse mediante el rango de los coeficientes de correlación de Pearson, el método de división por mitades de Spearman, la fórmula de Kuder-Richardson (KR-20) para escalas dicotómicas, y/o el coeficiente alfa de Cronbach para escalas policotómicas (Tangarife et al., 2015).

El coeficiente alfa de Cronbach es el método estadístico más empleado para evaluar la consistencia interna de un instrumento (Cronbach, 1951; da Silva et al., 2015; Soler Cárdenas, 2008). Su valor depende del número de ítems que constituyen la escala y a la correlación media entre ellos. Sus valores oscilan entre 0 y 1 (Bland y Altman, 1997; Henson, 2001; Soler-Cárdenas et al., 2012).

Un valor cercano a 1 indica una correlación casi perfecta entre los ítems, mientras que un valor de 0 denota ausencia de correlación, y un coeficiente de -1 sugiere una correlación negativa entre los ítems. La detección de coeficientes negativos, aunque infrecuente, es plausible al calcular este estadístico en escalas compuestas por solo dos ítems (Bland y Altman, 1997; Campo-Arias, 2006; Krus et al., 1993). Adicionalmente, el coeficiente alfa de Cronbach debe calcularse para los ítems de cada dominio (subescala) cuando el instrumento está estructurado por un grupo de dominios, en relación con el valor de las puntuaciones de dicho dominio (correlación ítem-dominio) (García de Yébenes et al., 2009). El valor mínimo aceptado para el alfa de Cronbach es de 0.70; valores inferiores implican una baja fiabilidad de la escala utilizada. Por otro lado, el valor máximo deseado es de 0.90, ya que valores superiores sugieren duplicidad o la evaluación del mismo elemento de un dominio o constructo por diferentes ítems. En consecuencia, los valores óptimos para el alfa de Cronbach se sitúan entre 0.80 y 0.90 (Streiner, 2003).

4.2.5 Fiabilidad

La estabilidad y consistencia de las puntuaciones obtenidas mediante un instrumento de evaluación a lo largo de mediciones continuadas constituyen un indicador crucial de la fiabilidad de dicho instrumento (Mokkink et al., 2010a; Mokkink et al., 2018b; Mokkink et al., 2019).

Desde una perspectiva estadística, el ICC se erige como el índice más apropiado y frecuentemente empleado para cuantificar la concordancia entre mediciones repetidas de una variable. Este coeficiente es particularmente relevante cuando las mediciones son adquiridas por diferentes evaluadores o mediante distintos instrumentos. Los valores del ICC varían en un rango de 0 a 1, donde un valor de 0 denota una ausencia total de concordancia, mientras que un valor de 1 indica una fiabilidad o concordancia absoluta en los resultados obtenidos.

La interpretación de los valores del ICC se rige por las siguientes categorías (Henson, 2001; Streiner et al., 2015):

- Baja fiabilidad: Valores de ICC inferiores a 0.4 ($ICC < 0.4$).
- Fiabilidad de regular a aceptable: Valores de ICC entre 0.4 y 0.69 ($0.4 \leq ICC \leq 0.69$).
- Fiabilidad excelente: Valores de ICC iguales o superiores a 0.70 ($ICC \geq 0.70$).

4.2.5.1 Fiabilidad intra-evaluador o fiabilidad test-retest

La repetibilidad de un instrumento se define por la capacidad de obtener puntuaciones equivalentes al aplicar el mismo instrumento, por los mismos evaluadores, empleando idénticos métodos y con la misma población, en dos

periodos de tiempo distintos (García de Yébenes et al., 2009; Mokkink et al., 2010a). Esta propiedad es fundamental para establecer la fiabilidad test-retest del instrumento.

Para la evaluación de esta propiedad, la selección del coeficiente estadístico apropiado depende de la naturaleza paramétrica de la muestra. Se pueden emplear el ICC, el coeficiente de correlación de Pearson o el coeficiente de correlación de Spearman (Carvajal et al., 2011). Es importante destacar que, al igual que el ICC, el coeficiente de Pearson se utiliza para cuantificar la correlación entre variables cuantitativas. Sin embargo, en el caso de variables cualitativas de naturaleza ordinal, el coeficiente de correlación de Spearman es más adecuado (Muller et al., 1994; Prieto et al., 1998).

Es crucial considerar que existirá una variación inherente en el constructo medido a lo largo del tiempo transcurrido entre la primera (test) y la segunda (retest) aplicación del instrumento. Por consiguiente, se recomienda que el intervalo entre ambas mediciones no sea excesivamente prolongado, a fin de minimizar posibles alteraciones en el fenómeno de interés. No obstante, tampoco debe ser excesivamente corto, ya que esto podría inducir un efecto de aprendizaje, donde los participantes recuerden las respuestas previamente proporcionadas, comprometiendo así la validez de la segunda medición. En ambos escenarios, el valor de la repetibilidad del instrumento se vería significativamente alterado. Adicionalmente, la negativa de algunos participantes a someterse a una segunda aplicación del instrumento representa otra posible limitación metodológica (Muller et al., 1994; Prieto et al., 1998; Ramada-Rodilla et al., 2013).

4.2.5.2 Fiabilidad inter-evaluador

La fiabilidad inter-observador alude al grado de consenso entre múltiples evaluadores al valorar a los mismos individuos, utilizando el mismo instrumento y en un mismo periodo de tiempo. Es fundamental reconocer que esta propiedad no es aplicable ni medible en el contexto de instrumentos autoadministrados, dado que la intervención de evaluadores externos no influye en los resultados obtenidos.

Sin embargo, existen limitaciones inherentes a la medición de esta consistencia que deben ser consideradas (García de Yébenes et al., 2009; Ramada-Rodilla et al., 2013):

- La posibilidad de acuerdo por azar entre los evaluadores.
- La presencia de un error sistemático introducido por alguno de los evaluadores.

Cuando la evaluación de la consistencia inter-evaluador es pertinente, los métodos estadísticos empleados para su cuantificación son idénticos a los utilizados para determinar la fiabilidad test-retest.

4.2.6 Error de medición

El error de medición se define como la discrepancia entre la puntuación observada y la puntuación teórica verdadera de un constructo (Mokkink et al., 2010a). Dentro de este marco, el Error Estándar de Medición (SEM) emerge como un método estadístico fundamental para cuantificar la magnitud de este error en un instrumento. El SEM proporciona una estimación de la variabilidad o dispersión inherente a los errores de medición de una prueba, integrándose en una amplia gama de técnicas estadísticas que incluyen el análisis de la varianza (ANOVA), el análisis factorial y el análisis de regresión múltiple. Asimismo, el SEM facilita la resolución de interrogantes que demandan métodos de análisis de regresión múltiple; por ejemplo, tanto el SEM como el ANOVA permiten realizar comparaciones inter-grupo e intra-grupo, convergiendo en resultados equivalentes (Mokkink et al., 2018b; Mokkink et al., 2019; Ortiz et al., 2018).

Parámetros estimados del SEM

En el contexto del SEM, la especificación y estimación de los parámetros son cruciales tanto para los modelos de medición como para los modelos estructurales. Una vez especificados, estos parámetros pueden clasificarse en tres categorías:

- **Parámetro libre:** Su valor es desconocido y se calcula a partir del análisis de los datos.
- **Parámetro fijo:** Su valor es preestablecido por el investigador, generalmente fijado en un valor determinado (comúnmente 10).
- **Parámetro restringido:** Aunque su valor es desconocido, el investigador le asigna una restricción que lo vincula a otros parámetros o a un valor específico (Jackson et al., 2009; Schmitt, 2011).

Etapas del SEM

El proceso de Análisis de Ecuaciones Estructurales se articula en cuatro etapas secuenciales: especificación, estimación, evaluación del ajuste global y re-especificación.

Esta etapa inicial implica la formulación de hipótesis concernientes a las relaciones entre las variables. Es importante discernir si estas relaciones son unidireccionales o bidireccionales, basándose en la evidencia empírica previa, la teoría sustantiva, o una combinación de ambas (Ortiz et al., 2018). Dichas relaciones pueden ser directas o indirectas, especialmente cuando el modelo

incorpora variables mediadoras que interceden en el efecto de una variable sobre otra (Bou et al., 2010; Byrne, 2008). Es fundamental destacar que el SEM se concibe primariamente como una técnica de naturaleza confirmatoria, más que exploratoria.

Una vez que el modelo ha sido debidamente especificado e identificado (un modelo se considera identificado cuando el número de parámetros a estimar no excede el número de datos observados, siendo únicamente los modelos identificados susceptibles de estimación), sus parámetros son evaluados con el objetivo de minimizar las discrepancias entre la matriz de covarianza observada y la matriz de covarianza poblacional teórica (Hayes, 2018).

Aunque existen diversos procedimientos de estimación adecuados, los métodos de Máxima Verosimilitud (ML), los métodos de Mínimos Cuadrados (SL) (Saris et al., 2009) y el método de Distribución Asintótica Libre (ADF) (Hu et al., 1992) son los más prevalentes.

De manera sistemática, el ajuste del modelo se evalúa mediante la prueba de chi-cuadrado (χ^2), siendo esta notablemente sensible a tamaños muestrales grandes. Por consiguiente, la evaluación global del modelo se complementa habitualmente con la consideración de otros indicadores de bondad de ajuste, los cuales oscilan entre 0 y 1, donde valores más altos denotan una mayor proporción de varianza explicada por el modelo. Entre estos, el Índice de Ajuste Comparativo (CFI) y el Índice de Tucker-Lewis (TLI) son ampliamente utilizados, contrastando el modelo propuesto con un modelo nulo. Valores de CFI>0.95 y TLI>0.90 son indicativos de un buen ajuste (Hu et al., 1999).

Un ajuste adecuado de los datos al modelo también se infiere de la proximidad a 0 de los valores residuales, lo que refleja el porcentaje de varianza no explicada por el modelo. En este sentido, se suelen reportar el Error Cuadrático Medio de Aproximación (RMSEA) y la Raíz Cuadrada Media Residual Estandarizada (SRMR). Se consideran aceptables valores de RMSEA<0.06 y SRMR<0.08 (Fan et al., 2007).

Tras la evaluación del ajuste global del modelo a los datos, se presentan dos opciones: la primera, y siempre preferible, consiste en la interpretación y comunicación de los resultados obtenidos. No obstante, si el ajuste global del modelo no cumple las expectativas, se hace necesaria una re-especificación del modelo. Las justificaciones para esta re-especificación pueden incluir la identificación de patrones de correlación inicialmente excluidos, el rastreo de residuos de gran magnitud, o la manipulación de algún algoritmo estadístico (Ortiz et al., 2018).

4.2.7 Validez de criterio convergente

La validez de constructo, un pilar fundamental en la psicometría, se establece mediante el análisis de las correlaciones entre las puntuaciones derivadas de diferentes escalas (Mokkink et al., 2018b; Mokkink et al., 2019). Este proceso permite determinar si un instrumento mide efectivamente el constructo teórico que pretende evaluar.

Cuando las puntuaciones obtenidas con un instrumento exhiben correlaciones significativas con las de otros instrumentos que miden el mismo constructo, se habla de validez convergente. Esta convergencia sugiere que las escalas son conceptualmente análogas o congruentes, confirmando que evalúan el mismo fenómeno subyacente (Carvajal et al., 2011; Mokkink et al., 2018b; Mokkink et al., 2019).

Por el contrario, la validez divergente se manifiesta cuando las correlaciones entre las puntuaciones de escalas que miden constructos diferentes resultan bajas o negativas. Una asociación no significativa entre estas variables confirma que los instrumentos evalúan dominios conceptuales distintos. Si, por el contrario, se obtienen correlaciones elevadas en este escenario, esto podría indicar que la escala en proceso de validación carece de la especificidad suficiente para medir el constructo de interés en la población seleccionada (Alarcon et al., 2008; Argibay, 2006; Carretero-Dios et al., 2005; Carvajal et al., 2011).

Para la cuantificación de esta propiedad, comúnmente se emplea el coeficiente de correlación de Pearson. En condiciones óptimas, se esperan resultados psicométricos que demuestren una correlación superior a 0.60 para la validez convergente y una correlación inferior a 0.20 para la validez divergente (DeVon et al., 2007).

4.2.8 Validez de criterio

La validez de criterio establece el grado en que las puntuaciones obtenidas de una escala son válidas al ser comparadas con un estándar o un patrón de referencia previamente establecido (García de Yébenes et al., 2009; Mokkink et al., 2010a; Mokkink et al., 2018b; Mokkink et al., 2019). Para determinar esta validez, el nuevo instrumento bajo evaluación debe ser comparado con una escala existente y reconocida, a menudo denominada "gold standard", la cual ha demostrado ser el instrumento de referencia más adecuado para medir el fenómeno de interés. La comparación subsiguiente de las puntuaciones obtenidas con ambas escalas permite determinar la existencia de una correlación adecuada (García de Yébenes et al., 2009; Lamprea et al., 2007).

Siempre que se disponga de un estándar o una escala alternativa que cumpla la misma función (es decir, que demuestre ser fiable, independiente, válida y capaz de medir la misma condición de interés), es recomendable seguir los siguientes pasos para evaluar esta propiedad (García de Yébenes et al., 2009; Ramada-Rodilla et al., 2013):

1. Seleccionar el estándar o su equivalente más apropiado.
2. Elegir una muestra representativa de la población objeto de estudio.
3. Administrar la escala en evaluación y obtener una puntuación para cada individuo.
4. Evaluar a cada sujeto con el "gold standard" y, posteriormente, comparar las puntuaciones obtenidas con ambos instrumentos.

Las características de esta propiedad pueden evaluarse, y según el momento en que se realice la comparación de resultados, la validez se clasificará como concurrente o predictiva (Lamprea et al., 2007; Streiner et al., 2015).

La validez concurrente determina el grado de correlación existente entre los resultados obtenidos por la escala considerada "criterio" o "estándar" y la escala que está siendo evaluada, cuando ambas son administradas simultáneamente (Lamprea et al., 2007). Esta comparación se realiza estadísticamente utilizando el coeficiente de correlación de Pearson (García de Yébenes et al., 2009; Sánchez-Pedraza et al., 1998) o el coeficiente de Spearman, en función de las características y la distribución de los datos (Streiner et al., 2015).

Por otro lado, la validez predictiva evalúa la capacidad de la nueva escala de medición para predecir las puntuaciones obtenidas por el "gold standard" cuando este último no se aplica en el mismo momento, sino en un punto futuro (Lamprea et al., 2007). Desde una perspectiva estadística, esta comparación se efectúa de manera análoga a la validez concurrente.

El objetivo primordial de la validez de criterio es establecer una correlación adecuada entre ambos instrumentos. Es fundamental considerar que si la finalidad de validar una nueva escala radica en su mayor utilidad (por ejemplo, sencillez en la aplicación, eficiencia en el proceso de calificación, economía, comodidad para el individuo, pertinencia y reducción del error de medida), la validación de criterio requerirá la obtención de correlaciones ≥ 0.8 . Esto indicaría que las dos escalas son psicométricamente similares. Sin embargo, si el propósito es demostrar que la nueva versión es superior y más válida que el instrumento de referencia, lo apropiado sería buscar correlaciones en un rango entre 0.3 y 0.7, lo que revelaría que, si bien ambos instrumentos miden el mismo atributo, presentan diferencias distintivas (Sánchez-Pedraza et al., 1998; Streiner, 1993; Streiner et al., 2015).

4.2.9 Prueba de hipótesis

Cuando un "gold standard" no está disponible para la evaluación de un instrumento, la validez de constructo de una nueva herramienta se investiga comúnmente mediante la prueba de hipótesis sobre relaciones esperadas con otras medidas de resultados de buena calidad y/o diferencias anticipadas entre grupos relevantes. Es de suma importancia que estas hipótesis sean definidas a priori al evaluar la validez de constructo de la nueva herramienta, con el fin de permitir a los autores extraer conclusiones imparciales después de la recopilación y el análisis de los datos (Mokkink et al., 2018b).

La prueba de hipótesis constituye un proceso iterativo y continuo (Strauss et al., 2009). Las hipótesis específicas deben incluir una indicación de la dirección esperada y la magnitud de las correlaciones o diferencias. De este modo, la prueba de hipótesis determinará si la dirección y la magnitud de una correlación o diferencia se asemejan a lo que cabría esperar en función de los constructos que se están midiendo. Cuantas más hipótesis se prueben sobre la correspondencia de los datos con las hipótesis formuladas a priori, mayor será la evidencia acumulada en favor de la validez de constructo (Mokkink et al., 2018b).

4.2.10 Sensibilidad al cambio

La sensibilidad al cambio se define como la capacidad inherente de un instrumento psicométrico para detectar variaciones temporales en el constructo que se pretende medir, ya sea a nivel individual o colectivo (Mokkink et al., 2018b; Mokkink et al., 2019). Esta propiedad es de particular relevancia en el diseño y aplicación de escalas diagnósticas, estudios prospectivos y ensayos clínicos. En estos contextos, la sensibilidad al cambio, en conjunto con la especificidad, permite evaluar la respuesta a una intervención o tratamiento determinado.

Cabe destacar que la evaluación de la sensibilidad al cambio no suele incluirse en estudios que abordan variables como la satisfacción, el bienestar, las actitudes o las percepciones, dadas las características inherentes de estabilidad o la naturaleza subjetiva de estos constructos (Argimon et al., 2004; García de Yébenes et al., 2009; Guyatt et al., 1987; Mokkink et al., 2010a; Sánchez et al., 2004).

4.2.11 Validez transcultural

La equivalencia lingüística y cultural de un instrumento traducido o adaptado se refiere al grado en que la interpretación de sus ítems refleja adecuadamente el desempeño de los ítems de la versión original (Mokkink et al., 2018b; Mokkink et al., 2019).

El proceso de validación transcultural es indispensable cuando un instrumento se adapta a una población específica, incluso si ya ha sido traducido a su idioma de destino. Este proceso requiere una traducción rigurosa que, posteriormente, debe ser revisada por un comité de expertos (Maneesriwongul et al., 2004; Mokkink et al., 2018b; Mokkink et al., 2019), seguida de una prueba piloto (pre-test) y, finalmente, un proceso formal de validación (Acquadro et al., 2008; Wild et al., 2005).

La traducción debe ser realizada por traductores expertos, idealmente dos traductores nativos del idioma de destino, uno de los cuales debe poseer experiencia y especialidad en el campo temático del instrumento. Ambos traductores deben trabajar de forma independiente y emplear un enfoque de traducción secuencial hacia adelante (forward-translation) y hacia atrás (back-translation). Para la traducción inversa, es crucial que los traductores sean nativos, trabajen de manera independiente y, preferiblemente, no tengan conocimiento previo del contenido específico del instrumento. Al concluir este proceso, se debe elaborar un informe detallado con los principales comentarios y observaciones sobre el proceso de traducción (Mokkink et al., 2018b; Mokkink et al., 2019).

Posteriormente, se convoca un panel de expertos para evaluar la validez de contenido del instrumento. Esta evaluación se realiza a través de indicadores específicos diseñados para asegurar que el nuevo instrumento es un reflejo semántico preciso de la versión original y que abarca la totalidad de su contenido (Maneesriwongul et al., 2004; Wild et al., 2005). Una vez obtenida la versión modificada del instrumento, el siguiente paso es someter esta herramienta a una prueba piloto (pre-test) con una muestra de personas que sea representativa de la población objetivo. El objetivo de esta fase es determinar la calidad metodológica del nuevo instrumento en términos de relevancia, exhaustividad y comprensibilidad en relación con el constructo de interés y la población de estudio. Esto también permite confirmar si el contenido resultante del instrumento refleja adecuadamente el constructo que se pretende medir (Mokkink et al., 2018b; Mokkink et al., 2019).

Una vez obtenida la versión final de la escala, traducida y adaptada culturalmente, se somete a un exhaustivo proceso de validación mediante la medición de una serie de propiedades psicométricas fundamentales (Mokkink et al., 2010a; Mokkink et al., 2010b; Mokkink et al., 2010c; Mokkink et al., 2018a; Mokkink et al., 2018b; Mokkink et al., 2019; Prinsen et al., 2018).

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

5. OBJETIVOS

Objetivo General

Analizar las variables psicométricas y validar escalas de evaluación funcional y de desempeño ocupacional en personas que han sufrido un ictus.

Objetivos específicos

Identificar la escala más recomendada entre las versiones del Índice de Barthel y de la escala FIM para evaluar la funcionalidad e independencia en las actividades de la vida diaria de personas que han sufrido un ictus.

Establecer la versión más recomendada de la escala Fugl-Meyer Assessment para la evaluación de la función sensoriomotora de miembro superior en población que ha sufrido un ictus.

Actualizar, traducir, adaptar transculturalmente y validar la Escala Nottingham de Vestido en Ictus – Revisada (ENVI-R) para la evaluación del vestido en población española que ha sufrido un ictus.

Validar la versión española de la escala Fugl-Meyer Assessment para poder evaluar la función motora, sensitiva, equilibrio, rango articular y dolor mediante esta escala en personas españolas post-ictus.

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

6. MARCO EMPÍRICO

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

6.1 REVISIÓN PSICOMÉTRICA DE CUESTIONARIOS DE EVALUACIÓN

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

6.1.1 ÍNDICE DE BARTHEL Y MEDIDA DE INDEPENDENCIA FUNCIONAL EN PERSONAS CON ICTUS: UNA REVISIÓN SISTEMÁTICA PSICOMÉTRICA

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

ÍNDICE DE BARTHEL Y MEDIDA DE INDEPENDENCIA FUNCIONAL EN PERSONAS CON ICTUS: UNA REVISIÓN SISTEMÁTICA PSICOMÉTRICA

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

El Artículo I se encuentra bajo revisión y pendiente para su publicación en la revista "Occupational Therapy International Journal".

Abstract

Aims: To identify, critically appraise, and summarize the different Barthel Index and Functional Independence Measure versions to synthesize the results and recognize which version is the most recommended for measuring independence and functionality in activities of daily living in people who have suffered a stroke.

Design: Psychometric systematic review.

Data Source: An exhaustive search was conducted during January and February 2024 in the following databases: Pubmed, Web of Science (WoS), Open Grey, Scielo, Cochrane, Dialnet, LILACS and OTSeeker.

Review Methods: By the most up-to-date COnsensus-based Standards for the selection of health status Measurement INstruments methodology (COSMIN) and Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statements for a systematic review of the measurement properties of existing patient-reported outcome measures to select the most appropriate outcome measurement instrument. The protocol of this systematic review was registered in Prospero.

Results: 20 versions of the Barthel Index over 31 articles and 15 versions of the Functional Independence Measure over 18 were identified. The Original Barthel Index was the questionnaire version included in the most significant volume of studies. The original version of the Barthel Index was the questionnaire version with the most evaluated psychometric characteristics (7 properties). Six versions (BI-5, BI-SS, MBI, Ira-MBI, FIM, Per-FIM) were evaluated with six psychometric characteristics. The original questionnaire version had a better score regarding the quality of evidence.

Conclusion: According to the quality methodological results, the Original Barthel Index is the recommended version to assess independence and functionality in activities of daily living in stroke patients.

Keywords: Barthel Index, Functional Independence Measure, Stroke, Independence, Activities of Daily Living

Resumen

Objetivos: Identificar, valorar críticamente y resumir las diferentes versiones del Índice de Barthel y de la Medida de Independencia Funcional para sintetizar los resultados y reconocer cuál versión es la más recomendada para medir la independencia y funcionalidad en las actividades de la vida diaria en personas que han sufrido un ictus.

Diseño: Revisión sistemática psicométrica.

Fuente de datos: Se realizó una búsqueda exhaustiva durante enero y febrero de 2024 en las siguientes bases de datos: Pubmed, Web of Science (WoS), Open Grey, Scielo, Cochrane, Dialnet, LILACS y OTSeeker.

Métodos de revisión: Según la metodología más actualizada de los Estándares de Consenso para la Selección de Instrumentos de Medición del Estado de Salud (COSMIN) y las declaraciones de los Elementos de Informe Preferidos para Revisiones Sistemáticas y Metaanálisis (PRISMA), se realizó una revisión sistemática de las propiedades de medición de las medidas de resultados existentes, reportadas por los pacientes, para seleccionar el instrumento de medición de resultados más adecuado. El protocolo de esta revisión sistemática se registró en Prospero.

Resultados: Se identificaron 20 versiones del Índice de Barthel en 31 artículos y 15 versiones de la Medida de Independencia Funcional en 18 artículos. El Índice de Barthel Original fue la versión del cuestionario incluida en el mayor volumen de estudios. Esta versión original del Índice de Barthel fue la versión del cuestionario con las características psicométricas más evaluadas (7 propiedades). Se evaluaron seis versiones (BI-5, BI-SS, MBI, Ira-MBI, FIM y Per-FIM) con seis características psicométricas. La versión original del cuestionario obtuvo una mejor puntuación en cuanto a la calidad de la evidencia.

Conclusión: De acuerdo con los resultados metodológicos de calidad, el Índice de Barthel Original es la versión recomendada para evaluar la independencia y funcionalidad en las actividades de la vida diaria en personas con ictus.

Palabras clave: Índice de Barthel, Medida de Independencia Funcional, Ictus, Independencia, Actividades de la Vida Diaria

Contribución del artículo a la Terapia Ocupacional:

- Encuentra las diferentes versiones y traducciones del Índice de Barthel y de las escalas de Medida de Independencia Funcional; esto permite a los terapeutas ocupacionales tener todas las versiones recopiladas y poder elegir la que necesitan y compararlas.
- Identificar las propiedades psicométricas analizadas en los diversos estudios de las diferentes versiones.
- Conocer la mejor versión para evaluar la independencia y funcionalidad en las actividades de la vida diaria en personas con ictus, un constructo comúnmente evaluado por los terapeutas ocupacionales.
- Reconocer el Índice de Barthel y la Medida de Independencia Funcional como la escala estándar de oro en la evaluación de la independencia y funcionalidad en las actividades de la vida diaria.

1.INTRODUCCIÓN

El ictus es la segunda causa de muerte (Feigin et al., 2023) y la primera causa de discapacidad a nivel mundial (Hua et al., 2023). El riesgo de mortalidad depende del tipo de ictus; por ejemplo, los accidentes isquémicos transitorios, en los que los síntomas se resuelven en menos de 24 horas, tienen el mejor pronóstico (Daffertshofer et al., 2004). Quince millones de personas sufren un ictus cada año en todo el mundo; de ellas, 5 millones fallecen y otros 5 millones quedan con discapacidad permanente, lo que supone una carga para la familia y la comunidad (Katan et al., 2018).

La carga de discapacidad tras un ictus es significativa y aumenta rápidamente, y su incidencia está aumentando a nivel mundial en personas jóvenes y de mediana edad menores de 55 años (Feigin et al., 2023). Se estima que la cantidad de ictus aumentará un 50% en los próximos años, pasando de 6,6 millones de casos en 2020 a 9,7 millones en 2050 (Feigin et al., 2023).

Las disfunciones del ictus, que se manifiestan principalmente como deterioro motor y cognitivo, con frecuencia afectan el desempeño de las actividades de la vida diaria, las tareas ocupacionales y, en última instancia, la independencia y la funcionalidad de quienes sobreviven al ictus (Soleimani et al., 2024). Tras el ictus, se produce una alteración física de las extremidades y deterioro cognitivo (Ingram et al., 2023; Li et al., 2023; Zhao et al., 2023) que provoca la incapacidad de usar las extremidades superiores para alcanzar, agarrar y manipular objetos, afectando a la realización de actividades de la vida diaria como vestirse, alimentarse o bañarse, lo que provoca la pérdida de autonomía en las actividades de la vida diaria (AVD) y la dependencia de otros para realizarlas (Ingram et al., 2023; Dawson et al., 2023). Estos déficits perjudican la calidad de vida de las personas que han sufrido un ictus (Dawson et al., 2023), además de tener un enorme impacto social y económico, con un alto coste humano y material, tanto intrahospitalario como extrahospitalario (Dobkin, 1995; Evers et al., 2004; Roschmah et al., 2021). Recuperar la independencia tras un ictus depende en gran medida de la recuperación motora, en particular de la recuperación de la actividad de las extremidades superiores (Shanks et al., 2024).

Con estos datos y estas tendencias, se espera que la demanda de servicios de rehabilitación del ictus aumente (Kwakkel et al., 2023), por lo que es necesario disponer de herramientas que permitan a los profesionales de la rehabilitación evaluar los diferentes déficits que provoca un ictus.

Dado que la independencia de la persona tras un ictus es un requisito, una evaluación y un trabajo principalmente en rehabilitación, existe una amplia variedad de herramientas para evaluar el constructo de independencia y funcionalidad en personas con ictus. Algunas de ellas son el Índice de Barthel (BI), la Medida de Independencia Funcional (MIF/FIM), la Escala de Rankin modificada (mRS), la National Institutes of Health Stroke Scale (NIHSS), el Freanchay Activity Index (FAI), la Nottingham Extended Activity of Daily Living

(NEADL), etc. (Colomer et al., 2023; Jhong et al., 2013; Langhammer et al., 2017; Miki et al., 2016; Sarker et al., 2012).

Elegir el instrumento más apropiado para evaluar la independencia y la funcionalidad en personas con ictus es un desafío, y la elección debe depender de las propiedades de medición y la viabilidad del instrumento (Prinsen et al., 2018).

El Índice de Barthel y la Medida de Independencia Funcional son conocidos internacionalmente, con diferentes versiones que varían en el número de ítems, el modo de evaluación, la escala de puntuación, etc. Ambas han sido adaptadas y traducidas a diferentes idiomas y culturas. Numerosos estudios evalúan sus propiedades psicométricas.

Por lo tanto, es necesaria una revisión que las combine todas y permita analizarlas y conocer las versiones más válidas y recomendadas. Ninguna revisión sistemática compara ni evalúa las propiedades psicométricas del Índice de Barthel ni de la Medida de Independencia Funcional. Sin embargo, son dos escalas ampliamente utilizadas para evaluar la independencia y la funcionalidad de los usuarios, tanto en la práctica clínica como en la investigación.

Ambas se consideran el estándar de oro para evaluar la independencia y la funcionalidad, pero aún no se ha demostrado. Existen varias versiones de ambas escalas en la literatura, pero es necesario un consenso claro sobre cuál es la más adecuada y recomendada. Esta revisión sistemática tiene como objetivo realizar una revisión psicométrica de ambos cuestionarios, utilizando la metodología COSMIN para conocer las diferentes versiones disponibles en la literatura y evaluar cuál sería la más recomendable para personas con ictus.

2.MÉTODOS

2.1 Objetivos

Esta revisión sistemática tuvo como objetivo identificar, evaluar críticamente y resumir las diferentes versiones del Índice de Barthel y la Medida de Independencia Funcional, para reconocer cuál es la versión más recomendada para medir la independencia y funcionalidad en las actividades de la vida diaria en personas con ictus.

2.2 Diseño

Se realizó una revisión sistemática de las propiedades de medición de las diferentes versiones del Índice de Barthel y la Medida de Independencia Funcional que evalúan la independencia y la funcionalidad en las AVD en personas con ictus, según la lista de verificación COSMIN más actualizada, con una escala de 4 puntos, y los estándares de Elementos de Informe Preferidos para Revisiones Sistemáticas y Metaanálisis (PRISMA) (Elsman et al., 2024; Moher et al., 2009; Mokkink et al., 2010; Terwee et al., 2012). Esta revisión sistemática identificará los estudios existentes, seleccionará y evaluará las contribuciones, analizará y sintetizará los datos, y presentará la evidencia de forma que permita extraer conclusiones razonablemente claras sobre lo que se conoce y lo que se desconoce. La lista de verificación COSMIN se utilizó para evaluar las propiedades de medición de las medidas de resultados existentes notificadas por los usuarios y seleccionar el instrumento de medición de resultados más adecuado (Elsman et al., 2024; Mokkink et al., 2010; Terwee et al., 2012). La declaración PRISMA mejora la presentación de informes de revisiones sistemáticas y metaanálisis. Este protocolo se ha registrado en Prospero.

2.3 Método de búsqueda

Tres revisores independientes realizaron una búsqueda exhaustiva durante enero y febrero de 2024 en diferentes bases de datos: Medline (Pubmed), Web of Science (WoS), Open Grey, Scielo, Cochrane, Dialnet, LILACS y OT Seeker. La estrategia de búsqueda incluyó una combinación de los siguientes cinco aspectos: 1) búsqueda de constructo, 2) búsqueda de población, 3) búsqueda de instrumentos, 4) filtro para propiedades de medición AND, y 5) filtro de exclusión "NOT" (Terwee et al., 2009). La mayor cantidad de datos disponibles fue mayor que el número de herramientas y sus diferentes versiones. Para realizar esta búsqueda, se utilizaron los siguientes términos clave: «Índice de Barthel», «BI», «Medida de Independencia Funcional», «FIM», «Accidente Cerebral», «evaluación», «herramienta», «instrumento», «prueba», «medida» y «escala», adaptándose esta estrategia a cada uno de los sistemas de búsqueda de las diferentes bases de datos. La búsqueda se limitó a artículos publicados en inglés

y español. Además, se revisó la lista de referencias de los artículos incluidos para identificar otros documentos relevantes [Tabla 1].

Tabla 1. Estrategia de búsqueda

Estrategia de búsqueda para el Índice de Barthel

Estrategia de búsqueda para Pubmed

1	("Barthel Index" OR "Barthel" OR "BI")
2	(Stroke OR Cerebral Stroke OR Cerebrovascular Accident OR Cerebral Strokes OR Strokes OR Cerebrovascular Stroke OR CVA OR Cerebrovascular OR Cerebrovascular Accidents OR Brain Accident OR Brain Accidents)
3	("Instrument" OR "instruments" OR "measure" OR "measures" OR "questionnaire" OR "questionnaires" OR "scale" OR "scales" OR "tool" OR "tools" OR "survey" OR "test")
4	("Instrumentation" OR "methods" OR "Validation Studies" OR "Comparative Study" OR "psychometrics" OR "psychometr*" OR "clinimetr*" OR "clinometr*" OR "outcome assessment" OR "outcome assessment" OR "outcome measure*" OR "observer variation" OR "observer variation" OR "Health Status Indicators" OR "reproducibility of results" OR "reproducib*" OR "discriminant analysis" OR "reliab*" OR "unreliab*" OR "valid*" OR "coefficient" OR "homogeneity" OR "homogeneous" OR "internal consistency" OR "cronbach*" OR "correlation*" OR "selection*" OR "reduction*" OR "agreement" OR "precision" OR "imprecision" OR "precise values" OR "test-retest" OR "reliab*" OR "stability" OR "interrater" OR "inter-rater" OR "intrarater" OR "intra-rater" OR "intertester" OR "inter-tester" OR "intratester" OR "intra-tester" OR "interobserver" OR "inter-observer" OR "intraobserver" OR "intraobserver" OR "intertechnician" OR "inter-technician" OR "intratechnician" OR "intra-technician" OR "interexaminer" OR "inter-examiner" OR "intraexaminer" OR "intra-examiner" OR "interassay" OR "inter-assay" OR "intraassay" OR "intra-assay" OR "interindividual" OR "inter-individual" OR "intraindividual" OR "intra-individual" OR "interparticipant" OR "inter-participant" OR "intraparticipant" OR "intra-participant" OR "kappa" OR "kappa's" OR "kappas" OR "repeatab*" OR "generaliza*" OR "generalisa*" OR "concordance" OR "intraclass correlation*" OR "discriminative" OR "known group" OR "factor analysis" OR "factor analyses" OR "dimension*" OR "subscale*" OR "item discriminant" OR "interscale correlation*" OR "error" OR "errors" OR "individual variability" OR "variability" OR "uncertainty" OR "standard

- error of measurement" OR "sensitiv*" OR "responsive*" OR "meaningful change" OR "ceiling effect" OR "floor effect" OR "Item response model" OR "IRT" OR "Rasch" OR "Differential item functioning" OR "DIF" OR "computer adaptive testing" OR "item bank" OR "cross-cultural equivalence")
- 5 #1 AND #2 AND #3 AND #4
- 6 ("Protocol" OR "addresses" OR "biography" OR "case reports" OR "comment" OR "editorial" OR "congresses" OR "consensus development conference" OR "consensus development conference" OR "practice guideline" OR "suffering from" OR "animals")
- 7 #5 NOT #6
- 8 FILTRO: Lenguaje (inglés y español)
-

Estrategia de búsqueda para Web of Science (WoS)

- 1 ("Barthel Index" OR "Barthel" OR "BI")
- 2 (Stroke OR Cerebral Stroke OR Cerebrovascular Accident OR Cerebral Strokes OR Strokes OR Cerebrovascular Stroke OR CVA OR Cerebrovascular OR Cerebrovascular Accidents OR Brain Accident OR Brain Accidents)
- 3 ("Instrument" OR "instruments" OR "measure" OR "measures" OR "questionnaire" OR "questionnaires" OR "scale" OR "scales" OR "tool" OR "tools" OR "survey" OR "test")
- 4 ("Instrumentation" OR "methods" OR "Validation Studies" OR "Comparative Study" OR "psychometrics" OR "psychometr*" OR "clinimetr*" OR "clinometr*" OR "outcome assessment" OR "outcome assessment" OR "outcome measure*" OR "observer variation" OR "observer variation" OR "Health Status Indicators" OR "reproducibility of results" OR "reproducib*" OR "discriminant analysis" OR "reliab*" OR "unreliab*" OR "valid*" OR "coefficient" OR "homogeneity" OR "homogeneous" OR "internal consistency" OR "cronbach*" OR "correlation*" OR "selection*" OR "reduction*" OR "agreement" OR "precision" OR "imprecision" OR "precise values" OR "test-retest" OR "reliab*" OR "stability" OR "interrater" OR "inter-rater" OR "intrarater" OR "intra-rater" OR "intertester" OR "inter-tester" OR "intratester" OR "intra-tester" OR "interobserver" OR "inter-observer" OR "intraobserver" OR "intraobserver" OR "intertechician" OR "inter-technician" OR "intratechnician" OR "intra-technician" OR "interexaminer" OR "inter-examiner" OR "intraexaminer" OR "intra-examiner" OR "interassay" OR "inter-assay" OR "intraassay" OR "intra-assay" OR "interindividual" OR "inter-individual" OR "intraindividual" OR "intra-individual" OR

- "interparticipant" OR "inter-participant" OR "intraparticipant" OR "intra-participant" OR "kappa" OR "kappa's" OR "kappas" OR "repeatab*" OR "generaliza*" OR "generalisa*" OR "concordance" OR "intraclass correlation*" OR "discriminative" OR "known group" OR "factor analysis" OR "factor analyses" OR "dimension*" OR "subscale*" OR "item discriminant" OR "interscale correlation*" OR "error" OR "errors" OR "individual variability" OR "variability" OR "uncertainty" OR "standard error of measurement" OR "sensitiv*" OR "responsive*" OR "meaningful change" OR "ceiling effect" OR "floor effect" OR "Item response model" OR "IRT" OR "Rasch" OR "Differential item functioning" OR "DIF" OR "computer adaptive testing" OR "item bank" OR "cross-cultural equivalence")
- 5 #1 AND #2 AND #3 AND #4
- 6 ("Protocol" OR "addresses" OR "biography" OR "case reports" OR "comment" OR "editorial" OR "congresses" OR "consensus development conference" OR "consensus development conference" OR "practice guideline" OR "suffering from" OR "animals")
- 7 #5 NOT #6
- 8 FILTRO: Lenguaje (inglés y español), Título y Abstract

Estrategia de búsqueda para Open Grey

- 1 ("Barthel Index" OR "Barthel" OR "BI")
- 2 (Stroke OR Cerebral Stroke OR Cerebrovascular Accident OR Cerebral Strokes OR Strokes OR Cerebrovascular Stroke OR CVA OR Cerebrovascular OR Cerebrovascular Accidents OR Brain Accident OR Brain Accidents)
- 3 ("Instrument" OR "instruments" OR "measure" OR "measures" OR "questionnaire" OR "questionnaires" OR "scale" OR "scales" OR "tool" OR "tools" OR "survey" OR "test")
- 4 ("Instrumentation" OR "methods" OR "Validation Studies" OR "Comparative Study" OR "psychometrics" OR "psychometr*" OR "clinimetr*" OR "clinometr*" OR "outcome assessment" OR "outcome assessment" OR "outcome measure*" OR "observer variation" OR "observer variation" OR "Health Status Indicators" OR "reproducibility of results" OR "reproducib*" OR "discriminant analysis" OR "reliab*" OR "unreliab*" OR "valid*" OR "coefficient" OR "homogeneity" OR "homogeneous" OR "internal consistency" OR "cronbach*" OR "correlation*" OR "selection*" OR "reduction*" OR "agreement" OR "precision" OR "imprecision" OR "precise values" OR "test-retest" OR

- "reliab*" OR "stability" OR "interrater" OR "inter-rater" OR "intrarater" OR "intra-rater" OR "intertester" OR "inter-tester" OR "intratester" OR "intra-tester" OR "interobserver" OR "inter-observer" OR "intraobserver" OR "intraobserver" OR "intertechician" OR "inter-technician" OR "intratechnician" OR "intra-technician" OR "interexaminer" OR "inter-examiner" OR "intraexaminer" OR "intra-examiner" OR "interassay" OR "inter-assay" OR "intraassay" OR "intra-assay" OR "interindividual" OR "inter-individual" OR "intraindividual" OR "intra-individual" OR "interparticipant" OR "inter-participant" OR "intraparticipant" OR "intra-participant" OR "kappa" OR "kappa's" OR "kappas" OR "repeatable" OR "generalizable" OR "generalisable" OR "concordance" OR "intraclass correlation" OR "discriminative" OR "known group" OR "factor analysis" OR "factor analyses" OR "dimension*" OR "subscale*" OR "item discriminant" OR "interscale correlation*" OR "error" OR "errors" OR "individual variability" OR "variability" OR "uncertainty" OR "standard error of measurement" OR "sensitive*" OR "responsive*" OR "meaningful change" OR "ceiling effect" OR "floor effect" OR "Item response model" OR "IRT" OR "Rasch" OR "Differential item functioning" OR "DIF" OR "computer adaptive testing" OR "item bank" OR "cross-cultural equivalence")
- 5 #1 AND #2 AND #3 AND #4
- 6 ("Protocol" OR "addresses" OR "biography" OR "case reports" OR "comment" OR "editorial" OR "congresses" OR "consensus development conference" OR "consensus development conference" OR "practice guideline" OR "suffering from" OR "animals")
- 7 #5 NOT #6
- 8 FILTRO: Lenguaje (inglés y español)

Estrategia de búsqueda para Scielo

- 1 ("Barthel Index" OR "Barthel")
- 2 ("Stroke")
- 3 ("Instrument" OR "instruments" OR "measure" OR "measures" OR "questionnaire" OR "questionnaires" OR "scale" OR "scales" OR "tool" OR "tools" OR "survey" OR "test")
- 4 ("psychometr*" OR "validat*" OR "feasibil*" OR "reliab*" OR "consistenc*" OR "clinimetr*")
- 5 #1 AND #2 AND #3 AND #4

Estrategia de búsqueda para Cochrane

- 1 (Barthel Index OR Barthel OR BI)
- 2 (Stroke OR Cerebral Stroke OR Cerebrovascular Accident OR Cerebral Strokes OR Strokes OR Cerebrovascular Stroke OR CVA OR Cerebrovascular OR Cerebrovascular Accidents OR Brain Accident OR Brain Accidents)
- 3 ("Instrument" OR "instruments" OR "measure" OR "measures" OR "questionnaire" OR "questionnaires" OR "scale" OR "scales" OR "tool" OR "tools" OR "survey" OR "test")
- 4 (Instrumentation OR methods OR Validation Studies OR Comparative Study OR psychometrics OR psychometr OR clinimetr OR clinometr OR outcome assessment OR outcome assessment OR outcome measure OR observer variation OR observer variation OR Health Status Indicators OR reproducibility of results OR reproducib OR discriminant analysis OR reliab OR unreliab OR valid OR coefficient OR homogeneity OR homogeneous OR internal consistency OR cronbach OR correlation OR selection OR reduction OR agreement OR precision OR imprecision OR precise values OR test-retest OR reliab OR stability OR interrater OR inter-rater OR intrarater OR intra-rater OR intertester OR inter-tester OR intratester OR intra-tester OR interobserver OR inter-observer OR intraobserver OR intraobserver OR intertechnician OR inter-technician OR intratechnician OR intra-technician OR interexaminer OR inter-examiner OR intraexaminer OR intra-examiner OR interassay OR inter-assay OR intraassay OR intra-assay OR interindividual OR inter-individual OR intraindividual OR intra-individual OR interparticipant OR inter-participant OR intraparticipant OR intra-participant OR kappa OR kappas OR kappas OR repeatab OR generaliza OR generalisa OR concordance OR intraclass correlation OR discriminative OR known group OR factor analysis OR factor analyses OR dimension OR subscale OR item discriminant OR interscale correlation OR error OR errors OR individual variability OR variability OR uncertainty OR standard error of measurement OR sensitiv OR responsive OR meaningful change OR ceiling effect OR floor effect OR Item response model OR IRT OR Rasch OR Differential item functioning OR DIF OR computer adaptive testing OR item bank OR cross-cultural equivalence)
- 5 #1 AND #2 AND #3 AND #4
- 6 ("Protocol" OR "addresses" OR "biography" OR "case reports" OR

"comment" OR "editorial" OR "congresses" OR "consensus development conference" OR "consensus development conference" OR "practice guideline" OR "suffering from" OR "animals")

7 #5 NOT #6

8 FILTRO: Lenguaje (inglés y español)

Estrategia de búsqueda para Dialnet

1 (Barthel)

2 (Stroke)

3 (Instrument OR measure OR questionnaire OR scale OR tool OR survey OR test)

4 (psychometr* OR reliab* OR consitenc* OR clinimetr*)

5 #1 AND #2 AND #3 AND #4

6 FILTRO: Lenguaje (inglés y español)

Estrategia de búsqueda para LILACS

1 ("Barthel Index" OR "Barthel" OR "BI")

2 (Stroke OR Cerebral Stroke OR Cerebrovascular Accident OR Cerebral Strokes OR Strokes OR Cerebrovascular Stroke OR CVA OR Cerebrovascular OR Cerebrovascular Accidents OR Brain Accident OR Brain Accidents)

3 ("Instrument" OR "instruments" OR "measure" OR "measures" OR "questionnaire" OR "questionnaires" OR "scale" OR "scales" OR "tool" OR "tools" OR "survey" OR "test")

4 ("Instrumentation" OR "methods" OR "Validation Studies" OR "Comparative Study" OR "psychometrics" OR "psychometr*" OR "clinimetr*" OR "clinometr*" OR "outcome assessment" OR "outcome assessment" OR "outcome measure*" OR "observer variation" OR "observer variation" OR "Health Status Indicators" OR "reproducibility of results" OR "reproducib*" OR "discriminant analysis" OR "reliab*" OR "unreliab*" OR "valid*" OR "coefficient" OR "homogeneity" OR "homogeneous" OR "internal consistency" OR "cronbach*" OR "correlation*" OR "selection*" OR "reduction*" OR "agreement" OR "precision" OR "imprecision" OR "precise values" OR "test-retest" OR

"reliab*" OR "stability" OR "interrater" OR "inter-rater" OR "intrarater" OR "intra-rater" OR "intertester" OR "inter-tester" OR "intratester" OR "intra-tester" OR "interobserver" OR "inter-observer" OR "intraobserver" OR "intraobserver" OR "intertechician" OR "inter-technician" OR "intratechnician" OR "intra-technician" OR "interexaminer" OR "inter-examiner" OR "intraexaminer" OR "intra-examiner" OR "interassay" OR "inter-assay" OR "intraassay" OR "intra-assay" OR "interindividual" OR "inter-individual" OR "intraindividual" OR "intra-individual" OR "interparticipant" OR "inter-participant" OR "intraparticipant" OR "intra-participant" OR "kappa" OR "kappa's" OR "kappas" OR "repeatab*" OR "generaliza*" OR "generalisa*" OR "concordance" OR "intraclass correlation*" OR "discriminative" OR "known group" OR "factor analysis" OR "factor analyses" OR "dimension*" OR "subscale*" OR "item discriminant" OR "interscale correlation*" OR "error" OR "errors" OR "individual variability" OR "variability" OR "uncertainty" OR "standard error of measurement" OR "sensitiv*" OR "responsive*" OR "meaningful change" OR "ceiling effect" OR "floor effect" OR "Item response model" OR "IRT" OR "Rasch" OR "Differential item functioning" OR "DIF" OR "computer adaptive testing" OR "item bank" OR "cross-cultural equivalence")

5 #1 AND #2 AND #3 AND #4

6 FILTRO: Lenguaje (inglés y español)

Estrategia de búsqueda para OT Seeker

- 1 ("Barthel Index" OR "Barthel" OR "BI")
- 2 (Stroke OR Cerebral Stroke OR Cerebrovascular Accident OR Cerebral Strokes OR Strokes OR Cerebrovascular Stroke OR CVA OR Cerebrovascular OR Cerebrovascular Accidents OR Brain Accident OR Brain Accidents)
- 3 ("Instrument" OR "instruments" OR "measure" OR "measures" OR "questionnaire" OR "questionnaires" OR "scale" OR "scales" OR "tool" OR "tools" OR "survey" OR "test")
- 4 ("Instrumentation" OR "methods" OR "Validation Studies" OR "Comparative Study" OR "psychometrics" OR "psychometr*" OR "clanimetr*" OR "clinometr*" OR "outcome assessment" OR "outcome assessment" OR "outcome measure*" OR "observer variation" OR "observer variation" OR "Health Status Indicators" OR "reproducibility of results" OR "reproducib*" OR "discriminant analysis" OR "reliab*" OR "unreliab*" OR "valid*" OR "coefficient" OR "homogeneity" OR "homogeneous" OR "internal consistency" OR "cronbach*" OR

"correlation*" OR "selection*" OR "reduction*" OR "agreement" OR "precision" OR "imprecision" OR "precise values" OR "test-retest" OR "reliab*" OR "stability" OR "interrater" OR "inter-rater" OR "intrarater" OR "intra-rater" OR "intertester" OR "inter-tester" OR "intratester" OR "intra-tester" OR "interobserver" OR "inter-observer" OR "intraobserver" OR "intraobserver" OR "intertechician" OR "inter-technician" OR "intratechnician" OR "intra-technician" OR "interexaminer" OR "inter-examiner" OR "intraexaminer" OR "intra-examiner" OR "interassay" OR "inter-assay" OR "intraassay" OR "intra-assay" OR "interindividual" OR "inter-individual" OR "intraindividual" OR "intra-individual" OR "interparticipant" OR "inter-participant" OR "intraparticipant" OR "intra-participant" OR "kappa" OR "kappa's" OR "kappas" OR "repeatab*" OR "generaliza*" OR "generalisa*" OR "concordance" OR "intraclass correlation*" OR "discriminative" OR "known group" OR "factor analysis" OR "factor analyses" OR "dimension*" OR "subscale*" OR "item discriminant" OR "interscale correlation*" OR "error" OR "errors" OR "individual variability" OR "variability" OR "uncertainty" OR "standard error of measurement" OR "sensitiv*" OR "responsive*" OR "meaningful change" OR "ceiling effect" OR "floor effect" OR "Item response model" OR "IRT" OR "Rasch" OR "Differential item functioning" OR "DIF" OR "computer adaptive testing" OR "item bank" OR "cross-cultural equivalence")

5 #1 AND #2 AND #3 AND #4

6 FILTRO: Lenguaje (inglés y español)

Estrategia de búsqueda para la Medida de Independencia Funcional

Estrategia de búsqueda para Pubmed

- 1 ("Functional Independence Measure" OR "FIM")
- 2 (Stroke OR Cerebral Stroke OR Cerebrovascular Accident OR Cerebral Strokes OR Strokes OR Cerebrovascular Stroke OR CVA OR Cerebrovascular OR Cerebrovascular Accidents OR Brain Accident OR Brain Accidents)
- 3 ("Instrument" OR "instruments" OR "measure" OR "measures" OR "questionnaire" OR "questionnaires" OR "scale" OR "scales" OR "tool" OR "tools" OR "survey" OR "test")
- 4 ("Instrumentation" OR "methods" OR "Validation Studies" OR "Comparative Study" OR "psychometrics" OR "psychometr*" OR "clinimetr*" OR "clinometr*" OR "outcome assessment" OR "outcome

assessment" OR "outcome measure*" OR "observer variation" OR "observer variation" OR "Health Status Indicators" OR "reproducibility of results" OR "reproducib*" OR "discriminant analysis" OR "reliab*" OR "unreliab*" OR "valid*" OR "coefficient" OR "homogeneity" OR "homogeneous" OR "internal consistency" OR "cronbach*" OR "correlation*" OR "selection*" OR "reduction*" OR "agreement" OR "precision" OR "imprecision" OR "precise values" OR "test-retest" OR "reliab*" OR "stability" OR "interrater" OR "inter-rater" OR "intrarater" OR "intra-rater" OR "intertester" OR "inter-tester" OR "intratester" OR "intra-tester" OR "interobserver" OR "inter-observer" OR "intraobserver" OR "intraobserver" OR "intertechician" OR "inter-technician" OR "intratechnician" OR "intra-technician" OR "interexaminer" OR "inter-examiner" OR "intraexaminer" OR "intra-examiner" OR "interassay" OR "inter-assay" OR "intraassay" OR "intra-assay" OR "interindividual" OR "inter-individual" OR "intraindividual" OR "intra-individual" OR "interparticipant" OR "inter-participant" OR "intraparticipant" OR "intra-participant" OR "kappa" OR "kappa's" OR "kappas" OR "repeatab*" OR "generaliza*" OR "generalisa*" OR "concordance" OR "intraclass correlation*" OR "discriminative" OR "known group" OR "factor analysis" OR "factor analyses" OR "dimension*" OR "subscale*" OR "item discriminant" OR "interscale correlation*" OR "error" OR "errors" OR "individual variability" OR "variability" OR "uncertainty" OR "standard error of measurement" OR "sensitiv*" OR "responsive*" OR "meaningful change" OR "ceiling effect" OR "floor effect" OR "Item response model" OR "IRT" OR "Rasch" OR "Differential item functioning" OR "DIF" OR "computer adaptive testing" OR "item bank" OR "cross-cultural equivalence")

5 #1 AND #2 AND #3 AND #4

6 ("Protocol" OR "addresses" OR "biography" OR "case reports" OR "comment" OR "editorial" OR "congresses" OR "consensus development conference" OR "consensus development conference" OR "practice guideline" OR "suffering from" OR "animals")

7 #5 NOT #6

8 FILTRO: Lenguaje (inglés y español)

Estrategia de búsqueda para Web of Science (WoS)

1 ("Functional Independence Measure" OR "FIM")

2 (Stroke OR Cerebral Stroke OR Cerebrovascular Accident OR Cerebral Strokes OR Strokes OR Cerebrovascular Stroke OR CVA OR Cerebrovascular OR Cerebrovascular Accidents OR Brain Accident OR

Brain Accidents)

- 3 ("Instrument" OR "instruments" OR "measure" OR "measures" OR "questionnaire" OR "questionnaires" OR "scale" OR "scales" OR "tool" OR "tools" OR "survey" OR "test")
- 4 ("Instrumentation" OR "methods" OR "Validation Studies" OR "Comparative Study" OR "psychometrics" OR "psychometr*" OR "clinimetr*" OR "clinometr*" OR "outcome assessment" OR "outcome assessment" OR "outcome measure*" OR "observer variation" OR "observer variation" OR "Health Status Indicators" OR "reproducibility of results" OR "reproducib*" OR "discriminant analysis" OR "reliab*" OR "unreliab*" OR "valid*" OR "coefficient" OR "homogeneity" OR "homogeneous" OR "internal consistency" OR "cronbach*" OR "correlation*" OR "selection*" OR "reduction*" OR "agreement" OR "precision" OR "imprecision" OR "precise values" OR "test-retest" OR "reliab*" OR "stability" OR "interrater" OR "inter-rater" OR "intrarater" OR "intra-rater" OR "intertester" OR "inter-tester" OR "intratester" OR "intra-tester" OR "interobserver" OR "inter-observer" OR "intraobserver" OR "intraobserver" OR "intertechnician" OR "inter-technician" OR "intratechnician" OR "intra-technician" OR "interexaminer" OR "inter-examiner" OR "intraexaminer" OR "intra-examiner" OR "interassay" OR "inter-assay" OR "intraassay" OR "intra-assay" OR "interindividual" OR "inter-individual" OR "intraindividual" OR "intra-individual" OR "interparticipant" OR "inter-participant" OR "intraparticipant" OR "intra-participant" OR "kappa" OR "kappa's" OR "kappas" OR "repeatab*" OR "generaliza*" OR "generalisa*" OR "concordance" OR "intraclass correlation*" OR "discriminative" OR "known group" OR "factor analysis" OR "factor analyses" OR "dimension*" OR "subscale*" OR "item discriminant" OR "interscale correlation*" OR "error" OR "errors" OR "individual variability" OR "variability" OR "uncertainty" OR "standard error of measurement" OR "sensitiv*" OR "responsive*" OR "meaningful change" OR "ceiling effect" OR "floor effect" OR "Item response model" OR "IRT" OR "Rasch" OR "Differential item functioning" OR "DIF" OR "computer adaptive testing" OR "item bank" OR "cross-cultural equivalence")
- 5 #1 AND #2 AND #3 AND #4
- 6 ("Protocol" OR "addresses" OR "biography" OR "case reports" OR "comment" OR "editorial" OR "congresses" OR "consensus development conference" OR "consensus development conference" OR "practice guideline" OR "suffering from" OR "animals")
- 7 #5 NOT #6
- 8 FILTRO: Lenguaje (inglés y español), Título y Abstract

Estrategia de búsqueda para Open Grey

- 1 ("Functional Independence Measure" OR "FIM")
- 2 (Stroke OR Cerebral Stroke OR Cerebrovascular Accident OR Cerebral Strokes OR Strokes OR Cerebrovascular Stroke OR CVA OR Cerebrovascular OR Cerebrovascular Accidents OR Brain Accident OR Brain Accidents)
- 3 ("Instrument" OR "instruments" OR "measure" OR "measures" OR "questionnaire" OR "questionnaires" OR "scale" OR "scales" OR "tool" OR "tools" OR "survey" OR "test")
- 4 ("Instrumentation" OR "methods" OR "Validation Studies" OR "Comparative Study" OR "psychometrics" OR "psychometr*" OR "clinimetr*" OR "clinometr*" OR "outcome assessment" OR "outcome assessment" OR "outcome measure*" OR "observer variation" OR "observer variation" OR "Health Status Indicators" OR "reproducibility of results" OR "reproducib*" OR "discriminant analysis" OR "reliab*" OR "unreliab*" OR "valid*" OR "coefficient" OR "homogeneity" OR "homogeneous" OR "internal consistency" OR "cronbach*" OR "correlation*" OR "selection*" OR "reduction*" OR "agreement" OR "precision" OR "imprecision" OR "precise values" OR "test-retest" OR "reliab*" OR "stability" OR "interrater" OR "inter-rater" OR "intrarater" OR "intra-rater" OR "intertester" OR "inter-tester" OR "intratester" OR "intra-tester" OR "interobserver" OR "inter-observer" OR "intraobserver" OR "intraobserver" OR "intertechnician" OR "inter-technician" OR "intratechnician" OR "intra-technician" OR "interexaminer" OR "inter-examiner" OR "intraexaminer" OR "intra-examiner" OR "interassay" OR "inter-assay" OR "intraassay" OR "intra-assay" OR "interindividual" OR "inter-individual" OR "intraindividual" OR "intra-individual" OR "interparticipant" OR "inter-participant" OR "intraparticipant" OR "intra-participant" OR "kappa" OR "kappa's" OR "kappas" OR "repeatab*" OR "generaliza*" OR "generalisa*" OR "concordance" OR "intraclass correlation*" OR "discriminative" OR "known group" OR "factor analysis" OR "factor analyses" OR "dimension*" OR "subscale*" OR "item discriminant" OR "interscale correlation*" OR "error" OR "errors" OR "individual variability" OR "variability" OR "uncertainty" OR "standard error of measurement" OR "sensitiv*" OR "responsive*" OR "meaningful change" OR "ceiling effect" OR "floor effect" OR "Item response model" OR "IRT" OR "Rasch" OR "Differential item functioning" OR "DIF" OR "computer adaptive testing" OR "item bank" OR "cross-cultural equivalence")
- 5 #1 AND #2 AND #3 AND #4

- 6 ("Protocol" OR "addresses" OR "biography" OR "case reports" OR "comment" OR "editorial" OR "congresses" OR "consensus development conference" OR "consensus development conference" OR "practice guideline" OR "suffering from" OR "animals")
- 7 #5 NOT #6
- 8 FILTRO: Lenguaje (inglés y español)

Estrategia de búsqueda para Scielo

- 1 ("Functional Independence Measure" OR "FIM")
- 2 ("Stroke")
- 3 ("Instrument" OR "instruments" OR "measure" OR "measures" OR "questionnaire" OR "questionnaires" OR "scale" OR "scales" OR "tool" OR "tools" OR "survey" OR "test")
- 4 ("psychometr*" OR "validat*" OR "feasibil*" OR "reliab*" OR "consitenc*" OR "clanimetr*")
- 5 #1 AND #2 AND #3 AND #4
- 6 FILTRO: Lenguaje (inglés y español)

Estrategia de búsqueda para Cochrane

- 1 (Functional Independence Measure OR FIM)
- 2 (Stroke OR Cerebral Stroke OR Cerebrovascular Accident OR Cerebral Strokes OR Strokes OR Cerebrovascular Stroke OR CVA OR Cerebrovascular OR Cerebrovascular Accidents OR Brain Accident OR Brain Accidents)
- 3 ("Instrument" OR "instruments" OR "measure" OR "measures" OR "questionnaire" OR "questionnaires" OR "scale" OR "scales" OR "tool" OR "tools" OR "survey" OR "test")
- 4 (Instrumentation OR methods OR Validation Studies OR Comparative Study OR psychometrics OR psychometr OR clinimetr OR clinometr OR outcome assessment OR outcome assessment OR outcome measure OR observer variation OR observer variation OR Health Status Indicators OR reproducibility of results OR reproducib OR discriminant analysis OR reliab OR unreliab OR valid OR coefficient OR homogeneity OR homogeneous OR internal consistency OR cronbach OR correlation OR selection OR reduction OR agreement OR precision)

- OR imprecision OR precise values OR test-retest OR reliab OR stability
OR interrater OR inter-rater OR intrarater OR intra-rater OR intertester
OR inter-tester OR intratester OR intra-tester OR interobserver OR
inter-observer OR intraobserver OR intraobserver OR intertechnician
OR inter-technician OR intratechnician OR intra-technician OR
interexaminer OR inter-examiner OR intraexaminer OR intra-examiner
OR interassay OR inter-assay OR intraassay OR intra-assay OR
interindividual OR inter-individual OR intraindividual OR intra-individual
OR interparticipant OR inter-participant OR intraparticipant OR intra-
participant OR kappa OR kappas OR kappas OR repeatab OR
generaliza OR generalisa OR concordance OR intraclass correlation
OR discriminative OR known group OR factor analysis OR factor
analyses OR dimension OR subscale OR item discriminant OR
interscale correlation OR error OR errors OR individual variability OR
variability OR uncertainty OR standard error of measurement OR
sensitivity OR responsive OR meaningful change OR ceiling effect OR
floor effect OR Item response model OR IRT OR Rasch OR Differential
item functioning OR DIF OR computer adaptive testing OR item bank
OR cross-cultural equivalence)
- 5 #1 AND #2 AND #3 AND #4
- 6 ("Protocol" OR "addresses" OR "biography" OR "case reports" OR
"comment" OR "editorial" OR "congresses" OR "consensus
development conference" OR "consensus development conference" OR
"practice guideline" OR "suffering from" OR "animals")
- 7 #5 NOT #6
- 8 FILTRO: Lenguaje (inglés y español)

Estrategia de búsqueda para Dialnet

-
- 1 (Functional Independence Measure OR FIM)
- 2 (Stroke)
- 3 (psychometr* OR validat* OR feasibil* OR reliab* OR consitenc* OR
clanimetr*)
- 4 #1 AND #2 AND #3
- 5 FILTRO: Lenguaje (inglés y español)
-

Estrategia de búsqueda para LILACS

-
- 1 ("Functional Independence Measure" OR "FIM")
 - 2 (Stroke OR Cerebral Stroke OR Cerebrovascular Accident OR Cerebral Strokes OR Strokes OR Cerebrovascular Stroke OR CVA OR Cerebrovascular OR Cerebrovascular Accidents OR Brain Accident OR Brain Accidents)
 - 3 ("Instrument" OR "instruments" OR "measure" OR "measures" OR "questionnaire" OR "questionnaires" OR "scale" OR "scales" OR "tool" OR "tools" OR "survey" OR "test")
 - 4 ("Instrumentation" OR "methods" OR "Validation Studies" OR "Comparative Study" OR "psychometrics" OR "psychometr*" OR "clinimetr*" OR "clinometr*" OR "outcome assessment" OR "outcome assessment" OR "outcome measure*" OR "observer variation" OR "observer variation" OR "Health Status Indicators" OR "reproducibility of results" OR "reproducib*" OR "discriminant analysis" OR "reliab*" OR "unreliab*" OR "valid*" OR "coefficient" OR "homogeneity" OR "homogeneous" OR "internal consistency" OR "cronbach*" OR "correlation*" OR "selection*" OR "reduction*" OR "agreement" OR "precision" OR "imprecision" OR "precise values" OR "test-retest" OR "reliab*" OR "stability" OR "interrater" OR "inter-rater" OR "intrarater" OR "intra-rater" OR "intertester" OR "inter-tester" OR "intratester" OR "intra-tester" OR "interobserver" OR "inter-observer" OR "intraobserver" OR "intraobserver" OR "intertechician" OR "inter-technician" OR "intratechnician" OR "intra-technician" OR "interexaminer" OR "inter-examiner" OR "intraexaminer" OR "intra-examiner" OR "interassay" OR "inter-assay" OR "intraassay" OR "intra-assay" OR "interindividual" OR "inter-individual" OR "intraindividual" OR "intra-individual" OR "interparticipant" OR "inter-participant" OR "intraparticipant" OR "intra-participant" OR "kappa" OR "kappa's" OR "kappas" OR "repeatab*" OR "generaliza*" OR "generalisa*" OR "concordance" OR "intraclass correlation*" OR "discriminative" OR "known group" OR "factor analysis" OR "factor analyses" OR "dimension*" OR "subscale*" OR "item discriminant" OR "interscale correlation*" OR "error" OR "errors" OR "individual variability" OR "variability" OR "uncertainty" OR "standard error of measurement" OR "sensitiv*" OR "responsive*" OR "meaningful change" OR "ceiling effect" OR "floor effect" OR "Item response model" OR "IRT" OR "Rasch" OR "Differential item functioning" OR "DIF" OR "computer adaptive testing" OR "item bank" OR "cross-cultural equivalence")
 - 5 #1 AND #2 AND #3 AND #4
 - 6 FILTRO: Lenguaje (inglés y español)
-

Estrategia de búsqueda para OT Seeker

-
1	("Functional Independence Measure" OR "FIM")
2	(Stroke OR Cerebral Stroke OR Cerebrovascular Accident OR Cerebral Strokes OR Strokes OR Cerebrovascular Stroke OR CVA OR Cerebrovascular OR Cerebrovascular Accidents OR Brain Accident OR Brain Accidents)
3	("Instrument" OR "instruments" OR "measure" OR "measures" OR "questionnaire" OR "questionnaires" OR "scale" OR "scales" OR "tool" OR "tools" OR "survey" OR "test")
4	("Instrumentation" OR "methods" OR "Validation Studies" OR "Comparative Study" OR "psychometrics" OR "psychometr*" OR "clanimetr*" OR "clinometr*" OR "outcome assessment" OR "outcome assessment" OR "outcome measure*" OR "observer variation" OR "observer variation" OR "Health Status Indicators" OR "reproducibility of results" OR "reproducib*" OR "discriminant analysis" OR "reliab*" OR "unreliab*" OR "valid*" OR "coefficient" OR "homogeneity" OR "homogeneous" OR "internal consistency" OR "cronbach*" OR "correlation*" OR "selection*" OR "reduction*" OR "agreement" OR "precision" OR "imprecision" OR "precise values" OR "test-retest" OR "reliab*" OR "stability" OR "interrater" OR "inter-rater" OR "intrarater" OR "intra-rater" OR "intertester" OR "inter-tester" OR "intratester" OR "intra-tester" OR "interobserver" OR "inter-observer" OR "intraobserver" OR "intraobserver" OR "intertechician" OR "inter-technician" OR "intratechnician" OR "intra-technician" OR "interexaminer" OR "inter-examiner" OR "intraexaminer" OR "intra-examiner" OR "interassay" OR "inter-assay" OR "intraassay" OR "intra-assay" OR "interindividual" OR "inter-individual" OR "intraindividual" OR "intra-individual" OR "interparticipant" OR "inter-participant" OR "intraparticipant" OR "intra-participant" OR "kappa" OR "kappa's" OR "kappas" OR "repeatab*" OR "generaliza*" OR "generalisa*" OR "concordance" OR "intraclass correlation*" OR "discriminative" OR "known group" OR "factor analysis" OR "factor analyses" OR "dimension*" OR "subscale*" OR "item discriminant" OR "interscale correlation*" OR "error" OR "errors" OR "individual variability" OR "variability" OR "uncertainty" OR "standard error of measurement" OR "sensitiv*" OR "responsive*" OR "meaningful change" OR "ceiling effect" OR "floor effect" OR "Item response model" OR "IRT" OR "Rasch" OR "Differential item functioning" OR "DIF" OR "computer adaptive testing" OR "item bank" OR "cross-cultural equivalence")
5	#1 AND #2 AND #3 AND #4
6	FILTRO: Lenguaje (inglés y español)
-

Fuente: elaboración propia

2.4 Criterios de elegibilidad

Los criterios de inclusión para este estudio fueron: a) personas que habían sufrido un ictus, b) el resultado debía ser la independencia y funcionalidad en las AVD con cualquier versión de BI o FIM, c) los estudios debían incluir una versión de BI o FIM y evaluar al menos una propiedad psicométrica de la lista de verificación COSMIN.

Los criterios de exclusión fueron: a) instrumento o artículo no disponible en idioma inglés o español, b) editoriales, discursos, conferencias, biografías, comentarios, reportes de casos, tesis doctorales, guías de práctica, artículos de opinión o artículos en los que no se utilizó metodología selectiva.

Tres revisores examinaron de forma independiente los títulos y resúmenes para comprobar si cumplían los criterios de inclusión. El artículo se recuperó cuando fue imposible tomar decisiones basándose únicamente en el resumen y el título, y fue revisado de nuevo por los revisores. Las discrepancias se resolvieron compartiendo los diferentes aspectos problemáticos y discutiéndolos entre los tres revisores hasta llegar a un acuerdo.

2.5 Evaluación de la calidad

Los estudios incluidos en esta revisión fueron evaluados por tres autores, de acuerdo con la lista de verificación de riesgo de sesgo COSMIN más actualizada con una escala de 4 puntos (Elsman et al., 2024; Terwee et al., 2012) (disponible en <https://www.cosmin.nl/>). Contiene nueve casillas con entre 5 y 18 ítems, y cada casilla proporciona una puntuación de calidad metodológica para la validez de contenido, validez estructural, consistencia interna, validez transcultural, error de medición, fiabilidad, validez de criterio, prueba de hipótesis y sensibilidad. El contenido evaluado en cada casilla puede calificarse como "muy bueno", "adecuado", "dudoso" e "inadecuado". La puntuación de cada propiedad psicométrica se obtiene tomando la puntuación más baja de cualquier ítem de cada propiedad (Terwee et al., 2012). Ningún artículo fue rechazado en función de la calidad metodológica del instrumento incluido.

2.6 Extracción de datos

Se revisaron los títulos y resúmenes de todos los artículos identificados en esta revisión. Tres expertos revisaron de forma independiente los artículos completos y las discrepancias se resolvieron por consenso. La información de cada artículo se extrajo con base en los siguientes aspectos: nombre de la herramienta y versión utilizada, población de estudio, características de la herramienta, análisis de propiedades psicométricas, puntuación COSMIN y resultados estadísticos de la investigación.

2.7 Síntesis

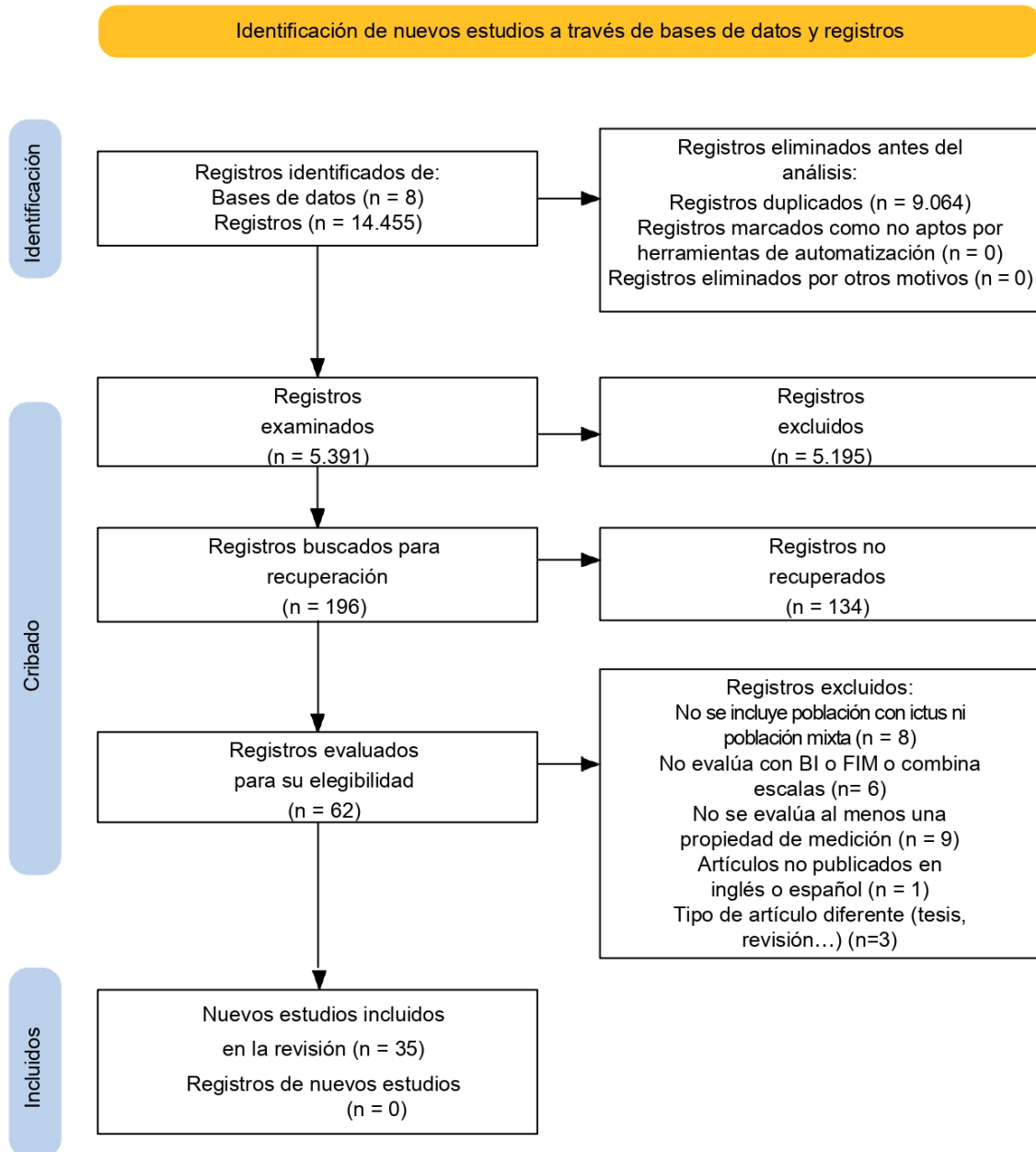
Los datos sobre las propiedades psicométricas de los artículos que cumplieron con los criterios de inclusión se calificaron según la lista de verificación COSMIN (Terwee et al., 2012). Tres revisores independientes establecieron las puntuaciones de las propiedades psicométricas de cada artículo y las diferencias se resolvieron mediante discusión y consenso. Posteriormente, el resultado de cada una de las propiedades de medición de cada estudio se calificó según los criterios de Terwee (Prinsen et al., 2018), clasificando cada uno de los resultados como “positivo” (+), “negativo” (-) o “indeterminado” (?). Finalmente, después de determinar la calidad metodológica de cada uno de los estudios individuales y evaluar las propiedades de medición, los resultados se resumieron cualitativamente para determinar tanto la calidad de la evidencia como la solidez de las recomendaciones, calificándolas como “altas”, “moderadas”, “bajas” o “muy bajas” según el riesgo de sesgo, inconsistencia, imprecisión e indirecteidad, utilizando el sistema GRADE (Guyatt et al., 2008). Tres revisores independientes asignaron puntuaciones para las propiedades psicométricas de cada artículo y las diferencias se resolvieron mediante discusión y consenso.

3.RESULTADOS

3.1 Resultados de la búsqueda

Se identificaron un total de 14.455 artículos, de los cuales se examinaron 5.391 títulos y resúmenes porque 9.064 se eliminaron por ser duplicados. De estos, 5.195 artículos fueron excluidos después de esta evaluación del título. Después de eso, de estos 196 artículos, 134 fueron excluidos con base en los criterios de elegibilidad en la revisión del resumen. Posteriormente, se examinaron 62 artículos de texto completo. De estos, 26 artículos fueron excluidos después de revisar el texto completo (8 estudios fueron excluidos debido a una población mixta, seis artículos debido a la falta de evaluación de BI o FIM o escalas combinadas, nueve artículos por no evaluar al menos una propiedad de medición incluida en COSMIN, 1 estudio no publicado en los idiomas seleccionados y tres artículos debido al tipo de estudio (revisión)). Finalmente, 35 estudios fueron incluidos en esta revisión sistemática de los dos instrumentos a evaluar. El diagrama de flujo se muestra en la Figura 1.

Figura 1. Diagrama de flujo de la selección de artículos



Fuente: elaboración propia

3.2 Versiones identificadas del Índice Barthel

Se identificaron veinte versiones del índice de Barthel entre 31 artículos diseñados para evaluar la independencia y la funcionalidad en las actividades de la vida diaria en personas con ictus: Índice de Barthel (Original) (BI) (Loewen et al., 1988; Shah et al., 1989; Wolfe et al., 1991; Hachisuka et al., 1997a; van der Putten et al., 1999; Gosman-Hedstrom et al., 2000; Green et al., 2001; Hsueh et al., 2001; Hsueh et al., 2002; van Hartingsveld et al., 2006; Hsieh et al., 2007; Sarker et al., 2012; Zhou et al., 2021; Yang et al., 2022), Índice de Barthel Iraní (Ira-BI) (Aminalroaya et al., 2021), Índice de Barthel Italiano (Ita-BI) (Galeoto et al., 2019); Índice de Barthel Persa (Per-BI) (Oveisgharan et al., 2006; Naghdi et al., 2016), Índice de Barthel Telefónico (Tel-BI) (Prasad et al., 2018), Índice de Barthel Entrevistado (Int-BI) (Sandaria et al., 2001), Índice de Barthel Auto-puntuado (SR-BI) (Hachisuka et al., 1997a; Hachisuka et al., 1997b), Índice de Barthel de 5 ítems (BI-5) (Park, 2018; Hsueh et al., 2002), Índice de Barthel de 8 ítems (BI-8) (van Hartingsveld et al., 2006), Índice de Barthel Extendido (16-ítems) (BI-16) (Jansa et al., 2004), Escalas Suplementarias del Índice de Barthel (Escala de dificultad autopercebida, Escala de habilidad) (BI-SS) (Lee et al., 2014; Lee et al., 2017), Modificación de Granger del Índice de Barthel (G-BI) (Hachisuka et al., 1997a), Entrevista telefónica con el Índice de Barthel modificado (Granger) (Tel-mBI) (Shinar et al., 1987), Evaluación de medida del desempeño con el Índice de Barthel modificado (Granger) (PMA-mBI) (Shinar et al., 1987), Índice de Barthel modificado (Shah) (MBI) (Shah et al., 1989; Yang et al., 2021; Yang et al., 2022), Versión china del Índice de Barthel modificado (Shah) (Chi-MBI) (Leung et al., 2007), Versión brasileña del Índice de Barthel modificado (Shah) (Bra-MBI) (Cincura et al., 2009), Versión iraní del Índice de Barthel modificado (Shah) (Ira-MBI) (Aminalroaya et al., 2021), Versión turca del Índice de Barthel modificado (Shah) (Tur-MBI) (Kuşukdeveci et al., 2000), y Herramienta de evaluación del desempeño del Índice de Barthel modificado (Shah) (PET-MBI) (Ohura et al., 2017). Esta información se refleja en la Tabla 2.

Tabla 2. Características de los estudios y descripción del índice de Barthel y la medida de independencia funcional en personas con ictus.

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Índice de Barthel (original) (BI)						
Loewen et al. (1988)	Siete pacientes hospitalizados con ictus en diferentes niveles de recuperación.	Hospital General St. Boniface, Canadá.	El índice de Barthel contiene 10 ítems que evalúan las actividades de la vida diaria (AVD), con una puntuación total de 0 a 100, que indica independencia en las AVD.	1. Fiabilidad 2. Prueba de hipótesis 3. Sensibilidad	1. Dudoso 2. Muy bueno 3. Muy bueno	1. Para el índice de Barthel, Los porcentajes de acuerdo variaron entre 69% a 100%; y los valores medios de Kappa oscilaron entre 0,47 y 1,00. 2. Para el índice de Barthel, el rango de coeficientes de correlación de orden de Spearman fue de 0,95 a 1,00 con una mediana de 0,99.
Shah et al. (1989)	Se utilizaron datos de un estudio prospectivo de rehabilitación de ictus de los	Hospitales de Brisbane, Australia, que fueron remitidos para rehabilitación integral para pacientes	Comparación del índice de Barthel original y el índice de Barthel modificado por Shah.	1. Consistencia interna	1. Muy bueno	1. El alfa de Cronbach del BI en la puntuación inicial es 0,867 y en el alta es 0,924.

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación de COSMIN	Valores de medición
	258 pacientes hospitalizados sobrevivientes durante de un primer ictus. 1984.					
Wolfe et al. (1991)	Se evaluaron cincuenta pacientes con diagnóstico clínico de haber sufrido un ictus al menos tres meses antes.	Hospital St. Thomas. Reino Unido.	Índice de Barthel como arriba.	1. Fiabilidad	1. Dudoso	1. La K no ponderada fue de 0,72 (intervalo de confianza del 95 %: 0,61-0,82) y los pesos de desacuerdo cuadráticos de 0,91 (0,87-0,95).
Hachisuka et al. (1997)a	Cuarenta y un pacientes participaron en el estudio comparativo. Ciento dos pacientes se	Se realizó un estudio comparativo en el Hospital Universitario de Salud Ocupacional y Ambiental de Japón. El estudio test-retest	Índice de Barthel como arriba.	1. Consistencia interna 2. Validez de criterio 3. Prueba de hipótesis 4. Sensibilidad	1. Muy bueno 2. Muy bueno 3. Muy bueno 4. Muy	1. El alfa de Cronbach del BI es 0,805. 2. La correlación de Spearman entre BI y SR-BI es 0,994. La correlación de Spearman entre BI y G-BI es 0,993. 3. La correlación de Spearman

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
	incluyeron en el estudio test-retest.	se realizó en siete hospitales afiliados en Japón.			bueno	entre BI y FIM-Mot-13 es 0,901. La correlación de Spearman entre BI y FIM-Mot-3lev es 0,850.
van der Putten et al. (1999)	Pacientes con esclerosis múltiple (n=201) y ictus (n=82) en una unidad de neurorrehabilitación entre 1994 y 1997.	Hospital Nacional de Neurología y Neurocirugía, Reino Unido, entre 1994 y 1997.	Índice de Barthel como arriba.	1. Prueba de hipótesis 2. Sensibilidad	1. Inadecuado 2. Inadecuado	1. Comparación de las puntuaciones de cambio de BI y FIM: BI=5.2 FIM Total=21,9 FIM Motor= 19.1 FIM Cognitivo=2.8
Gosman-Hedstrom et al. (2000)	Doscientos cuatro pacientes con ictus.	Hospital Universitario Sahlgrenska. Suecia.	Índice de Barthel como arriba.	1. Prueba de hipótesis 2. Sensibilidad	1. Dudoso 2. Dudoso	1. Porcentaje de concordancia (PA) como medida de concordancia entre escalas. El PA en los niveles categóricos entre las escalas calibradas es de 0,94 (PA = (11 + 32 + 148) /

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Green et al. (2001)	Veintidós pacientes con problemas de movilidad al menos un año después del ictus.	Hospital St. Luke, NHS, Bradford, West Yorkshire, Reino Unido.	Índice de Barthel como arriba.	1. Fiabilidad 2. Prueba de hipótesis 3. Sensibilidad	1. Inadecuado 2. Dudoso 3. Dudoso	204 = 0,94), lo que significa una concordancia del 94 % entre los 3 niveles condensados. 1. El porcentaje de acuerdo fue alto (75% para el BI; 85% para el RMI; 60% para el NEADL y el FAI). 2. Los porcentajes de acuerdo fueron en general altos, especialmente para el BI (475%) y el RMI (485%), pero hubo una variación considerable en los coeficientes kappa.

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Hsueh et al. (2001)	Se evaluaron ciento veintiún pacientes utilizando el BI a los 14, 30, 90 y 180 días después del ictus.	Hospital Universitario Nacional de Taiwán entre diciembre-1 de julio de 1999 y 30 de septiembre de 2000.	Índice de Barthel como arriba.	1. Consistencia interna 2. Fiabilidad 3. Prueba de hipótesis 4. Sensibilidad	1. Muy bueno 2. Dudoso 3. Muy bueno 4. Muy bueno	1. Las puntuaciones de alfa de Cronbach oscilaron entre 0,89 y 0,92, lo que indica una excelente consistencia interna. 2. La fiabilidad entre evaluadores en los ítems individuales, las estadísticas kappa ponderadas para cada uno de los 10 ítems variaron de 0,53 a 0,94 (mediana, 0,72), lo que indica un acuerdo de moderado a excelente. En cuanto a la fiabilidad entre evaluadores de la puntuación total, el ICC fue 0,94 (IC del 95 %: 0,91-0,96, $p < 0,0001$), lo que indica un acuerdo entre evaluadores muy alto. 3. Las puntuaciones de BI estuvieron altamente correlacionadas con las

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
						puntuaciones de BBS y FM en cada etapa de recuperación del ictus, lo que indica una buena validez de constructo convergente (r de Pearson $\geq 0,78$, $p < 0,0001$). Las puntuaciones de BI a 14, 30 y 90 días tras el ictus se correlacionaron moderadamente con las puntuaciones de FAI a 180 días tras el ictus (r de Pearson $\geq 0,59$, $p < 0,0001$).
Hsueh et al. (2002)	En el estudio participaron 118 pacientes hospitalizados con ictus en una unidad de rehabilitación.	Hospital Nacional Universitario de Taiwán entre el 1 de diciembre de 1999 y Mayo de 2000.	Índice de Barthel como arriba.	1. Consistencia interna 2. Fiabilidad 3. Validez de criterio 4. Prueba de hipótesis	1. Muy bueno 2. Inadecuado 3. Muy bueno 4. Muy	1. Alfa de Cronbach = 0,85 2. CCI=0,86 3. La correlación de Spearman entre BI y BI-5 fue de 0,94 4. La correlación de Spearman entre BI y FIM-Mot fue de 0,94

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
van Hartingsveld et al. (2006)	Se evaluaron 559 pacientes con ictus a las 3 semanas (n=89) y a los 6 meses (n=470) después del ictus.	Los autores utilizaron información de dos bases de datos de dos estudios previos realizados en los Países Bajos.	Índice de Barthel como arriba.	1. Validez estructural	5. Sensibilidad bueno 5. Muy bueno	1. El ajuste general de la escala revisada fue excelente (estadística R1c P0,35), lo que indica que el modelo de Rasch extendido es un modelo de medición válido para Barthel y que los ítems juntos definen un continuo de construcción de escala de AVD unidimensional.

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	de	Puntuación COSMIN	Valores de medición
Hsieh et al. (2007)	Se reclutaron 43 pacientes con ictus en el estudio basado en anclas. Se reclutaron 56 pacientes con ictus en el estudio basado en la distribución.	Los participantes fueron reclutados de los Departamentos de Medicina Física y Rehabilitación de 3 hospitales en Taiwán entre septiembre de 2004 y julio de 2005.	Índice de Barthel como arriba.	1. Fiabilidad 2. Error de medición		1. Inadecuado 2. Dudoso	1. Si la puntuación media de mejora del BI dentro de un grupo de ACV es menor que el MCID de 1,85 puntos, aunque el cambio de puntuación ha alcanzado un nivel estadísticamente significativo ($P < .05$), el cambio no debe considerarse como un cambio verdadero e importante. Por el contrario, si la puntuación media de mejora del BI supera los 1,85 puntos, incluso si la puntuación de cambio no ha alcanzado la significación estadística ($P > .05$), el cambio puede ser clínicamente relevante.

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Sarker et al. (2012)	Doscientos treinta y ocho pacientes del Registro de ictus del Sur de Londres, basado en la población.	Fundación del Hospital St. Thomas. Hospital del King's College. Hospital de Westminster.	Índice de Barthel como arriba.	1. Fiabilidad 2. Prueba de hipótesis 3. Sensibilidad	1. Dudoso 2. Muy bueno 3. Muy bueno	1. ICC BI vs FAI = 0,27 ICC BI frente a NEADL = 0,55 2. BI frente a FAI $r=0,80$ BI frente a NEADL $r=0,88$
Zhou et al. (2021)	Se incluyeron en este estudio 5.475 pacientes.	El estudio se realizó en el Departamento de Rehabilitación de 88 hospitales y centros de rehabilitación comunitarios en toda China desde diciembre de 2018 hasta mayo de 2020.	Índice de Barthel como arriba.	1. Prueba de hipótesis 2. Sensibilidad	1. Muy bueno 2. Muy bueno	1. Las calificaciones de LS se asociaron positivamente con BI ($r = 0,925$, $p < 0,001$).
Yang et al. (2022)	Participaron un total de 60 pacientes	El estudio se llevó a cabo en un departamento	Índice de Barthel como arriba.	1. Fiabilidad 2. Error de medición	1. Adecuado 2. Adecuado 3.	1. Tanto BI como MBI tuvieron ICC altos (0,94 para ambos). 3. Se encontraron correlaciones

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	de	Puntuación COSMIN	Valores de medición
	ambulatorios con ictus crónico.	ambulatorio de un centro médico en Taichung, Taiwán, desde agosto de 2018 hasta octubre de 2018.		3. Validez de criterio 4. Sensibilidad	de	Inadecuado 4. Inadecuado	moderadas ($r < 0,30$) entre las medias de los pacientes y los valores absolutos de cambio tanto para el BI como para el MBI.
Índice Barthel iraní (Ira-BI)							
Aminalroaya y otros (2021)	100 pacientes con ictus hospitalizados ≥ 60 años.	Hospital Shohadaye Tajrish en Teherán, Irán, de febrero a junio de 2019.	Los autores utilizaron las versiones iraníes de BI y mBI para validar ambas en personas con ictus.	1. Validez de criterio 2. Sensibilidad	de	1. Muy bueno 2. Muy bueno	1. Ira-BI frente a Ira-MBI = 0,946
Índice Barthel italiano (Ita-BI)							
Galeoto et al. (2019)	99 pacientes hospitalizados con ictus isquémico.	Unidad de Ictus del Departamento de Neurociencias, Policlínico Tor Vergata, Roma, Italia, entre mayo de 2016 y	Los autores utilizaron una versión italiana del índice de Barthel para validarlo en pacientes con ictus.	1. Validez estructural 2. Consistencia interna 3. Fiabilidad		1. Dudoso 2. Muy bueno 3. Inadecuado	1. El análisis EFA de factores de eje principal reveló una estructura de dos factores correlacionados. Los índices de ajuste del CFA confirmaron una estructura de dos factores.

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
		septiembre de 2016.				2. El alfa de Cronbach fue 0,901 ($p < 0,001$). 3. Fiabilidad intraevaluador: El ICC fue de 0,987 (IC del 95 %: 0,975-0,993), lo que indica una fiabilidad intraevaluador óptima. El índice Kappa de Cohen fue de 0,69, lo que indica una concordancia sustancial. Fiabilidad entre evaluadores: ICC = 0,909; CI del 95 %: 0,852-0,948.
Índice Barthel Persa (Per-BI)						
Oveisgharan et al. (2006)	Los cuestionarios traducidos fueron completados mediante entrevistas	Unidad de Registro de ictus del Centro de Investigación Cardiovascular de Isfahán, un centro colaborador de la OMS en el centro de	Los autores realizaron una traducción del Índice Barthel original mencionado con anterioridad a una versión persa para validarlo en la población con ictus.	1. Consistencia interna 2. Validez transcultural 3. Fiabilidad 4. Prueba de hipótesis	1. Muy bueno 2. Inadecuado 3. Inadecuado 4. Muy	1. El alfa de Cronbach de la versión traducida al persa del índice de Barthel fue 0,9354. 3. El coeficiente kappa fue $0,616 \pm 0,084$. El coeficiente de correlación de Spearman entre 2

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
	con 459 pacientes con ictus. El mismo entrevistador preguntó a 58 pacientes dos veces durante dos días consecutivos para evaluar la fiabilidad.	Irán.		5. Sensibilidad	bueno 5. Muy bueno	puntuaciones fue 0,989 ($p < 0,001$). Se calculó la correlación entre las puntuaciones Per-BI y las puntuaciones MRS. El coeficiente de correlación de Spearman fue de -0,912 ($p < 0,001$).
Naghdi et al. (2016)	Se incluyeron en este estudio 40 pacientes con ictus.	Centro Tabassom de Rehabilitación de ictus en Teherán, Irán.	Los autores hicieron una comparación entre Per-BI y Per-FIM.	1. Prueba de hipótesis 2. Sensibilidad	1. Muy bueno 2. Muy bueno	1. Hubo una correlación excelente significativa entre Per-FIM y Per-BI demostrada por un análisis de correlación de Pearson de 0,95 ($p < 0,001$).
Índice Barthel Telefónico (Tel-BI)						

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Prasad et al. (2018)	Se reclutaron un total de 120 pacientes (60 intervenciones y 60 controles) con ictus subagudos de cinco hospitales docentes de diferentes partes de la India entre noviembre de 2008 y junio de 2010.	Los cinco hospitales docentes fueron: Instituto de Ciencias Médicas de la India (AIIMS), Nueva Delhi (Centro 1). Instituto de Posgrado de Ciencias Médicas Sanjay Gandhi, Lucknow (Centro 2). Facultad de Medicina de las Fuerzas Armadas (AFMC), Pune (Centro 3). Instituto de Postgrado de Ciencias Médicas (PGI), Chandigarh (Centro 4). Investigación y referencia	Los autores analizan las propiedades psicométricas de la evaluación con BI a través de una llamada telefónica.	1. Consistencia interna 2. Fiabilidad 3. Validez de criterio 4. Prueba de hipótesis 5. Sensibilidad	1. Muy bueno 2. Inadecuado 3. Adecuado 4. Adecuado 5. Adecuado	1. Alfa de Cronbach a 3 meses: 0,8914 Alfa de Cronbach a los 6 meses: 0,9464 2. Se observó un acuerdo sustancial entre los dos métodos a los 3 y 6 meses con una wK de 0,63 (IC del 95 %: 0,61-0,68) y 0,72 (IC del 95 %: 0,70-0,77), respectivamente. 3. Considerando la evaluación presencial como el estándar de oro, investigamos la validez del IB administrado por teléfono. Se observó una sensibilidad del 89,2 % al establecer el punto de corte en una puntuación de $IB \leq 60$ y ≥ 20 , y la mayor especificidad (95 %) al establecer el punto de corte en una puntuación de $IB \leq 80$ y $>$

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
		Hospital (R&R), Nueva Delhi (Centro 5).				80 durante 3 y 6 meses, respectivamente. 4. La BI telefónica está altamente correlacionada con el mRS y el NIHSS a los 3 y 6 meses. Se observó una excelente correlación (- 0,9) entre mRS y BI evaluados telefónicamente.
Índice de Barthel Entrevistado (Int-BI)						
Sandaria et al.	82 pacientes con ictus.	Unidad neurológica del Hospital Hartford, Hartford, Connecticut, EE.UU.	Los autores utilizaron un índice de Barthel entrevistado con pacientes con ictus.	1. Consistencia interna 2. Prueba de hipótesis 3. Sensibilidad	1. Muy bueno 2. Muy bueno 3. Muy bueno	1. El alfa de Cronbach fue: BI en la entrevista previa al ictus = 0,828, y BI en la entrevista posterior al ictus = 0,952. 2. La correlación entre las puntuaciones totales de estos instrumentos antes del ictus fue moderada (0,645). Sin embargo, las correlaciones

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
						entre las puntuaciones del BI y la FIM después del ictus fueron excelentes (0,913 a 0,971). La correlación de Spearman entre las puntuaciones de la FIM para la entrevista y la observación del terapeuta indicó una correlación entre moderada y buena entre estos dos métodos de administración. Estas correlaciones oscilaron entre 0,508 y 0,753 para cada ítem. La correlación entre las puntuaciones totales del subconjunto de la FIM obtenidas mediante la entrevista y la observación fue de 0,751. Todas las correlaciones fueron significativas con un valor de $p < 0,001$.

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Autoevaluación del Índice de Barthel (SR-BI)						
Hachisuka et al. (1997)a	Cuarenta y un pacientes participaron en el estudio comparativo. Ciento dos pacientes se incluyeron en el estudio test-retest.	Se realizó un estudio comparativo en el Hospital Universitario de Salud Ocupacional y Ambiental de Japón. El estudio test-retest se realizó en siete hospitales afiliados en Japón.	El SR-BI consta de 13 actividades de la vida diaria (AVD) principales y 2 actividades complementarias (modificación de Granger). Cada ítem se puntúa con 0, 5 o 10, según el nivel de dependencia de la persona. La puntuación total oscila entre 0 y 100.	1. Consistencia interna 2. Fiabilidad 3. Validez de criterio 4. Prueba de hipótesis 5. Sensibilidad	1. Muy bueno 2. Adecuado 3. Muy bueno 4. Muy bueno 5. Muy bueno	1. El alfa de Cronbach del SR-BI es 0,842. 2. Los coeficientes Kappa varían entre 0,587 (controlar los movimientos intestinales) y 0,823 (entrar y salir de la bañera o la ducha). 3. La correlación de Spearman entre SR-BI y BI es 0,994. 4. La correlación de Spearman entre SR-BI y FIM-Mot-13 es 0,904.
Hachisuka et al. (1997)b	Datos obtenidos de diferentes bases de datos	El primer estudio involucró a 31 pacientes del Hospital Universitario en 1992, de el segundo estudio	Índice de Barthel de autoevaluación como el anterior.	1. Fiabilidad 2. Prueba de hipótesis 3. Sensibilidad	1. Dudoso 2. Inadecuado 3. Inadecuado	1. Los coeficientes de correlación intraclass del índice de autocuidado, índice de movilidad y puntuación total del SR-BI versión 3, entre las tres

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
	estudios, con un total de 236 pacientes.	involucró a 140 pacientes (44 del Hospital Universitario, 55 del Hospital Chikuho Rosai y 41 del Hospital Moji Rosai) en 1993, y el tercer estudio involucró a 65 pacientes del Hospital Universitario en 1994.				calificaciones fueron 0,854, 0,889 y 0,867, respectivamente. 2. Los coeficientes de correlación de Spearman del índice de autocuidado, índice de movilidad y puntuación total del SR-BI versión 1, fueron 0,454, 0,824 y 0,768, respectivamente. Los coeficientes de correlación de Spearman del índice de autocuidado, índice de movilidad y puntuación total del SR-BI versión 2 fueron 0,790, 0,717 y 0,791, respectivamente.

Índice de Barthel de 5 ítems (BI-5)

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Park (2018)	24 pacientes con ictus crónico.	Hospital General de Corea, de agosto de 2016 a diciembre de 2016.	Reducción de 5 ítems de la versión original de 10 ítems del Índice de Barthel. Los 5 ítems seleccionados fueron traslados, baño, uso del inodoro, subir escaleras y deambulaci3n.	1. Consistencia interna 2. Fiabilidad 3. Error de medici3n 4. Validez de criterio 5. Prueba de hip3tesis 6. Sensibilidad	1. Muy bueno 2. Inadecuado 3. Inadecuado 4. Muy bueno 5. Muy bueno 6. Muy bueno	1. El coeficiente α de Cronbach de las puntuaciones totales del SF-BI fue 0,87. 2. CCI=0,91 (0,86–0,95). 4. Se analiz3 el coeficiente de correlaci3n de las puntuaciones totales del BI-5 y las puntuaciones totales del BI original, y fue muy alto ($r = 0,95$). 5. Las puntuaciones totales del BI-5 se correlacionaron significativamente con las puntuaciones totales del PASS ($r=0,81$) y las puntuaciones totales del FMA ($r=0,76$), respectivamente.
Hsueh et al. (2002)	En el estudio participaron 118 pacientes hospitalizados	Hospital Nacional Universitario de Taiwán entre el 1 de diciembre de 1999 y	Índice de Barthel de 5 ítems como el anterior.	1. Consistencia interna 2. Fiabilidad 3. Validez de	1. Muy bueno 2. Inadecuado	1. Alfa de Cronbach = 0,73 2. ICC=0,74 3. La correlaci3n de Spearman entre BI-5 y BI fue de 0,94

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
	con ictus en el 31 una unidad de rehabilitación.	Mayo de 2000.		criterio 4. Prueba de hipótesis 5. Sensibilidad	3. Muy bueno 4. Muy bueno 5. Muy bueno	4.La correlación de Spearman entre BI-5 y FIM-Mot fue de 0,92
Índice de Barthel de 8 ítems (BI-8)						
van Hartingsveld et al. (2006)	Se evaluaron 559 pacientes con ictus a las 3 semanas (n=89) y a los 6 meses (n=470) después del ictus.	Los autores utilizaron información de dos bases de datos de dos estudios previos realizados en los Países Bajos.	Versión de 8 ítems del Índice Barthel. Similar a la versión original, sin elementos de control de intestino ni de vejiga.	1. Validez estructural 2. Consistencia interna	1. Muy bueno 2. Muy bueno	1. El ajuste general de la escala revisada fue excelente (estadística R1c P0,35), lo que indica que el modelo de Rasch extendido es un modelo de medición válido para Barthel y que los ítems juntos definen un continuo de construcción de escala de AVD unidimensional. 2. La fiabilidad de consistencia interna (Cronbach) de la escala revisada de 8 ítems resultante no se alteró (0,93).

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Índice de Barthel ampliado (16 ítems) (BI-16)						
Jansa et al. (2004)	33 pacientes con ictus isquémico agudo.	Universidad Médica Centro Ljubljana, Departamento de Neurología. Eslovenia.	Versión ampliada de 16 elementos del Índice Barthel, con algunos elementos del BI original y algunos elementos del FIM.	1. Validez estructural 2. Consistencia interna 3. Validez de criterio 4. Prueba de hipótesis 5. Sensibilidad	1. Inadecuado 2. Muy bueno 3. Muy bueno 4. Muy bueno 5. Muy bueno	1. La validez de constructo se realizó mediante análisis factorial, el cual arrojó tres factores: físico con un autovalor de 8,24, cognitivo-comunicación con un autovalor de 2,75 y cognitivo con un autovalor de 0,9. El factor 1 agrupó los 10 ítems de la parte física del BI ampliado, el factor 2 agrupó cuatro ítems de la parte cognitiva (comprensión, expresión, interacción social y resolución de problemas) y el factor 3 agrupó ítems sobre memoria y negligencia. El análisis factorial de la parte

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
						física del BI ampliado mostró dos factores: el factor 1, con un autovalor de 7,38, agrupó los ítems sobre alimentación, aseo personal, movilidad, etc., y el factor 2, con un autovalor de 1,18, agrupó los ítems sobre función intestinal y vesical. 2. La fiabilidad interna de la puntuación total del BI-16 aumentó de la semana 1 a la 6 (0,89-0,94). La consistencia interna de la parte física del BI-16 aumentó de la semana 1 a la 6 (0,90-0,96). La consistencia interna de la parte cognitiva del BI ampliado-16 disminuyó de la semana 1 a la 6 (0,84-0,79). 3. BI-16 vs BI, de $r=0,89$ (semana 1) a $r=0,98$ (semana

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
						6). 4. BI-16 vs FMA, de $r=0,66$ (semana 1) a $r=0,87$ (semana 6).
Escalas suplementarias del índice de Barthel (escala de dificultad autopercebida, escala de capacidad) (BI-SS)						
Lee et al. (2014)	12 pacientes con ictus participaron en la prueba piloto. Otros 306 participantes participaron en este estudio.	Se reclutaron pacientes sometidos a rehabilitación ambulatoria o hospitalaria de siete departamentos de rehabilitación de Taiwán (incluidas las zonas norte (4 hospitales), central (2	Desarrollo de dos escalas complementarias para el Índice de Barthel: Escala de Dificultad Autopercebida y Escala de Habilidad.	1. Validez de contenido 2. Validez estructural 3. Validez de criterio 4. Sensibilidad	1. Dudoso 2. Inadecuado 3. Muy bueno 4. Muy bueno	2. Los resultados mostraron que la unidimensionalidad de las dos escalas del constructo AVD fue fuerte ($H \geq 0,63$). 3. Los resultados mostraron un alto grado de correlación entre el BI original y las dos escalas ($\rho = 0,78$ y $0,90$, respectivamente), indicando que ambos constructos

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
						medidos por el BI-SS, dificultad y capacidad autopercebidas, estaban altamente relacionados con el constructo de desempeño real en pacientes con ictus.
Lee et al. (2017)	En el estudio participaron 84 pacientes ambulatorios con ictus crónico.	Departamentos de Medicina Física y Rehabilitación de siete hospitales de Taiwán, entre marzo de 2011 y mayo de 2012.	Escalas suplementarias del índice de Barthel como las anteriores.	1. Fiabilidad 2. Error de medición 3. Validez de criterio 4. Sensibilidad	1. 2. Inadecuado 3. Inadecuado 3. Muy bueno 4. Muy bueno	1. Los ICC para la Escala de Capacidad y la Escala de Dificultad Autopercebida fueron 0,97 y 0,78, respectivamente. 3. La r de Pearson entre BI y BI-SS (Escala de Capacidad) fue de 0,62 (0,42-0,76, P<0,001). La r de Pearson entre BI y BI-SS (Escala de dificultad autopercebida) fue de 0,23 (-0,03-0,47, P=0,080).
Modificación de Granger del índice						

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
de Barthel (G-BI)						
Hachisuka et al. (1997)a	Cuarenta y un pacientes participaron en el estudio comparativo. Ciento dos pacientes se incluyeron en el estudio test-retest.	Se realizó un estudio comparativo en el Hospital Universitario de Salud Ocupacional y Ambiental de Japón. El estudio test-retest se realizó en siete hospitales afiliados en Japón.	El Índice de Barthel de Granger es una modificación del Índice de Barthel original. Consta de 13 ítems principales y 2 complementarios de las AVD.	1. Consistencia interna 2. Validez de criterio 3. Prueba de hipótesis 4. Sensibilidad	1. Muy bueno 2. Muy bueno 3. Muy bueno 4. Muy bueno	1. El alfa de Cronbach del G-BI es 0,791. 2. La correlación de Spearman entre G-BI y BI es 0,993. 3. La correlación de Spearman entre G-BI y FIM-Mot-13 es 0,901.
Entrevista telefónica con el Índice de Barthel modificado (Granger) (Tel-mBI)						
Shinar et al. (1987)	En el estudio participaron 33 hombres y 39 mujeres con ictus.	Hospital Reese, Illinois, Estados Unidos.	Michael Chicago, Estados Unidos. Un índice de Barthel modificado, basado en el Stroke Data Bank, con 15 ítems, similar a la versión de Granger; evaluando entrevista	1. Fiabilidad 2. Validez de criterio 3. Sensibilidad	1. Inadecuado 2. Muy bueno 3. Muy	1. La fiabilidad entre los administradores de las AVD y entre los observadores de la administración de las AVD fue r = 0,99 para las puntuaciones

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Evaluación de medidas de desempeño con el índice de Barthel modificado (Granger) (PMA-mBI)			telefónica y medida de desempeño.		bueno	totales y $r > 0,90$ para la mayoría de los elementos de los componentes individuales. 2. Las correlaciones entre PMA-mBI y Tel-mBI fueron $r > 0,97$ para la puntuación total y $r > 0,85$ para la mayoría de los ítems individuales.
Shinar et al. (1987)	En el estudio participaron 33 hombres y 39 mujeres con ictus.	Hospital Reese, Illinois, Unidos.	Michael Chicago, Estados Unidos. Un índice de Barthel modificado, basado en el Stroke Data Bank, con 15 ítems, similar a la versión de Granger; evaluando entrevista telefónica y medida de	1. Fiabilidad 2. Validez de criterio 3. Sensibilidad	1. Inadecuado 2. Muy bueno 3. Muy bueno	1. La Fiabilidad entre los administradores de las AVD y entre los observadores de la administración de las AVD fue $r = 0,99$ para las puntuaciones totales y $r > 0,90$ para la

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
			desempeño.			mayoría de los elementos de los componentes individuales. 2. Las correlaciones entre PMA-mBI y Tel-mBI fueron $r > 0,97$ para la puntuación total y $r > 0,85$ para la mayoría de los ítems individuales.
Índice de Barthel Modificado (Shah) (MBI)						
Shah et al. (1989)	Datos de un estudio prospectivo sobre rehabilitación de ictus. Se utilizaron los 258 pacientes supervivientes de un ictus.	Hospitales de Brisbane, Australia, que fueron remitidos para rehabilitación integral para pacientes hospitalizados durante 1984.	Una modificación del índice de Barthel con los mismos ítems y un rango de puntuación mayor (1=no puede realizar la tarea, 2=intenta realizar la tarea, pero no es seguro, 3=requiere ayuda moderada, 4=requiere ayuda mínima, 5=totalmente independiente), para proporcionar más sensibilidad a la escala.	1. Consistencia interna	1. Muy bueno	1. El alfa de Cronbach del MBI en la puntuación inicial es de 0,900 y en el momento del alta es de 0,927.

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Yang et al. (2021)	Revisión retrospectiva de los datos de estudios consecutivos de pacientes reclutados en el estudio CIRCLE entre noviembre de 2018 y noviembre de 2019.	Departamento de Neurología de The Second Affiliated Hospital, Facultad de Medicina de la Universidad de Zhejiang. China.	Índice de Barthel modificado por Shah como se indica arriba.	1. Validez estructural 2. Fiabilidad	1. Muy bueno 2. Inadecuado	1. Los resultados de los análisis de Rasch mostraron que la fiabilidad del MBI en esta muestra de pacientes con IS fue buena tanto para personas como para ítems, con valores de 0,88 y 1,00, respectivamente, estos resultados indican que los ítems de la escala fueron válidos y confiables para evaluar las AVD del paciente. 2. El valor del índice de separación de personas obtenido fue de 3,73, lo que se asoció con un coeficiente de fiabilidad de 0,88. Asimismo, el valor del índice de separación de ítems fue de 20,87, lo que se asoció con un coeficiente de fiabilidad de 1,00.

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Yang et al. (2022)	Participaron un total de 60 pacientes ambulatorios con ictus crónico.	El estudio se llevó a cabo en un departamento ambulatorio de un centro médico en Taichung, Taiwán, desde agosto de 2018 hasta octubre de 2018.	Índice de Barthel modificado por Shah como se indica arriba.	1. Fiabilidad 2. Error de medición 3. Validez criterio 4. Sensibilidad	1. Adecuado 2. Adecuado 3. 4. Inadecuado	1. Tanto BI como MBI tuvieron ICC altos (0,94 para ambos). 3. Se encontraron correlaciones moderadas ($r < 0,30$) entre las medias de los pacientes y los valores absolutos de cambio tanto para el BI como para el MBI.
Índice de Barthel modificado chino (Shah) (Chi-MBI)						
Leung et al. (2007)	En el análisis de validez estructural participaron 116 personas con ictus. En la evaluación de fiabilidad	Un hospital general regional y un hospital de día en Hong Kong. Centro Médico Cáritas.	Los autores realizaron una traducción del Índice Barthel modificado por Shah mencionado con anterioridad a una versión china utilizando un panel de expertos.	1. Validez estructural 2. Consistencia interna 3. Validez transcultural 4. Fiabilidad	1. Adecuado 2. Muy bueno 3. Muy bueno 4. Dudoso	1. Los resultados sugirieron una solución de dos factores, que explicó el 75,7% de la varianza total. 2. La consistencia interna, estimada mediante el alfa de Cronbach, fue similar entre el Chi-MBI y el MBI. El valor alfa para el primero fue de 0,93 ($p \leq$

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
	participaron 15 pacientes con ictus.					0,001), mientras que para el segundo fue de 0,92 ($p \leq 0,001$). 4. Los valores kappa, como medida de fiabilidad entre evaluadores a nivel de ítem, estuvieron entre 0,81 y 1,00.
Índice de Barthel Modificado Brasileño (Shah) (Bra-MBI)						
Cinturón et al. (2009)	Se evaluaron 132 pacientes (84 con mRS, 57 con BI y 62 con PV-NIHSS).	Universidad Federal de Bahía, Brasil y Universidad de São Paulo en Ribeirão Preto, Brasil. Entre julio de 2005 y julio de 2006.	Los autores realizaron una traducción del Índice Barthel Modificado de Shah mencionado con anterioridad a una versión brasileña mediante una reunión entre investigadores.	1. Validez transcultural 2. Fiabilidad 3. Prueba de hipótesis 4. Sensibilidad	1. Inadecuado 2. Inadecuado 3. Muy bueno 4. Muy bueno	2. En el análisis Bra-MBI, el coeficiente de correlación intraclase fue de 0,967 (IC del 95 % = 0,94-0,98). Al considerar las cinco categorías de la escala, el kappa ponderado fue de 0,70 (IC del 95 % = 0,61-0,96). 3. El rho de Spearman entre

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Índice de Barthel Modificado Iraní (Shah) (Ira-MBI)						PV-NIHSS y Bra-MBI fue - 0,438 (p=0,003).
Aminalroaya y otros (2021)	100 pacientes con ictus hospitalizados ≥60 años.	Hospital Shohadaye Tajrish en Teherán, Irán, de febrero a junio de 2019.	Los autores utilizaron las versiones iraníes de BI y mBI para validar ambas en personas con ictus.	1. Validez estructural 2. Consistencia interna 3. Fiabilidad 4. Validez de criterio 5. Prueba de hipótesis 6. Sensibilidad	1. Adecuado 2. Muy bueno 3. Inadecuado 4. Muy bueno 5. Muy bueno 6. Muy bueno	1. En el CFA, los autores encontraron un factor 2. α de Cronbach = 0,955, $p < 0,001$. 3. La fiabilidad del MBI iraní fue significativa en 0,955. 4. Ira-MBI frente a Ira-BI = 0,946 5. Ira-MBI frente a MMSE = 0,469 Ira-MBI frente a GDS-15 = 0,359 Ira-MBI frente a KATZ = 0,93
Índice turco de Barthel modificado (Shah) (Tur-MBI)						

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Kuçukdeveci et al. (2000)	100 pacientes (50 con ictus y 50 con lesión medular).	Departamento de Medicina Física y Rehabilitación del Instituto Médico Facultad de la Universidad de Ankara, Turquía, de 1993 a 1997.	Los autores realizaron una traducción del Índice de Barthel modificado por Shah mencionado con anterioridad a una versión turca realizada por cuatro profesionales de la salud.	1. Validez estructural 2. Consistencia interna 3. Validez transcultural 4. Fiabilidad	1. Inadecuado 2. Muy bueno 3. Inadecuado 4. Inadecuado	1. Existe cierta duda sobre la unidimensionalidad de los ítems dados que no se ajustan al modelo de Rasch. El análisis de Rasch mostró que los ítems de vejiga e intestino comprometen la unidimensionalidad. 2. Al ingreso el alfa fue de 0,927 y al alta de 0,930 para el grupo con ictus. 4. El nivel de acuerdo entre los dos evaluadores fue suficiente, reflejado en niveles Kappa superiores a 0,6 con ictus.
Herramienta de evaluación del desempeño: Índice de Barthel Modificado (Shah) (PET-MBI)						

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación de COSMIN	Valores de medición
Ohura et al. (2017)	30 pacientes con ictus evaluados utilizando tanto la prueba directa y métodos de evaluación de vídeo.	Los pacientes fueron reclutados y evaluados en cinco hospitales japoneses diferentes.	Es la versión de evaluación funcional del índice de Barthel modificado por Shah. Se evalúa mediante la observación del desempeño del paciente en las actividades de la vida diaria (AVD).	1. Fiabilidad 2. Validez de criterio 3. Sensibilidad	1. Inadecuado 2. Muy bueno 3. Muy bueno	1. Se obtuvo un ICC de 0,99-1,00 para las puntuaciones totales del PET-MBI. Finalmente, se calcularon de forma independiente los coeficientes κ y κ_w de las puntuaciones de 10 ítems del PET-MBI como resultado secundario, con valores de 0,78-0,92 y 0,85-0,96, respectivamente. 2. El coeficiente de correlación de Pearson entre la IB y la PET-MBI fue de 0,95, y el límite inferior del intervalo de confianza (CI) del 95% fue de 0,90. Se obtuvieron valores idénticos mediante el coeficiente de correlación de rangos de Spearman.

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Medida de Independencia Funcional (Original) (MIF)						
Heinemann et al. (1994)	27.699 pacientes hospitalizados por rehabilitación por ictus.	Más de 190 instalaciones se suscribieron al Sistema Uniforme de Datos (UDS) para rehabilitación médica desde enero de 1987 hasta junio de 1989.	FIM tiene 18 ítems ADL con una escala de siete niveles.	1. Validez estructural 2. Error de medición	1. Muy bueno 2. Inadecuado	1. El análisis factorial de los ítems motores arrojó un factor principal que explica el 96,0 % de la varianza en la dificultad de los ítems (valor propio: 8,42). Las cargas factoriales variaron de 0,43 para el grupo de quemados a 1,08 para el grupo de amputados. Gracias a este estadístico, se confirmó una estructura bifactorial.
van der Putten et al. (1999)	Pacientes con esclerosis múltiple (n=201) y ictus (n=82) en una unidad de	Hospital Nacional de Neurología y Neurocirugía, Reino Unido, entre 1994 y 1997.	Medida de Independencia Funcional como la anterior.	1. Validez de criterio 2. Prueba de hipótesis 3. Sensibilidad	1. Inadecuado 2. Inadecuado 3. Inadecuado	1. Comparación de las puntuaciones de cambio de BI y FIM: FIM total=21,9 FIM motor= 19.1 FIM cognitivo=2.8

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
	neurorehabilitación entre 1994 y 1997.					2. Comparación de las puntuaciones de cambio de BI y FIM: BI=5.2
FIM turco (Tur-FIM)						
Kuçekdeveci et al. (2001)	51 pacientes con ictus. 62 pacientes con lesión de la médula espinal.	Departamento de Medicina Física y Rehabilitación en la Facultad de Medicina de la Universidad de Ankara, Turquía, de 1994 a 1998.	Los autores realizaron una traducción de la Medida de Independencia Funcional original mencionada con anterioridad a una versión turca con tres traductores independientes que la unificaron en una versión definitiva.	1. Validez estructural 2. Consistencia interna 3. Validez transcultural 4. Fiabilidad	1. Dudoso 2. Muy bueno 3. Inadecuado 4. Inadecuado	1. El análisis de Rasch mostró que los elementos de vejiga e intestino comprometen la unidimensionalidad en la escala motora. 2. Los valores α al ingreso fueron 0,957 para la FIM motora y 0,921 para la FIM cognitiva, y al alta fueron 0,962 para la FIM motora y 0,922 para la FIM cognitiva. 4. El nivel de acuerdo entre los dos evaluadores fue adecuado, reflejado en niveles kappa superiores a 0,44 para el ictus.

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
FIM persa (Per-FIM)						
Naghdi et al. (2016)	Se incluyeron en este estudio 40 pacientes con ictus.	Centro Tabassom de Rehabilitación de ictus en Teherán, Irán.	Los autores realizaron una traducción de la Medida de Independencia Funcional original mencionada con anterioridad a una versión persa utilizando un método de traducción de avance a retroceso y una validación de un panel de expertos.	1. Validez estructural 2. Consistencia interna 3. Validez transcultural 4. Fiabilidad 5. Prueba de hipótesis 6. Sensibilidad	1. Inadecuado 2. Muy bueno 3. Inadecuado 4. Inadecuado 5. Muy bueno 6. Muy bueno	1. Los hallazgos indican la estructura unidimensional de los factores Per-FIM extraídos. 2. Los coeficientes alfa de Cronbach para Per-FIM y subescalas fueron altos, oscilando entre 0,70 y 0,96. 4. Los valores del ICC para la Fiabilidad inter e intraevaluador estuvieron entre 0,88-0,97 y 0,93-0,98, respectivamente, para la puntuación total, la puntuación de la subescala motora y la puntuación de la subescala cognitiva de Per-FIM, lo que indica una excelente fiabilidad. 5. Hubo una correlación excelente significativa entre Per-FIM y Per-BI demostrada

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
						por un análisis de correlación de Pearson de 0,95 ($p < 0,001$).
Motor FIM (13 elementos) (FIM-Mot-13)						
Hachisuka et al. (1997)a	Cuarenta y un pacientes participaron en el estudio comparativo. Ciento dos pacientes se incluyeron en el estudio test-retest.	Se realizó un estudio comparativo en el Hospital Universitario de Salud Ocupacional y Ambiental de Japón. El estudio test-retest se realizó en siete hospitales afiliados en Japón.	Es como el FIM original de 18 ítems, pero sin los 5 ítems de los dominios de comunicación y reconocimiento social; por lo que finalmente tiene 13 ítems del dominio motor.	1. Consistencia interna 2. Validez de criterio 3. Prueba de hipótesis 4. Sensibilidad	1. Muy bueno 2. Muy bueno 3. Muy bueno 4. Muy bueno	1. El alfa de Cronbach del FIM-Mot-13 es 0,901. 2. La correlación de Spearman entre FIM-Mot 13 y FIM-Mot-3lev es 0,909. 3. La correlación de Spearman entre FIM-Mot-13 y SR-BI es 0,904. La correlación de Spearman entre FIM-Mot-13 y G-BI es 0,901. La correlación de Spearman entre FIM-Mot-13 y BI es 0,901.
van der Putten et al. (1999)	Pacientes con esclerosis	Hospital Nacional de Neurología y de la	Subescala motora de 13 ítems de la Medida de	1. Validez de criterio	1. Inadecuado	1. Comparación de las puntuaciones de cambio de BI

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
	múltiple (n=201) y ictus (n=82) en una unidad de neurorrehabilitación entre 1994 y 1997.	Neurocirugía, Reino Unido, entre 1994 y 1997.	Independencia Funcional como la anterior.	2. Prueba de hipótesis 3. Sensibilidad	2. Inadecuado 3. Inadecuado	y FIM: FIM total=21,9 FIM motor= 19.1 FIM cognitivo=2.8 2. Comparación de las puntuaciones de cambio de BI y FIM: BI=5.2
Gosman-Hedstrom et al. (2000)	Doscientos cuatro pacientes con ictus.	Hospital Universitario Sahlgrenska. Suecia.	Subescala motora de 13 ítems de la Medida de Independencia Funcional como la anterior.	1. Prueba de hipótesis 2. Sensibilidad	1. Dudoso 2. Dudoso 3. Dudoso	1. Porcentaje de concordancia (PA) como medida de concordancia entre escalas. El PA en los niveles categóricos entre las escalas calibradas es de 0,94 ($PA = (11 + 32 + 148) / 204 = 0,94$), lo que significa una concordancia del 94 % entre los 3 niveles condensados.
Hsueh et al. (2002)	En el estudio participaron	Nacional Hospital Universitario	Subescala motora de 13 ítems de la Medida de	1. Consistencia interna	1. Muy bueno	1. Alfa de Cronbach = 0,91 2. ICC=0,86

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
	118 pacientes hospitalizados con ictus en una unidad de rehabilitación.	de Taiwán entre el 1 de diciembre de 1999 y el 31 de mayo de 2000.	Independencia Funcional como la anterior.	2. Fiabilidad 3. Prueba de hipótesis 4. Sensibilidad	2. Inadecuado 3. Muy bueno 4. Muy bueno	3. La correlación de Spearman entre FIM-Mot y BI fue de 0,94 La correlación de Spearman entre FIM-Mot y BI-5 fue de 0,92
Báscula FIM-Motor de 3 niveles (FIM-Mot 3-lev)						
Hachisuka et al. (1997)a	Cuarenta y un pacientes participaron en el estudio comparativo. Ciento dos pacientes se incluyeron en el estudio test-retest.	Se realizó un estudio comparativo en el Hospital Universitario de Salud Ocupacional y Ambiental de Japón. El estudio test-retest se realizó en siete hospitales afiliados en Japón.	Esta versión tiene los mismos elementos que FIM-Motor, pero la puntuación está clasificada en una escala de tres niveles de 1 (1 y 2 del original), 4 (3, 4 y 5 del original) o 7 (6 o 7 del original).	1. Consistencia interna 2. Validez de criterio 3. Prueba de hipótesis 4. Sensibilidad	1. Muy bueno 2. Muy bueno 3. Muy bueno 4. Muy bueno	1. El alfa de Cronbach del FIM-Mot-3lev es 0,864. 2. La correlación de Spearman entre FIM-Mot-3lev y FIM-Mot-13 es 0,909. 3. La correlación de Spearman entre FIM-Mot-3lev y BI es 0,850.
FIM-Cognitivo (FIM-						

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Cog)						
van der Putten et al. (1999)	Pacientes con esclerosis múltiple (n=201) y ictus (n=82) en una unidad de neurorrehabilitación entre 1994 y 1997.	Hospital Nacional de Neurología y Neurocirugía, Reino Unido, entre 1994 y 1997.	Esta versión tiene los 5 ítems cognitivos de los dominios de comunicación y reconocimiento social.	1. Validez de criterio 2. Prueba de hipótesis 3. Sensibilidad	1. Inadecuado 2. Inadecuado 3. Inadecuado	1. Comparación de las puntuaciones de cambio de BI y FIM: FIM total=21,9 FIM motor= 19.1 FIM cognitivo=2.8 2. Comparación de las puntuaciones de cambio de BI y FIM: BI=5.2
Adunsky y otros (2002)	Se incluyeron ciento cincuenta y un pacientes en el análisis final.	Centro Médico Sheba, Ramat Gan, Israel.	Subescala cognitiva de la medida de independencia funcional como la anterior.	1. Prueba de hipótesis 2. Sensibilidad	1. Muy bueno 2. Muy bueno	1. Cog-FIM frente a MMSE r=0,57 Cog-FIM frente a CDT r=0,51
FIM Entrevistado (Int-FIM)						

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Brosseau et al. (1995)	Se reclutaron 89 adultos supervivientes de un ictus.	Hospital Notre-Dame (HND) de Montreal, entre mayo y agosto de 1992.	Medida de Independencia Funcional entrevistada por un evaluador, sin observación del desempeño.	1. Fiabilidad 2. Validez de criterio 3. Sensibilidad	1. Inadecuado 2. Muy bueno 3. Muy bueno	1. Se encuentran ICC relativamente altos para la mayoría de las secciones FIM (valores R que varían de 0,56 a 0,81) para la muestra total de sujetos. 2. Los valores de sensibilidad y especificidad son relativamente altos para los métodos de evaluación FIM 1 y 2. En el método FIM 1 (entrevista estructurada con supervivientes de ictus), la sensibilidad (0,79 a 0,94) fue mayor que la especificidad (0,67 a 0,87). En el método FIM 2 (entrevista estructurada con personal de enfermería), la sensibilidad (0,63 a 0,94) fue menor que la especificidad (0,73 a 1,00), excepto para los ítems N-O.

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Daving et al. (2001)	En este estudio se evaluaron 63 sobrevivientes de un ictus.	Hospital Sahlgrenska en Gotemburgo, Suecia	Versión FIM entrevistada como arriba.	1. Fiabilidad	1. Inadecuado	1. Las dos entrevistas también se han estudiado utilizando valores de ICC entre todos los evaluadores y estos oscilaron entre 0,62 y 0,88 para los 13 elementos motores (intestino 0,62 y baño 0,88) y 0,60 y 0,72 para los cinco ítems sociales y cognitivos, con una excepción: el ítem de interacción social tuvo un valor ICC de tan solo 0,44. Las puntuaciones totales del ICC para los ítems motores de todos los evaluadores fueron de 0,92 y para los ítems sociocognitivos de 0,75, respectivamente.

Entrevista FIM de 3 ítems (Int-3item-FIM)

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Sandaria et al. (2001)	82 pacientes con ictus.	Unidad neurológica del Hospital Hartford, Hartford, Connecticut, EE.UU.	Los autores solo utilizaron 3 elementos FIM que se correlacionan con BI (transferencia de cama a silla, locomoción y alimentación).	1. Consistencia interna 2. Prueba de hipótesis 3. Sensibilidad	1. Muy bueno 2. Muy bueno 3. Muy bueno	1. Los valores alfa de Cronbach fueron: FIM de la entrevista previa al ictus (subconjunto) = 0,660, FIM de la entrevista posterior al ictus (subconjunto) = 0,939 y FIM de la observación posterior al ictus (subconjunto) = 0,903. 2. La correlación entre las puntuaciones totales de estos instrumentos antes del ictus fue moderada (0,645). Sin embargo, las correlaciones entre las puntuaciones del BI y la FIM después del ictus fueron excelentes (0,913 a 0,971). La correlación de Spearman entre las puntuaciones de la FIM para la entrevista y la observación del terapeuta indicó una correlación entre

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
AlphaFIM de 6 elementos (Alpha-FIM)						moderada y buena entre estos dos métodos de administración. Estas correlaciones oscilaron entre 0,508 y 0,753 para cada ítem. La correlación entre las puntuaciones totales del subconjunto de la FIM obtenidas mediante la entrevista y la observación fue de 0,751. Todas las correlaciones fueron significativas con un valor de $p < 0,001$.
Stillman et al. (2009)	Se incluyó una muestra de conveniencia de 144 pacientes con	Unidad de ictus de un hospital de agudos de Nueva York, EE. UU. Entre octubre de 2004 y noviembre de	Una versión abreviada de 6 ítems del FIM original de 18 ítems, con una calificación de 7 niveles. Consta de 4 ítems motores y 2 ítems cognitivos.	1. Validez estructural 2. Consistencia interna	1. Muy bueno 2. Muy bueno	1. El modelo fue significativo indicando que los factores se cargaron en un solo componente, lo que representó el 62% de la varianza.

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
	ictus.	2006.				2. Estadística alfa de 0,87, lo que indica que la medida era confiable.
FIM de 5 elementos (FIM-5 elementos)						
Bohannon et al. (2003)	451 pacientes ingresados con un ictus isquémico que se diagnosticó sobre la base de la presentación clínica y procedimientos de neuroimagen confirmatorios.	El mismo hospital docente de atención terciaria aguda entre el 1 de marzo de 2000 y el 30 de junio de 2002.	Una selección de 5 elementos de los 18 elementos del FIM original realizada por un experto clínico para caracterizar la independencia funcional.	1. Consistencia interna	1. Muy bueno	1. La fiabilidad del subconjunto de 5 ítems del FIM quedó demostrada por la alta consistencia interna de los ítems al ingreso (alfa = 0,894) y al alta (alfa = 0,918).
FIM agudo (FIM-Acu)						

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Roberts et al. (2016)	481 pacientes con ictus en cuidados agudos.	Diferentes hospitales en EE. UU. Entre el 1 de enero y el 30 de septiembre de 2013.	Una versión FIM con 6 elementos del FIM original de 18 elementos, con una escala de calificación de 3 niveles.	1. Consistencia interna 2. Fiabilidad	1. Muy bueno 2. Inadecuado	1. Se observó un alfa de Cronbach de 0,94, lo que indica que la medida tiene una alta consistencia interna. 2. Coeficientes de correlación entre elementos (todos superiores a 0,6).
FIM de motor autoinformado (SR-FIM-Mot)						
Vadassery et al. (2019)	Una muestra de análisis final de 33 pacientes con ictus.	Departamento de Medicina de Rehabilitación del Hospital Tan Tock Seng, un hospital de atención terciaria de Singapur, de octubre de 2013 a mayo de 2014.	Un autoinforme de los 13 ítems de la subescala motora del FIM, en lugar de una entrevista y una evaluación del desempeño.	1. Fiabilidad	1. Dudoso	1. ICC de 0,651 (intervalo de confianza del 95% [CI] 0,404–0,811).
Versión española de FIM+FAM (Spa-						

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
FIM+FAM)						
Colomer et al. (2023)	En total se incluyeron 122 participantes en el estudio.	Instituto de Rehabilitación Neurológica (IRENEA) (Valencia, España). Desde mayo de 2022 hasta noviembre de 2022.	Los autores utilizan la versión en español de la Medida de Independencia Funcional + Medida de Evaluación Funcional (FIM+FAM), para determinar las propiedades psicométricas.	1. Consistencia interna 2. Fiabilidad 3. Prueba de hipótesis 4. Sensibilidad	1. Muy bueno 2. Inadecuado 3. Muy bueno 4. Muy bueno	1. Alfa de Cronbach de FIM+FAM (30 ítems originales) = 0,973. Alfa de Cronbach de FIM+FAM (30 ítems originales + 6 ítems EADL) = 0,974. 2. Puntuación total ICC FIM+FAM = 0,998. Subescala del motor ICC FIM+FAM = 0,997. Subescala cognitiva ICC FIM+FAM = 0,994. ICC FIM+FAM EADL = 1.000. 3. La correlación de Spearman entre FIM+FAM y BI fue de 0,937. La correlación de Spearman entre FIM+FAM y la Escala de Calificación de Discapacidad

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación de COSMIN	Valores de medición
						fue -0,905.
						La correlación de Spearman entre FIM+FAM y mRS fue -0,728.
						La correlación de Spearman entre FIM+FAM y la Escala de resultados de Glasgow extendida fue de 0,540.
						La correlación de Spearman entre FIM+FAM y NIHSS fue -0,757.
						La correlación de Spearman entre FIM+FAM y la Escala de Variables de Resultados Clínicos fue de 0,861.
						La correlación de Spearman entre FIM+FAM y la prueba de detección de afasia de Mississippi fue de 0,541.
						La correlación de Spearman entre FIM+FAM y MMSE fue de

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
FIM+FAM japonesa (Jap-FIM+FAM)						0,542.
Miki et al. (2016)	En este estudio participaron 42 pacientes con ictus.	Centro de Ataques Cerebrales, Hospital Memorial Oota y Hospital de Rehabilitación Nishi-Hiroshima (Japón), entre mayo y diciembre de 2013.	Los autores utilizan la versión japonesa de la Medida de Independencia Funcional + Medida de Evaluación Funcional (FIM+FAM), para determinar sus propiedades psicométricas.	1. Consistencia interna 2. Fiabilidad 3. Prueba de hipótesis 4. Sensibilidad	1. Muy bueno 2. Inadecuado 3. Muy bueno 4. Muy bueno	1. Los alfa de Cronbach fueron 0,968 para la escala completa, 0,954 para la subescala motora y 0,949 para la subescala cognitiva. 2. Buena fiabilidad para la escala completa y la subescala motora, y excelente fiabilidad para la subescala cognitiva (los taus de Kendall fueron 0,83, 0,80 y 0,98, respectivamente). 3. La correlación de Spearman entre FIM+FAM y BI fue de 0,84 La correlación de Spearman entre FIM+FAM y NIHSS fue de -0,76 La correlación de Spearman

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Reino Unido	FIM+FAM (UK-FIM+FAM)					entre FIM+FAM y mRS fue de -0,88 La correlación de Spearman entre FIM+FAM y BRS U/E fue de 0,68 La correlación de Spearman entre FIM+FAM y BRS L/E fue de 0,81 La correlación de Spearman entre FIM+FAM y MMSE fue de 0,44 La correlación de Spearman entre FIM+FAM y FAB fue de 0,34 La correlación de Spearman entre FIM+FAM y CBS fue -0,56.

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Configuración	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Nayar et al. (2016)	Se analizaron datos de 1.539 pacientes con ictus.	Datos de 62 unidades de rehabilitación especializadas recopilados a través de la base de datos UKROC 2010-2013.	Los autores realizaron un análisis de las propiedades psicométricas de la versión del Reino Unido de la Functional Independence Measure + Functional Assessment Measure (FIM+FAM), que consta de 30 ítems con una clasificación de puntuación de 7 niveles.	1. Validez estructural 2. Consistencia interna	1. Muy bueno 2. Muy bueno	1. Se realizó un análisis factorial exploratorio (EFA) en la muestra A. Todos los ítems presentaron una carga razonablemente alta en el primer componente principal, con las 30 cargas >0,3. La inspección del diagrama de sedimentación sugirió una solución de tres factores con tres componentes principales con valores propios >1,5. 2. La consistencia interna fue alta para toda la escala, con un alfa de Cronbach de 0,96. Los coeficientes alfa para las subescalas Motora, Psicosocial y de Comunicación fueron de 0,97, 0,93 y 0,88, respectivamente.

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Abreviaturas: (BI) Índice de Barthel; (ADL) Actividades de la vida diaria; (UK) Reino Unido; (Ira-BI) Índice de Barthel iraní; (Ita-BI) Índice de Barthel italiano; (Per-BI) Índice de Barthel persa; (WHO) Organización Mundial de la Salud; (Tel-BI) Índice de Barthel telefónico; (Int-BI) Índice de Barthel entrevistado; (SR-BI) Índice de Barthel de autoevaluación; (USA) Estados Unidos de América; (BI-5) Índice de Barthel de 5 ítems; (BI-8) Índice de Barthel de 8 ítems; (BI-16) Índice de Barthel extendido (16 ítems); (BI-SS) Escalas suplementarias del índice de Barthel; (G-BI) Modificación de Granger del índice de Barthel; (Tel-mBI) Índice de Barthel modificado telefónico; (PMA-mBI) Evaluación de la medida del desempeño con el índice de Barthel modificado; (MBI) Índice de Barthel modificado por Shah; (Chi-MBI) Índice de Barthel modificado chino; (Bra-MBI) Índice de Barthel modificado brasileño; (Ira-MBI) Índice de Barthel modificado iraní; (Tur-MBI) Índice de Barthel modificado turco; (PET-MBI) Herramienta de evaluación del desempeño: índice de Barthel modificado; (mRS) Escala de Rankin modificada; (PV-NIHSS) Versión portuguesa de la Escala de ictus de los Institutos Nacionales de Salud; (UDS) Sistema uniforme de datos; (HND) Hospital Notre-Dame; (AIIMS) Instituto de Ciencias Médicas de toda la India; (AFMC) Facultad de Medicina de las Fuerzas Armadas; (PGI) Instituto de posgrado de ciencias médicas; (R&R) Hospital de investigación y referencia; (FIM) Medida de independencia funcional; (Tur-FIM) Medida de independencia funcional turca; (Per-FIM) Medida de independencia funcional persa; (FIM-Mot-13) Medida de independencia funcional motora (13 ítems); (FIM-Mot-3lev) Medida de Independencia Funcional-Motora-escala de 3 niveles; (FIM-Cog) Medida de Independencia Funcional-Cognitiva; (Int-FIM) Medida de Independencia Funcional Entrevistada; (Int-3ítem-FIM) Medida de Independencia Funcional Entrevistada-3 ítems; (FIM-Obs) Medida de Independencia Funcional por observación del desempeño; (Alpha-FIM) Medida de Independencia Funcional Alfa de 6 ítems; (FIM-5ítems) Medida de Independencia Funcional de 5 ítems; (FIM-5lev) Medida de Independencia Funcional de 5 niveles; (FIM-Acu) Medida de Independencia Funcional Aguda; (SR-FIM-Mot) Medida de Independencia Funcional Motora Auto-Reportada; (FAM) Medida de Evaluación Funcional; (Spa-FIM+FAM) FIM+FAM Español; (Jap-FIM+FAM) FIM+FAM Japonés; (UK-FIM+FAM) FIM+FAM Reino Unido; (IRENEA) Instituto de Rehabilitación Neurológica; (UKROC) Colaboración de Resultados de Rehabilitación del Reino Unido; (K) Kappa; (wK) Kappa ponderada; (RMI) Índice de movilidad de Rivermead; (NEADL) Actividades extendidas de la vida diaria de Nottingham; (FAI) Índice de actividades de Frenchay; (ICC) Coeficiente de correlación intraclass; (BBS) Escala de equilibrio de Berg; (MCID) Diferencia clínica mínima importante; (LS) Escala de Longshi; (EFA) Análisis factorial exploratorio; (CFA) Análisis factorial confirmatorio; (PASS) Escala de evaluación postural; (FMA) Evaluación de Fugl-Meyer; (IS) ictus isquémico; (CI) Intervalo de confianza; (MMSE) Mini examen del estado mental; (GDS-15) Escala de depresión geriátrica de 15 ítems; (KATZ) Índice de Katz; (CDT) Tarea de dibujo de reloj; (EADL) Actividades extendidas de la vida diaria; (CBS) Escala de Catherine Bergego; (BRS U/E) Etapas de recuperación de Brunnstrom de extremidades superiores; (BRS L/E) Etapas de recuperación de Brunnstrom de extremidades inferiores; (FAB) Batería de evaluación frontal.

Fuente: elaboración propia

3.3 Versiones identificadas de la medida de independencia funcional

Se identificaron quince versiones de la Medida de Independencia Funcional entre 18 artículos que fueron diseñados para evaluar la independencia y funcionalidad en actividades de la vida diaria en personas con ictus: Medida de Independencia Funcional (Original) (FIM) (Heinemann et al., 1994; van der Putten et al., 1999), FIM Turca (Tur-FIM) (Kuşukdeveci et al., 2001), FIM Persa (Per-FIM) (Naghdi et al., 2016), FIM-Motora (13 ítems) (FIM-Mot-13) (Hachisuka et al., 1997a; van der Putten et al., 1999; Gosman-Hedstrom et al., 2000; Hsueh et al., 2002), Escala de 3 niveles de FIM-Motora (FIM-Mot 3-lev) (Hachisuka et al., 1997a), FIM-Cognitivo (FIM-Cog) (van der Putten et al., 1999; Adunsky et al., 2002), FIM Entrevistado (Int-FIM) (Brosseau et al., 1995; Daving et al., 2001), FIM Entrevistado de 3 ítems (Int-3item-FIM) (Sandaria et al., 2001), AlphaFIM de 6 ítems (Alpha-FIM) (Stillman et al., 2009), FIM de 5 ítems (FIM-5-items) (Bohannon et al., 2003), FIM Agudo (FIM-Acu) (Roberts et al., 2016), FIM Motor Auto-reportado (SR-FIM-Mot) (Vadassery et al., 2019), Versión en español de FIM+FAM (Spa-FIM+FAM) (Colomer et al., 2023), FIM+FAM japonés (Jap-FIM+FAM) (Miki et al., 2016) y FIM+FAM del Reino Unido (UK-FIM+FAM) (Nayar et al., 2016). Esta información se refleja en la Tabla 2.

3.4 Características de las versiones BI y FIM incluidas

Las características de las versiones de BI y FIM incluidas para medir la independencia y funcionalidad en las AVD en personas con ictus y estudios se presentan en la Tabla 2. En todos los casos el idioma original utilizado fue el inglés, aunque hubo varios de estos cuestionarios en los que se realizó una validación transcultural: Índice de Barthel iraní (Ira-BI) (Aminaloaya et al., 2021), Índice de Barthel italiano (Ita-BI) (Galeoto et al., 2019), Índice de Barthel persa (Oveisgharan et al., 2006; Naghdi et al., 2016), Índice de Barthel modificado chino (Shah) (Chi-MBI) (Leung et al., 2007), Índice de Barthel modificado brasileño (Shah) (Bra-MBI) (Cincura et al., 2009), Índice de Barthel modificado iraní (Shah) (Ira-MBI) (Aminaloaya et al., 2021), Índice de Barthel modificado turco (Tur-MBI) (Kuşukdeveci et al., 2000), FIM turco (Tur-FIM) (Kuşukdeveci et al., 2001), FIM persa (Per-FIM) (Naghdi et al., 2016), versión española de FIM+FAM (Spa-FIM+FAM) (Colomer et al., 2023), japonés FIM+FAM (Jap-FIM+FAM) (Miki et al., 2016) y Reino Unido FIM+FAM (UK-FIM+FAM)

La mayoría de los estudios se realizaron en hospitales y centros de rehabilitación neurológica en algunos lugares de Asia, Europa y América. La ubicación más común desde donde se desarrollaron los estudios fue Taiwán (Hsueh et al., 2001; Hsueh et al., 2002; Hsieh et al., 2007; Yang et al., 2022; Lee et al., 2014; Lee et al., 2017), Reino Unido (Wolfe et al., 1991; van der Putten et al., 1999; Green et al., 2001; Sarker et al., 2012; Nayar et al., 2016), Estados Unidos de América (Sandaria et al., 2011; Shinar et al., 1987; Stillman et al., 2009; Bohannon et al., 2003; Roberts et al., 2016) y Japón (Hachisuka et al., 1997a; otros, 1997b; Ohura et al., 2017; Miki et al., 2016).

En cuanto al tamaño muestral de los estudios, estos variaron desde los 7 usuarios incluidos en Loewen et al. (1988) hasta los 27.699 usuarios en

Heinemann et al. (1994), con una heterogeneidad significativa en el tamaño muestral.

En cuanto al número de ítems, las versiones variaron desde 5 ítems (Park, 2018; Hsueh et al., 2002) hasta 16 ítems (Jansa et al., 2004) en el índice de Barthel; y desde 3 ítems (Sandaria et al., 2001), hasta 18 ítems en la versión original y la mayoría de las versiones FIM.

El tipo de pregunta/respuesta utilizada en cada versión varía desde los 3 niveles de la escala de 3 niveles FIM-Motor (Hachisuka et al., 1997a) hasta los 7 niveles de las versiones originales y la mayoría de las FIM.

3.5 Propiedades psicométricas

Las principales propiedades evaluadas fueron la fiabilidad, la sensibilidad, la prueba de hipótesis y la consistencia interna. Los resultados de la lista de verificación COSMIN de 4 puntos de las versiones de BI y FIM para medir la independencia y la funcionalidad en las actividades de la vida diaria (AVD) en personas con ictus se muestran en la Tabla 3. La Tabla 4 muestra los datos de las versiones de BI y FIM utilizando el método GRADE.

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Tabla 3. Calidad metodológica de los estudios y calidad de los resultados reportados por propiedad de medición, instrumento y estudio

Instrumento	Artículos	Validez de contenido		Validez estructural		Consistencia interna		Validez transcultural/Invariancia de la medición		Error de medición		Fiabilidad		Validez criterio		Prueba de hipótesis		Sensibilidad	
		M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q
Índice de Barthel (original) (BI)	Loewen et al. (1988)											D	?			VG	?	VG	?
	Shah et al. (1989)					VG	+												
	Wolfe et al. (1991)											D	+						
	Hachisuka et al. (1997)a					VG	+							VG	+	VG	?	VG	?
	van der Putten et al. (1999)															I	?	I	?
	Gosman-Hedstrom et al. (2000)															D	?	D	?
	Green et al. (2001)											I	?			D	?	D	?

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Artículos	Validez de contenido		Validez estructural		Consistencia interna		Validez transcultural/Invariancia de la medición		Error de medición		Fiabilidad		Validez criterio		Prueba de hipótesis		Sensibilidad	
		M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q
	Hsueh et al. (2001)					VG	+					D	+			VG	?	VG	?
	Hsueh et al. (2002)					VG	+					I	+	VG	+	VG	?	VG	?
	van Hartingsveldt et al. (2006)			VG	?	VG	+												
	Hsieh et al. (2007)									D	?	I	?						
	Sarker et al. (2012)											D	-			VG	?	VG	?
	Zhou et al. (2021)															VG	?	VG	?
	Yang et al. (2022)									A	?	A	+	I	-			I	-
Índice Barthel iraní (Ira-BI)	Aminalroaya et al. (2021)													VG	+			VG	+

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Artículos	Validez de contenido		Validez estructural		Consistencia interna		Validez transcultural/Invariancia de la medición		Error de medición		Fiabilidad		Validez criterio		Prueba de hipótesis		Sensibilidad	
		M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q
Índice Barthel italiano (Ita-BI)	Galeoto et al. (2019)			D	?	VG	+					I	+						
Índice Barthel Persa (Per-BI)	Oveisgharan et al. (2006)					VG	+	I	?			I	-			VG	?	VG	?
	Naghdi et al. (2016)															VG	+	VG	+
Índice Barthel Telefónico (Tel-BI)	Prasad et al. (2018)					VG	+					I	?	A	?	A	?	A	?
Índice de Barthel Entrevistado (Int-BI)	Sandaria et al. (2001)					VG	+									VG	?	VG	?
Autoevaluación del índice de	Hachisuka et al. (1997)a					VG	+					A	?	VG	+	VG	?	VG	?

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Artículos	Validez de contenido		Validez estructural		Consistencia interna		Validez transcultural/Invariancia de la medición		Error de medición		Fiabilidad		Validez criterio		Prueba de hipótesis		Sensibilidad	
		M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q
Barthel (SR-BI)	Hachisuka et al. (1997)b											D	+			I	?	I	?
Índice de Barthel de 5 ítems (BI-5)	Park (2018)					VG	+			I	?	I	+	VG	+	VG	?	VG	?
	Hsueh et al. (2002)					VG	+					I	+	VG	+	VG	?	VG	?
Índice de Barthel de 8 ítems (BI-8)	van Hartingsveldt et al. (2006)			VG	?	VG	+												
Índice de Barthel ampliado (16 ítems) (BI-16)	Jansa et al. (2004)			I	?	VG	+							VG	+	VG	?	VG	?

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Artículos	Validez de contenido		Validez estructural		Consistencia interna		Validez transcultural/Invariancia de la medición		Error de medición		Fiabilidad		Validez criterio		Prueba de hipótesis		Sensibilidad	
		M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q
Escalas suplementarias del índice de Barthel (escala de dificultad autopercibida escala de capacidad) (BI-SS)	Lee et al. (2014)	D	-	I	?									VG	+			VG	+
	Lee et al. (2017)									I	?	I	+	VG	-			VG	-
Modificación de Granger del índice de Barthel (G-BI)	Hachisuka et al. (1997)a					VG	+							VG	+	VG	?	VG	?

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Artículos	Validez de contenido		Validez estructural		Consistencia interna		Validez transcultural/Invariancia de la medición		Error de medición		Fiabilidad		Validez criterio		Prueba de hipótesis		Sensibilidad	
		M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q
Entrevista telefónica con el Índice de Barthel modificado (Granger) (Tel-mBI)	Shinar et al. (1987)											I	?	VG	+			VG	+
Evaluación de medidas de desempeño con el índice de Barthel modificado (Granger) (PMA-mBI)	Shinar et al. (1987)											I	?	VG	+			VG	+

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Artículos	Validez de contenido		Validez estructural		Consistencia interna		Validez transcultural/Invariancia de la medición		Error de medición		Fiabilidad		Validez criterio		Prueba de hipótesis		Sensibilidad	
		M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q
Índice de Barthel Modificado (Shah) (MBI)	Shah et al. (1989)					VG	+												
	Yang et al. (2021)			VG	?							I	?						
	Yang (2022)									A	?	A	+	I	-			I	-
Índice de Barthel modificado chino (Shah) (Chi-MBI)	Leung et al. (2007)			A	?	VG	+	VG	+			D	+						
Índice de Barthel Modificado Brasileño (Shah) (Bra-	Cincura et al. (2009)							I	+			I	+			VG	?	VG	?

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Artículos	Validez de contenido		Validez estructural		Consistencia interna		Validez transcultural/Invariancia de la medición		Error de medición		Fiabilidad		Validez criterio		Prueba de hipótesis		Sensibilidad	
		M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q
MBI)																			
Índice de Barthel Modificado Iraní (Shah) (Ira-MBI)	Aminalroaya et al. (2021)			A	?	VG	+					I	+	VG	+	VG	?	VG	?
Índice de Barthel modificado turco (Ajedrez) (Tour-UP)	Kuşukdeveci et al. (2000)			I	?	VG	+	I	-			I	-						
Herramienta de evaluación del desempeño: Índice de Barthel Modificado	Ohura et al. (2017)											I	+	VG	+			VG	+

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Artículos	Validez de contenido		Validez estructural		Consistencia interna		Validez transcultural/Invariancia de la medición		Error de medición		Fiabilidad		Validez criterio		Prueba de hipótesis		Sensibilidad	
		M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q
(Shah) (PET-MBI)																			
Medida de Independencia Funcional (Original) (MIF)	Heinemann et al. (1994)			VG	?					I	?								
	van der Putten et al. (1999)													I	?	I	?	I	?
FIM turco (Tur-FIM)	Kuçekdeveci et al. (2001)			D	?	VG	+	I	-			I	-						
FIM persa (Per-FIM)	Naghdi et al. (2016)			I	?	VG	+	I	-			I	+			VG	+	VG	+
Motor FIM (13)	Hachisuka et al. (1997)a					VG	+							VG	+	VG	?	VG	?

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Artículos	Validez de contenido		Validez estructural		Consistencia interna		Validez transcultural/Invariancia de la medición		Error de medición		Fiabilidad		Validez criterio		Prueba de hipótesis		Sensibilidad	
		M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q
elementos) (FIM-Mot-13)	van der Putten et al. (1999)													I	?	I	?	I	?
	Gosman-Hedstrom et al. (2000)															D	?	D	?
	Hsueh et al. (2002)					VG	+					I	+	VG	+	VG	?	VG	?
FIM-Motor de 3 niveles (FIM-Mot 3-lev)	Hachisuka et al. (1997)a					VG	+							VG	+	VG	?	VG	?
FIM-Cognitivo (FIM-Cog)	van der Putten et al. (1999)													I	?	I	?	I	?
	Adunsky et al. (2002)															VG	?	VG	?

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Artículos	Validez de contenido		Validez estructural		Consistencia interna		Validez transcultural/Invariancia de la medición		Error de medición		Fiabilidad		Validez criterio		Prueba de hipótesis		Sensibilidad	
		M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q
FIM Entrevistado (Int-FIM)	Brosseau et al. (1995)											I	?	VG	?			VG	?
	Daving et al. (2001)											I	?						
Entrevista FIM de 3 ítems (Int-3item-FIM)	Sandaria et al. (2001)					VG	+									VG	?	VG	?
AlphaFI M de 6 elementos (Alpha-FIM)	Stillman et al. (2009)			VG	?	VG	+												
5 elementos FIM (FIM-5 elementos)	Bohannon et al. (2003)					VG	+												

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Artículos	Validez de contenido		Validez estructural		Consistencia interna		Validez transcultural/Invariancia de la medición		Error de medición		Fiabilidad		Validez criterio		Prueba de hipótesis		Sensibilidad	
		M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q
FIM agudo (END-Acu)	Roberts et al. (2016)					VG	+					I	-						
FIM de motor autoinformado (SR-FIM-Mot)	Vadassery et al. (2019)											D	-						
Versión española de FIM+FA M (Spa-FIM+FA M)	Colomer et al. (2023)					VG	+					I	+			VG	?	VG	?
FIM+FA M japonesa (Jap-FIM+FA M)	Miki et al. (2016)					VG	+					I	?			VG	?	VG	?

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Artículos	Validez de contenido		Validez estructural		Consistencia interna		Validez transcultural/Invariancia de la medición		Error de medición		Fiabilidad		Validez criterio		Prueba de hipótesis		Sensibilidad	
		M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q
Reino Unido FIM+FA M (UK-FIM+FA M)	Nayar et al. (2016)			VG	-	VG	+												

Abreviaturas: M: Calidad metodológica del estudio calificada como I = Inadecuada, D = Dudosa, A = Adecuada, VG = Muy buena, casillas vacías = no reportada.

P: Calidad de los resultados calificados como + = calificación positiva, ? = calificación indeterminada, - = calificación negativa, casillas vacías = no informado.

Fuente: elaboración propia

Tabla 4. Calidad de la evidencia reportada por instrumento

Instrumento	Validez de contenido	Validez estructural	Consistencia interna	Validez transcultural/invariancia de la medición	Error de medición	Fiabilidad	Validez de criterio	Prueba de hipótesis	Sensibilidad
Índice de Barthel (original) (BI)		H	H		I	VL	VL	VL	VL
Índice Barthel iraní (Ira-BI)							H		H
Índice Barthel italiano (Ita-BI)		VL	M			VL			
Índice Barthel Persa (Per-BI)			H	VL		VL		H	H
Índice Barthel Telefónico (Tel-BI)			H			VL	M	M	M
Índice de Barthel Entrevistado (Int-BI)			M					M	M

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Validez de contenido	Validez estructural	Consistencia interna	Validez transcultural/invariancia de la medición	Error de medición	Fiabilidad	Validez de criterio	Prueba de hipótesis	Sensibilidad
Índice de autoevaluación de Barthel (SR-BI)			H			I	H	VL	VL
Índice de Barthel de 5 ítems (BI-5)			H		VL	I	H	M	M
Índice de Barthel de 8 ítems (BI-8)		H	H						
Índice de Barthel ampliado (16 ítems) (BI-16)		VL	I				I	VL	VL
Escalas suplementarias del índice de Barthel (escala de dificultad autopercebida,	I	VL			VL	VL	M		M

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Validez de contenido	Validez estructural	Consistencia interna	Validez transcultural/invariancia de la medición	Error de medición	Fiabilidad	Validez de criterio	Prueba de hipótesis	Sensibilidad
escala de capacidad) (BI-SS)									
Modificación de Granger del índice de Barthel (G-BI)			H			H	H	H	H
Entrevista telefónica con el Índice de Barthel modificado (Granger) (Tel-mBI)						VL	H		H
Evaluación de medidas de desempeño con el índice de Barthel						VL	M		M

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Validez de contenido	Validez estructural	Consistencia interna	Validez transcultural/invariancia de la medición	Error de medición	Fiabilidad	Validez de criterio	Prueba de hipótesis	Sensibilidad
modificado (Granger) (PMA-mBI)									
Índice de Barthel Modificado (Shah) (MBI)		H	H		I	I	VL		VL
Índice de Barthel modificado chino (Shah) (Chi-MBI)		M	H	H		VL			
Índice Barthel brasileño especializado (Shah) (Bra-MBI)				VL		VL		I	I
Índice de Barthel		M	H			VL	H	I	I

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Validez de contenido	Validez estructural	Consistencia interna	Validez transcultural/invariancia de la medición	Error de medición	Fiabilidad	Validez de criterio	Prueba de hipótesis	Sensibilidad
Modificado Iraní (Shah) (Ira-MBI) Índice de Barthel modificado turco (Ajedrez) (Tour-UP) Herramienta de evaluación del desempeño: Índice de Barthel Modificado (Shah) (PET- MBI) Medida de Independencia Funcional (Original) (MIF)		VL	I	VL		VL	I		I
		H			VL		VL	VL	VL

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Validez de contenido	Validez estructural	Consistencia interna	Validez transcultural/invariancia de la medición	Error de medición	Fiabilidad	Validez de criterio	Prueba de hipótesis	Sensibilidad
FIM turco (Tur-FIM)		VL	I	VL		VL			
FIM persa (Per-FIM)		VL	I	VL		VL		I	I
Motor FIM (13 elementos) (FIM-Mot-13)			H			VL	VL	VL	VL
FIM-Motor de 3 niveles (FIM-Mot 3-lev)			I				I	I	I
FIM-Cognitivo (FIM-Cog)							VL	VL	VL
FIM éntrevistado (Int-FIM)						I	M		M
Entrevista FIM de 3 ítems (Int-3ítem-FIM)			M					M	M
AlphaFIM de 6 elementos		H	H						

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Validez de contenido	Validez estructural	Consistencia interna	Validez transcultural/invariancia de la medición	Error de medición	Fiabilidad	Validez de criterio	Prueba de hipótesis	Sensibilidad
(Alpha-FIM)									
5 elementos FIM (FIM-5 elementos)			H						
FIM agudo (FIM-Acu)			H			VL			
FIM de motor autoinformado (SR-FIM-Mot)						VL			
Versión española de FIM+FAM (Spa-FIM+FAM)			H			VL		I	I
FIM+FAM japonesa (Jap-FIM+FAM)			I			VL		VL	VL

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño Ocupacional desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Validez de contenido	Validez estructural	Consistencia interna	Validez transcultural/invariancia de la medición	Error de medición	Fiabilidad	Validez de criterio	Prueba de hipótesis	Sensibilidad
Reino Unido FIM+FAM (UK-FIM+FAM)		H	H						

Abreviaturas: QE, Calidad de la evidencia calificada como H=Alta, M=Moderada, L=Baja, VL=Muy baja.

Fuente: elaboración propia

3.6 Índice Barthel (Original) (BI)

La versión original del Índice de Barthel fue la más evaluada, presente en 14 estudios (Loewen et al., 1988; Shah et al., 1989; Wolfe et al., 1991; Hachisuka et al., 1997a; van der Putten et al., 1999; Gosman-Hedstrom et al., 2000; Green et al., 2001; Hsueh et al., 2001; Hsueh et al., 2002; van Hartingsveld et al., 2006; Hsieh et al., 2007; Sarker et al., 2012; Zhou et al., 2021; Yang et al., 2022) y prácticamente todas las propiedades psicométricas fueron evaluadas excepto la validez transcultural/invarianza de la medida, debido a que esta propiedad solo se evalúa en las versiones traducidas; y la validez de contenido, la propiedad psicométrica más crucial. En cuanto a la validez estructural, que solo se evalúa en un estudio (van Hartingsveld et al., 2006), la calidad metodológica se calificó como "muy buena", mostrando solo un factor en el análisis factorial. La consistencia interna se evaluó en cinco estudios (Shah et al., 1989; Hachisuka et al., 1997a; Hsueh et al., 2001; Hsueh et al., 2002; van Hartingsveld et al., 2006), obteniendo una puntuación de "muy buena" con un alfa de Cronbach $\geq 0,70$. El error de medición se calificó como "dudoso" (Hsieh et al., 2007) y "adecuado" (Yang et al., 2022). La fiabilidad test-retest se midió en diez estudios, calificados como «inadecuados» (Green et al., 2001; Hsueh et al., 2002; Hsieh et al., 2007), «dudosos» (Loewen et al., 1988; Wolfe et al., 1991; Hsueh et al., 2001; Sarker et al., 2012) y «adecuados» (Yang et al., 2022). La validez de criterio fue «muy buena» en dos estudios (Hachisuka et al., 1997a; Hsueh et al., 2002) e «inadecuados» en un estudio (Yang et al., 2022), con un rango de $r < 0,30$ (Yang et al., 2022) a $r = 0,994$ (Hachisuka et al., 1997a). La prueba de hipótesis fue "inadecuada" (van der Putten et al., 1999), "dudosa" (Godman-Hedstrom et al., 2000; Green et al., 2001) y "muy buena" (Loewen et al., 1988; Hachisuka et al., 1997a; Hsueh et al., 2001; Hsueh et al., 2002; Sarker et al., 2012; Zhou et al., 2021), con un rango de correlación total de $r = 0,80$ (Sarker et al., 2012) y $r = 0,99$ (Loewen et al., 1988).

3.7 Índice Barthel iraní (Ira-BI)

La versión original cuenta con varias traducciones. Una de ellas es la versión iraní, elaborada por Amini et al. (2021), que evalúa dos propiedades psicométricas: la validez de criterio, calificada como «muy buena», con una correlación entre Ira-BI e Ira-MBI de $r = 0,946$; y la sensibilidad.

3.8 Índice Barthel italiano (Ita-BI)

La versión italiana del BI original (Galeoto et al., 2019) evaluó tres propiedades psicométricas: validez estructural, calificada como "dudosa"; consistencia interna, calificada como "muy buena"; y fiabilidad, con una puntuación de "inadecuada".

3.9 Índice Barthel Persa (Per-BI)

Esta versión persa se evaluó en dos estudios (Oveisgharan et al., 2006; Naghdi et al., 2016), que midieron cinco propiedades psicométricas. La primera fue la consistencia interna, calificada como "muy buena" con un alfa de Cronbach de 0,9354; la validez transcultural, calificada como "inadecuada"; la fiabilidad, como

"inadecuada"; y la prueba de hipótesis y la sensibilidad, calificadas como "muy buenas", con una correlación de $r = -0,912$ con la MRS y de $r = 0,95$ con la Per-FIM.

3.10 Índice Barthel Telefónico (Tel-BI)

La evaluación telefónica con el Índice de Barthel se evaluó en esta versión telefónica del Índice de Barthel (Prasad et al., 2018), evaluando cinco propiedades psicométricas: consistencia interna, calificada como "muy buena", con un alfa de Cronbach que oscila entre 0,8914 a los tres meses y 0,9464 a los seis meses; fiabilidad test-retest puntuada como "inadecuada"; y la correlación de validez de criterio, calificada como "adecuada". La prueba de hipótesis y la sensibilidad también se calificaron como "adecuadas", con una correlación alta y excelente con la NIHSS y la mRS.

3.11 Índice de Barthel Entrevistado (Int-BI)

El Índice de Barthel Entrevistado se incluyó en un único estudio (Sandaria et al., 2001), que evaluó tres propiedades calificadas como "muy buenas". La consistencia interna mostró un alfa de Cronbach de entre 0,828 antes del ictus y 0,952 después. Las pruebas de hipótesis y la sensibilidad mostraron una excelente correlación de Spearman entre el Int-BI y el FIM ($r = 0,913-0,971$).

3.12 Índice de Barthel de autoevaluación (SR-BI)

Este instrumento se incluyó en dos estudios (Hachisuka et al., 1997a; Hachisuka et al., 1997b) que evaluaron cinco propiedades psicométricas. La consistencia interna se calificó como "muy buena", con un alfa de Cronbach de 0,842. La fiabilidad se calificó como "adecuada" y "dudosa", evaluada con kappa (Hachisuka et al., 1997a) y con ICC (Hachisuka et al., 1997b). El coeficiente de correlación entre SR-BI y BI fue $r = 0,994$. La prueba de hipótesis y la sensibilidad se calificaron como "inadecuadas" y "muy buenas", con rangos de $r = 0,454$ (Hachisuka et al., 1997b) a $r = 0,904$ (Hachisuka et al., 1997a).

3.13 Índice de Barthel de 5 ítems (BI-5)

El índice de Barthel de 5 ítems se incluyó en dos estudios (Park, 2018; Hsueh et al., 2002), donde se evaluaron todas las propiedades psicométricas, excepto la validez de contenido, la validez estructural y la validez transcultural. La consistencia interna fue muy buena, con un alfa de Cronbach de 0,73 (Hsueh et al., 2002) a 0,87 (Park, 2018). La fiabilidad se calificó en ambos estudios como inadecuada, con un ICC entre 0,74 (Hsueh et al., 2002) y 0,91 (Park, 2018). El error de medición también fue inadecuado. El coeficiente de correlación fue $r = 0,94-0,95$, con una validez de criterio muy buena. La prueba de hipótesis fue "muy buena" para ambos estudios, con una correlación de $r=0,76$ (Park, 2018) a 0,92 (Hsueh et al., 2002).

3.14 Índice de Barthel de 8 ítems (BI-8)

El índice de Barthel de 8 ítems solo aparece en un estudio (van Hartingsveld et al., 2006) que mide la validez estructural, calificado como "muy bueno", con una escala unidimensional. La consistencia interna también se calificó como "muy buena", con un alfa de Cronbach de 0,93.

3.15 Índice de Barthel ampliado (16 ítems) (BI-16)

Jansa et al. (2004) estudiaron una versión extendida del Índice de Barthel de 16 ítems, evaluando cinco propiedades psicométricas. La validez estructural se calificó como "inadecuada" y mostró una escala de tres factores. Las cuatro propiedades restantes obtuvieron una puntuación "muy buena". La consistencia interna aumentó de 0,89 en la semana 1 a 0,94 en la semana 6. El BI-16 correlaciona con el BI con $r = 0,89$ en la semana 1 y $r = 0,98$ en la semana 6 (validez de criterio). Además, el BI-16 correlaciona con el FMA para la prueba de hipótesis con $r = 0,66$ en la semana 1 a $r = 0,87$ en la semana 6.

3.16 Escalas suplementarias del índice de Barthel (Escala de dificultad autopercebida, Escala de capacidad) (BI-SS)

Dos estudios evaluaron la BI-SS (Lee et al., 2014; Lee et al., 2017). Este es el único estudio en el que se evaluó la validez de contenido y se calificó como "dudosa". La validez estructural se calificó como "inadecuada" y mostró unidimensionalidad en ambas escalas. La fiabilidad y el error de medición también se calificaron como "inadecuados". El ICC fue de 0,97 en la Escala de Capacidad y de 0,78 en la Escala de Dificultad Autopercebida. La validez de criterio fue "muy buena" en ambos estudios, con una correlación entre $r = 0,23$ (Lee et al., 2017) y $r = 0,90$ (Lee et al., 2014).

3.17 Modificación de Granger del índice de Barthel (G-BI)

El G-BI se evaluó sólo en un estudio (Hachisuka et al., 1997a). La consistencia interna fue muy buena, con un alfa de Cronbach de 0,791. La validez de criterio obtuvo una puntuación muy buena, con una correlación entre el G-BI y el BI de $r = 0,993$. Las pruebas de hipótesis también fueron muy buenas, con una correlación de Spearman entre el G-BI y el FIM-Mot-13 de 0,901.

3.18 Entrevista telefónica con índice de Barthel modificado (Granger) (Tel-mBI)

Tel-mBI se incluyó en un único estudio (Shinar et al., 1987), en el que se evaluaron tres propiedades psicométricas: fiabilidad, calificada como "inadecuada"; y validez de criterio y sensibilidad calificadas como "muy buenas", con una correlación entre Tel-mBI y PMA-mBI de $r > 0,97$.

3.19 Evaluación de medidas de desempeño con el índice de Barthel modificado (Granger) (PMA-mBI)

Al igual que en la versión anterior, PMA-mBI fue incluido en un único estudio de Shinar et al. (1987), evaluando las mismas propiedades: fiabilidad, calificada

como "inadecuada" y validez de criterio y sensibilidad calificada como "muy buena" con una correlación entre Tel-mBI y PMA-mBI de $r > 0,97$.

3.20 Índice de Barthel Modificado (Shah) (MBI)

La modificación de Shah del Índice de Barthel es una de las versiones más evaluadas y traducidas, después de la versión original. Se evaluaron prácticamente todas las propiedades psicométricas, excepto la validez de contenido, la validez transcultural y la prueba de hipótesis. En cuanto a la validez estructural, la calidad metodológica se calificó como "muy buena" (Yang et al., 2021). La consistencia interna fue "muy buena", con un alfa de Cronbach de 0,900 al inicio y 0,927 al alta (Shah et al., 1989). La fiabilidad se calificó como "inadecuada" (Yang et al., 2021) y "adecuada" (Yang et al., 2022), con un ICC de 0,94. El error de medición se calificó como "adecuado". La validez de criterio y la sensibilidad se calificaron como "inadecuadas".

3.21 Índice de Barthel modificado chino (Shah) (Chi-MBI)

Chi-MBI es una traducción al chino del BI modificado de Shah. El estudio de Leung et al. (2007) evaluó cuatro propiedades psicométricas: validez estructural, calificada como "adecuada", mostrando una solución de dos factores; consistencia interna calificada como "muy buena" con un alfa de Cronbach de 0,92-0,93; validez transcultural también calificada como "muy buena"; y fiabilidad calificada como "dudosa" con un valor kappa entre 0,81 y 1,00.

3.22 Índice de Barthel Modificado Brasileño (Shah) (Bra-MBI)

Bra-MBI (Cincura et al., 2009) midió cuatro propiedades psicométricas: validez transcultural, calificada como "inadecuada". La fiabilidad test-retest se calificó como "inadecuada", con un valor de ICC de 0,967. La prueba de hipótesis y la sensibilidad se calificaron como "muy buenas", con una correlación con PV-NIHSS de $r = -0,438$.

3.23 Índice de Barthel Modificado Iraní (Shah) (Ira-MBI)

La versión iraní del MBI evaluó casi todas las propiedades, excepto la validez de contenido, la validez transcultural y el error de medición. Los autores calificaron la validez estructural como "adecuada", encontrando solo un factor. El alfa de Cronbach fue de 0,955, mostrando una consistencia interna "muy buena". El coeficiente de correlación intraclase fue de 0,955, con una puntuación de fiabilidad "inadecuada". El coeficiente de correlación entre Ira-MBI e Ira-BI fue $r = 0,946$, mostrando una validez de criterio "muy buena". Las pruebas de hipótesis y la sensibilidad se calificaron como "muy buenas", con una correlación entre $r = 0,359$ y $r = 0,93$.

3.24 Índice de Barthel modificado turco (Shah) (Tur-MBI)

El Tur-MBI (Kuçukdeveci et al., 2000) evaluó cuatro de las nueve propiedades psicométricas del COSMIN: la validez estructural se calificó como "inadecuada",

con dudas sobre su unidimensionalidad. La consistencia interna fue "muy buena", con un alfa de Cronbach entre 0,927 y 0,930. La validez y la fiabilidad transculturales fueron "inadecuadas", con un índice kappa superior a 0,6.

3.25 Herramienta de evaluación del desempeño: Índice de Barthel Modificado (Shah) (PET-MBI)

Esta versión solo fue evaluada en un estudio (Ohura et al., 2017), midiendo únicamente la fiabilidad, con una puntuación "inadecuada" y un ICC de 0,99-1,00; validez de criterio y sensibilidad, calificadas como "muy buenas" con correlaciones de Pearson y Spearman de $r=0,90$ a $r=0,95$.

3.26 Medida de Independencia Funcional (Original) (MIF)

La versión original de la Medida de Independencia Funcional es una de las más utilizadas por los terapeutas para evaluar la independencia y la funcionalidad. Esta versión se evaluó en dos estudios (Heinemann et al., 1994; van der Putten et al., 1999), evaluando cinco propiedades. La validez estructural se calificó como "muy buena". Las otras cuatro propiedades se calificaron como "inadecuadas": error de medición, validez de criterio (comparación entre la FIM total, la FIM motora y la FIM cognitiva), prueba de hipótesis (comparación con la BI) y sensibilidad.

3.27 FIM turco (Tur-FIM)

El Tur-FIM (Kuşukdeveci et al., 2001) evalúa cuatro propiedades psicométricas: validez estructural, calificada como "dudosa", lo que indica que hay dos ítems que comprometen la unidimensionalidad; consistencia interna calificada como "muy buena", con un alfa de Cronbach entre 0,921 y 0,962. La validez transcultural se calificó como "inadecuada", al igual que la fiabilidad, con un nivel kappa superior a 0,44.

3.28 FIM persa (Per-FIM)

La traducción persa de FIM (Naghdi et al., 2016) presenta más propiedades psicométricas analizadas (seis de nueve propiedades). La validez estructural obtuvo una puntuación "insuficiente", lo que indica la estructura unidimensional de la escala. La consistencia interna fue "muy buena", con un alfa de Cronbach de 0,70 a 0,96. La validez y la fiabilidad transculturales se calificaron como "insuficientes", con un ICC interevaluador de 0,88-0,97 y un ICC intraevaluador de 0,93-0,98. La prueba de hipótesis y la sensibilidad se calificaron como "muy buenas", con una correlación de $r = 0,95$ entre Per-FIM y Per-BI.

3.29 Motor FIM (13 elementos) (FIM-Mot-13)

FIM-Motor (13 ítems) es la versión de FIM incluida en más estudios, y se ha analizado en cuatro estudios (Hachisuka et al., 1997a; van der Putten et al., 1999; Gosman-Hedstrom et al., 2000; Hsueh et al., 2002). La consistencia interna se calificó dos veces como «muy buena», con un alfa de Cronbach de

0,901 (Hachisuka et al., 1997a) a 0,91 (Hsueh et al., 2002). La fiabilidad se calificó como «inadecuada» (Hsueh et al., 2002). La validez de criterio fue "muy buena" para dos estudios (Hachisuka et al., 1997a; van der Putten et al., 1999), con una correlación entre FIM-Mot-13 y FIM-Mot-3lev de $r = 0,904$. Las pruebas de hipótesis se calificaron como "inadecuadas" (van der Putten et al., 1999), "dudosas" (Gosman-Hedstrom et al., 2000) y "muy buenas" (Hachisuka et al., 1997a; Hsueh et al., 2002), con una correlación entre $r = 0,901$ (Hachisuka et al., 1997a) y $r = 0,94$ (Hsueh et al., 2002).

3.30 Escala FIM-Motor de 3 niveles (FIM-Mot 3-lev)

La escala FIM Motor, similar a la anterior, pero con una puntuación de tres niveles, se evaluó en un único estudio (Hachisuka et al., 1997a) que evaluó cuatro propiedades psicométricas. La consistencia interna se calificó como "muy buena", con un alfa de Cronbach de 0,864. El coeficiente de correlación entre FIM-Mot-3lev y FIM-Mot-13 fue $r = 0,909$, con una validez de criterio "muy buena". La prueba de hipótesis y la sensibilidad se calificaron como "muy buenas", con una correlación entre FIM-Mot-3lev y BI de $r = 0,850$.

3.31 FIM-Cognitivo (FIM-Cog)

La otra parte de la FIM, además de la motora, es la cognitiva, lo que da lugar a esta versión (FIM-Cog), evaluada en dos estudios (van der Putten et al., 1999; Adunsky et al., 2002). La fiabilidad se calificó como "inadecuada". La validez de criterio también fue "inadecuada", con una comparación entre el cambio de la FIM en las puntuaciones motoras, cognitivas y totales (van der Putten et al., 1999). La prueba de hipótesis se calificó como "inadecuada" (van der Putten et al., 1999) y como "muy buena" (Adunsky et al., 2002), con una correlación que oscila entre $r = 0,51$ y $r = 0,57$ (Adunsky et al., 2002).

3.32 FIM Entrevistado (Int-FIM)

La entrevista FIM se analizó dos veces (Brosseau et al., 1995; Daving et al., 2001) y se evaluó su fiabilidad, calificada como "inadecuada" en ambos estudios, con un ICC de entre 0,44 (Daving et al., 2001) y 0,92 (Daving et al., 2001). La validez de criterio y la sensibilidad se calificaron como "muy buenas", con una sensibilidad de entre 0,79 y 0,94 y una especificidad de entre 0,67 y 1,00 (Brosseau et al., 1995).

3.33 Entrevista FIM de 3 ítems (Int-3item-FIM)

La versión de entrevista de 3 ítems del FIM se evaluó en un solo estudio (Sandaria et al., 2001), con tres propiedades psicométricas evaluadas: consistencia interna con un alfa de Cronbach "muy bueno" de 0,660 a 0,939. La prueba de hipótesis y la sensibilidad también fueron "muy buenas", con una correlación de Spearman de $r = 0,751$.

3.34 AlphaFIM de 6 ítems (Alpha-FIM)

Esta versión abreviada de 6 ítems se denominó Alpha-FIM y se analizó en un estudio (Stillman et al., 2009). La validez estructural se calificó como "muy buena" al mostrar un solo factor. La consistencia interna se calificó como "muy buena", con un alfa de Cronbach de 0,87.

3.35 5 artículos FIM (FIM-5 artículos)

En un estudio (Bohannon et al., 2003) se evaluó una versión abreviada con 5 de los 18 ítems originales del FIM (FIM-5-items), midiendo únicamente la consistencia interna, calificada como "muy buena" con un alfa de Cronbach de 0,894 al ingreso y de 0,918 al alta.

3.36 FIM aguda (FIM aguda)

La FIM aguda se evaluó en un único estudio (Roberts et al., 2016), midiendo dos propiedades: la consistencia interna se calificó como «muy buena» con un alfa de Cronbach de 0,94. La fiabilidad fue «inadecuada», con un coeficiente de correlación interítem superior a 0,6.

3.37 FIM motor autoevaluado (SR-FIM-Mot)

SR-FIM-Mot solo se incluyó en un estudio (Vadassery et al., 2019), midiendo solo una propiedad psicométrica, la fiabilidad test-retest, calificada como "dudosa" con un ICC de 0,651.

3.38 Versión española de FIM+FAM (Spa-FIM+FAM)

La versión en español de FIM+FAM (Colomer et al., 2023) evaluó cuatro propiedades psicométricas. En cuanto a la consistencia interna, el alfa de Cronbach fue de 0,973, lo que indica una puntuación "muy buena". La fiabilidad se calificó como "insuficiente", con un ICC de 0,998. Se observaron algunas correlaciones, lo que indica una prueba de hipótesis y una sensibilidad "muy buenas", con una correlación de Spearman que osciló entre $r = 0,542$ (en comparación con la Prueba de Detección de Afasia de Mississippi) y $r = 0,937$ (en comparación con el Índice de Barthel).

3.39 FIM+FAM japonesa (Jap-FIM+FAM)

La versión japonesa de FIM+FAM (Miki et al., 2016) evaluó cuatro propiedades psicométricas: la consistencia interna fue muy buena, con un alfa de Cronbach de 0,949 a 0,968. La fiabilidad obtuvo una puntuación deficiente, con un κ de Kendall de 0,80 a 0,98. Los coeficientes de correlación oscilaron entre $r = 0,34$ (en comparación con FAB) y $r = 0,84$ (en comparación con BI), lo que indica una prueba de hipótesis y una sensibilidad muy buenas.

3.40 Reino Unido FIM+FAM (UK-FIM+FAM)

Esta versión (UK-FIM+FAM) (Nayar et al., 2016) analizó dos propiedades psicométricas. Para la validez estructural, los autores realizaron un análisis

factorial exploratorio (AFE), que reveló tres componentes principales en la escala, calificados como "muy buenos". Para la consistencia interna, el alfa de Cronbach fue de 0,96, calificando esta propiedad como "muy buena".

4.DISCUSIÓN

Esta es la primera revisión sistemática que identifica las versiones del Índice de Barthel y la Medida de Independencia Funcional para medir la independencia y la funcionalidad en las actividades de la vida diaria en personas con ictus, y evalúa las propiedades psicométricas y la calidad metodológica de las versiones, basándose en la metodología COSMIN más actualizada (Elsman et al., 2024; Mokkink et al., 2010; Terwee et al., 2012). Se identificaron veinte versiones del Índice de Barthel y quince versiones de la Medida de Independencia Funcional diseñadas para evaluar la independencia y la funcionalidad en las AVD en personas con ictus mediante la inclusión de 35 estudios.

Los profesionales clínicos e investigadores deben considerar diversos factores, como el entorno y la población donde se utilizará el instrumento, sus dimensiones, el número de ítems y la evidencia demostrada en la evaluación de cada propiedad psicométrica, para decidir qué instrumento es el más adecuado para medir este constructo. Además, deben tenerse en cuenta el idioma y la cultura de las diferentes versiones, incluyendo la versión original de BI o FIM que se utilizará.

No existía ninguna versión de BI ni de FIM para medir la independencia y la funcionalidad en las AVD en las que se evaluaron todas las propiedades psicométricas. Las propiedades evaluadas en la mayoría de los estudios fueron la sensibilidad, la fiabilidad, la prueba de hipótesis y la consistencia interna, como se muestra en la tabla 3.

4.1 Calidad metodológica

Hubo cinco estudios (Aminalroaya et al., 2021; Hachisuka et al., 1997a; Leung et al., 2007; Prasad et al., 2018; Yang et al., 2022) en los que la calidad metodológica se calificó como "adecuada", para la validez estructural (Aminalroaya et al., 2021; Leung et al., 2007), debido a un análisis factorial desarrollado, un análisis de Rasch y un buen tamaño de muestra; para el error de medición (Yang et al., 2022), debido a las características similares de los participantes en el estudio y un tamaño de muestra suficiente de supervivientes de ictus incluidos; para la fiabilidad (Hachisuka et al., 1997a; Yang et al., 2022), debido a la estabilidad de la muestra incluida, el intervalo de tiempo apropiado y las condiciones de prueba similares; para la validez de criterio (Prasad et al., 2018), debido a una buena correlación entre versiones; y para la prueba de hipótesis (Prasad et al., 2018), debido a una comparación adecuada entre escalas.

Hubo diez estudios (Adunsky et al., 2002; Aminalroaya et al., 2021; Bohannon et al., 2003; Naghdi et al., 2016; Nayar et al., 2016; Sandaria et al., 2001; Shah et al., 1989; Stillman et al., 2009; van Hartingveld et al., 2006; Zhou et al., 2021) donde la puntuación de calidad metodológica se calificó solo como "muy buena" para todas sus propiedades psicométricas analizadas.

A pesar de estos cinco estudios con una calidad metodológica media-alta, la calidad metodológica de la mayoría de los estudios no fue la esperada, resultando “inadecuada” o “dudosa” en muchas versiones y propiedades debido a la falta de evaluación de contenido, y en el caso de la validez de contenido, debido a la relevancia del ítem en la población de estudio, como se muestra en la tabla 3.

En el caso de la validez estructural, la baja calidad metodológica se debe a que no se calculó para cada dimensión por separado o a la falta de descripción del porcentaje de gestión de los valores faltantes. Para la validez transcultural, la principal razón de la baja calidad es la pequeña muestra y la falta de comparación entre los diferentes grupos después de realizar el análisis transcultural. En el caso de la fiabilidad y la validez de criterio, la baja calidad metodológica se debe a la falta de descripción de los porcentajes y la gestión de los valores faltantes y a la fase posterior al ictus, ya que la mayoría de los estudios se realizaron en personas en fase aguda post-ictus que no estaban neurológicamente estables en el momento de la evaluación. Finalmente, la baja calidad metodológica en la prueba de hipótesis resultó de que las hipótesis debían formularse a priori, aunque se podía deducir lo que se esperaba encontrar en ese estudio.

Ira-BI (Aminaloaya et al., 2021), Ita-BI (Galeoto et al., 2019), Tel-BI (Prasad et al., 2018), Int-BI (Sandaria et al., 2001), BI-8 (van Hartingsveld et al., 2006), BI-16 (Jansa et al., 2004), G-BI (Hachisuka et al., 1997a), Tel-mBI (Shinar et al., 1987), PMA-mBI (Shinar et al., 1987), Chi-MBI (Leung et al., 2007), Bra-MBI (Cincura et al., 2009), Ira-MBI (Aminaloaya et al., 2021), Tur-MBI (Kuşukdeveci et al., 2000), PET-MBI (Ohura et al., 2017), Tur-FIM (Kuşukdeveci et al., 2001), Per-FIM (Naghdi et al., 2016), FIM-Mot 3-Irv (Hachisuka et al., 1997a), Int-3item-FIM (Sandaria et al., 2001), Alpha-FIM (Stillman et al., 2009), FIM-5 ítems (Bohannon et al., 2003), FIM-Acu (Roberts et al., 2016), SR-FIM-Mot (Vadassery et al., 2019), Spa-FIM+FAM (Colomer et al., 2023), Jap-FIM+FAM (Miki et al., 2016) y UK-FIM+FAM (Nayar et al., 2016) solo se evaluaron en un estudio. Por lo tanto, se requieren estudios adicionales para evaluar las propiedades psicométricas de estas versiones.

Las propiedades psicométricas calificadas con mayor frecuencia como “muy buenas” fueron la consistencia interna (31 veces) y la sensibilidad (31 veces), seguidas de la prueba de hipótesis (24 veces) y la validez de criterio (18 veces). Las propiedades calificadas con mayor frecuencia como “inadecuadas” fueron principalmente la fiabilidad (28 veces), seguida de la sensibilidad (6 veces), la prueba de hipótesis y la validez de criterio (5 veces cada una), como se muestra en la tabla 3.

4.2 Calidad de la evidencia

En la tabla 4, la calidad de la evidencia fue «baja» o «muy baja» para la mayoría de las propiedades de medición analizadas en todas las versiones incluidas en

esta revisión sistemática. La validez de contenido, la validez estructural y la consistencia interna se consideran las propiedades de medición más importantes (Mokkink et al., 2010; Prinsen et al., 2018).

4.2.1 Validez del contenido

La validez de contenido solo se presentó en el BI-SS (Lee et al., 2014), y se calificó como "dudosa" debido a la probabilidad de que la propiedad de medición real se aleje de la estimada, aun cuando existe la posibilidad de que sea sustancialmente diferente. Además, la calidad de este resultado se calificó como "-", lo que indica que no se han evaluado la relevancia, la comprensibilidad ni la exhaustividad de los ítems en las diferentes versiones.

4.2.2 Validez estructural

La validez estructural se evaluó en trece estudios para catorce versiones: BI (2006), BI-8, Ita-BI (Galeoto et al., 2019), BI-16 (Jansa et al., 2004), BI-SS (Lee et al., 2014), MBI (Yang et al., 2021), Chi-MBI (Leung et al., 2007), Ira-MBI (Aminalroaya et al., 2021), Tur-MBI (Kuşukdeveci et al., 2000), FIM (Heinemann et al., 1994), Tur-FIM (Kuşukdeveci et al., 2001), Per-FIM (Naghdi et al., 2001). al., 2016), Alpha-FIM (Stillman et al., 2009) y UK-FIM+FAM (Nayar et al., 2016). Sólo seis versiones (BI, BI-8, MBI, FIM, Alpha-FIM y UK-FIM+FAM) obtuvieron una calidad de evidencia calificada como "muy buena". En dos estudios (Chi-MBI, Ira-MBI) se calificó como "adecuado". Para dos estudios (Ita-BI, Tur-FIM) se calificó como "dudoso"; y cuatro estudios (BI-16, BI-SS, Tur-MBI, Per-FIM) se calificaron como "inadecuados". Para todos ellos, excepto UK-FIM+FAM, la calidad de los resultados se calificó como "?" porque se proporciona poca información sobre el procedimiento de evaluación de esta propiedad.

4.2.3 Consistencia interna

Finalmente, para lograr coherencia interna, treinta y un estudios evaluaron esta propiedad psicométrica para veinticinco versiones: BI (Shah et al., 1989; Hachisuka et al., 1997a; al., 2019), Per-BI (Oveisgharan et al., 2006), Tel-BI (Prasad et al., 2018), Int-BI (Sandaria et al., 2001), SR-BI (Hachisuka et al., 1997a), BI-5 (Park, 2018; Hsueh et al., 2002), BI-8 (van Hartingsveld et al., 2006), BI-16 (Jansa et al., 2004), G-BI (Hachisuka et al., 1997a), MBI (Shah et al., 1989), Chi-MBI (Leung et al., 2007), Ira-MBI (Aminalroaya et al., 2021), Tur-MBI (Kuşukdeveci et al., 2021). al., 2000), Tur-FIM (Kuşukdeveci et al., 2001), Per-FIM (Naghdi et al., 2016), FIM-Mot-13 (Hachisuka et al., 1997a), Int-3item-FIM (Sandaria et al., 2001), Alpha-FIM (Stillman et al., 2009), Artículos FIM-5 (Bohannon et al., 2003), FIM-Acu (Roberts et al., 2016), Spa-FIM+FAM (Colomer et al., 2023), Jap-FIM+FAM (Miki et al., 2016) y UK-FIM+FAM (Nayar et al., 2016). Todos ellos obtuvieron una calidad de evidencia calificada como "muy buena" y mostraron una calificación de criterio "+", indicando que el alfa de Cronbach es $\geq 0,70$.

4.3 Conclusiones finales

El índice de Barthel original (IB) (Green et al., 2001; Gosman-Hedstrom et al., 2000; Hachisuka et al., 1997a; Hsieh et al., 2007; Hsueh et al., 2001; Hsueh et al., 2002; Loewen et al., 1988; Shah et al., 1989; Sarker et al., 2012; van der Putten et al., 1999; van Hartingsveld et al., 2006; Wolfe et al., 1991; Yang et al., 2022; Zhou et al., 2021) ha sido la versión del índice de Barthel incluida en más estudios (14 en total, y tres traducciones). La segunda versión fue FIM Motor de 13 ítems (FIM-Mot-13) (Hachisuka et al., 1997a; van der Putten et al., 1999; Gosman-Hedstrom et al., 2000; Hsueh et al., 2002), apareciendo en 4 estudios. La versión con más traducciones es el índice de Barthel modificado por Shah (MBI) con cuatro traducciones (Chi-MBI (Leung et al., 2007), Bra-MBI (Cincura et al., 2009), Ira-MBI (Aminalroaya et al., 2021), Tur-MBI (Kuşukdeveci et al., 2000)), seguido del índice de Barthel (BI) con tres traducciones en 4 estudios; (Ira-BI (Aminalroaya et al., 2021), Ita-BI (Galeoto et al., 2019), Per-BI (Oveisgharan et al., 2006; Naghdi et al., 2016)). La versión del índice de Barthel o de la medida de independencia funcional con más propiedades psicométricas evaluadas ha sido la versión original del índice de Barthel (Green et al., 2001; Gosman-Hedstrom et al., 2000; Hachisuka et al., 1997a; Hsieh et al., 2007; Hsueh et al., 2001; Hsueh et al., 2002; Loewen et al., 1988; Shah et al., 1989; Sarker et al., 2012; van der Putten et al., 1999; van Hartingsveld et al., 2006; Wolfe et al., 1991; Yang et al., 2022; Zhou et al., 2021) en los cuales se han evaluado siete propiedades psicométricas, casi todas ellas, excepto validez de contenido y validez transcultural, que se evalúa en las traducciones de esta versión. Hubo seis versiones con seis propiedades psicométricas evaluadas, casi todas fueron evaluadas, excepto la validez de contenido y la validez transcultural (BI-5 (Hsueh et al., 2002; Park, 2018), BI-SS (Lee et al., 2014; Lee et al., 2017), MBI (Shah et al., 1989; Yang et al., 2021; Yang et al., 2022), Ira-MBI (Aminalroaya et al., 2021), FIM (Heinemann et al., 1994; van der Putten et al., 1999) y Per-FIM (Naghdi et al., 2016)).

Esto convierte a esta versión del Índice de Barthel (Índice de Barthel Original) en la más validada y recomendada para evaluar la independencia y la funcionalidad en las actividades de la vida diaria en personas con ictus, tanto en la práctica clínica como en la investigación. Esta versión (Índice de Barthel Original) se destaca como la más válida y recomendada porque, según los resultados de la presente investigación, es la más utilizada, con más propiedades psicométricas analizadas y ha demostrado mejores puntuaciones. Esto convierte a esta versión del Índice de Barthel en la más validada y recomendada, así como en el estándar de oro para evaluar la independencia y la funcionalidad en las actividades de la vida diaria en personas con ictus.

4.4 Fortalezas y limitaciones

Este estudio constituye la revisión sistemática más reciente que identifica, evalúa y resume la evidencia sobre las versiones del Índice de Barthel y la Medida de Independencia Funcional para medir la independencia y la funcionalidad en las

actividades de la vida diaria en personas con ictus. La revisión se realizó según los estándares COSMIN y PRISMA más actualizados; por lo tanto, se empleó un método adecuado, siguiendo las recomendaciones de expertos (Elsman et al., 2024; Moher et al., 2009; Mokkink et al., 2010; Prinsen et al., 2018).

Aunque se realizó una búsqueda exhaustiva, es posible que no se hayan identificado todas las versiones disponibles de BI y FIM, debido a que solo se seleccionaron artículos en inglés y español, por lo que es posible que no se hayan identificado artículos desarrollados y utilizados en otros idiomas.

5. CONCLUSIONES

Se han desarrollado diferentes versiones de BI y FIM y traducciones transculturales para evaluar la independencia y la funcionalidad en las actividades de la vida diaria en personas con ictus.

Este es uno de los constructos principales que se evalúan en la rehabilitación actual, por lo que es necesario contar con herramientas válidas, sensibles y precisas que permitan a los clínicos e investigadores comprender este constructo en detalle y los cambios que se producen con el tiempo. Realizar una evaluación exhaustiva permite a los profesionales comprender mejor el perfil, las dificultades y las fortalezas del paciente y, por lo tanto, tomar decisiones adecuadas, establecer objetivos precisos y elaborar un plan de intervención lo más cercano posible a la realidad de la persona. Por ello, se ha realizado esta revisión sistemática para recopilar todas las versiones de BI y FIM, analizar sus propiedades psicométricas, examinar su calidad metodológica y la calidad de la evidencia, con el objetivo de resumir sus ventajas y desventajas, y poder encontrar la versión más recomendada para evaluar este constructo, según la lista de verificación COSMIN más actualizada.

La evaluación de este aspecto resulta beneficiosa en la práctica clínica y la investigación, lo que requiere nuevos conocimientos en esta área. La evaluación de esta variable es crucial, ya que permite identificar y evaluar la intervención enfocada en mejorar la independencia y la funcionalidad en las actividades de la vida diaria en personas con ictus. Con base en la evidencia actual sobre las propiedades de medición de las versiones que evalúan este constructo en personas con ictus, la versión Índice de Barthel Original parece ser la más recomendada según los resultados obtenidos y las directrices COSMIN, pero esto requiere mayor evaluación. Se necesitan más estudios de alta calidad metodológica que evalúen las propiedades de medición de las versiones del Índice de Barthel Original (BI) y de la Medida de Independencia Funcional (FIM) para medir la independencia y la funcionalidad en las actividades de la vida diaria en personas con ictus.

6.DECLARACIÓN DE CONFLICTO DE INTERESES

Los autores han declarado no tener ningún conflicto de intereses.

7. CONTRIBUCIONES DEL AUTOR

Todos los autores han acordado la versión final y cumplen al menos uno de los siguientes criterios (recomendados por el ICMJE): contribuciones sustanciales a la concepción y el diseño, adquisición de datos o análisis e interpretación de datos; redacción del artículo; o revisión crítica del mismo para determinar su contenido intelectual importante.

8.FONDOS

Esta investigación no recibió ninguna subvención específica de agencias de financiación de los sectores público, comercial o sin fines de lucro.

9.REFERENCIAS

- Adunsky, A., Fleikssig, Y., Levenkrohn, S., Arad, M., & Noy, S. (2002). Clock drawing task, mini-mental state examination and cognitive-functional independence measure: Relation to functional outcome of stroke patients. *Archives of Gerontology and Geriatrics*, 35(2), 153–160. [https://doi.org/10.1016/s0167-4943\(02\)00018-3](https://doi.org/10.1016/s0167-4943(02)00018-3)
- Aminalroaya, R., Mirzadeh, F. S., Heidari, K., Alizadeh-Khoei, M., Sharifi, F., Effatpanah, M., Angooti-Oshnari, L., Fadaee, S., Saghebi, H., & Hormozi, S. (2021). The Validation Study of Both the Modified Barthel and Barthel Index, and Their Comparison Based on Rasch Analysis in the Hospitalized Acute Stroke Elderly. *International Journal of Aging & Human Development*, 93(3), 864–880. <https://doi.org/10.1177/0091415020981775>
- Bohannon, R. W., Ahlquist, M., Lee, N., & Maljanian, R. (2003). Functional gains during acute hospitalization for stroke. *Neurorehabilitation and Neural Repair*, 17(3), 192–195. <https://doi.org/10.1177/0888439003256045>
- Brosseau, L., Philippe, P., & Boulanger, Y.-L. (1995). The Functional Independence Measure: Validity of selected assessment methods applied to stroke survivors. *Topics in Geriatric Rehabilitation*, 11(1), 75–86.
- Cincura, C., Pontes-Neto, O. M., Neville, I. S., Mendes, H. F., Menezes, D. F., Mariano, D. C., Pereira, I. F., Teixeira, L. A., Jesus, P. A. P., de Queiroz, D. C. L., Pereira, D. F., Pinto, E., Leite, J. P., Lopes, A. A., & Oliveira-Filho, J. (2009). Validation of the National Institutes of Health Stroke Scale, modified Rankin Scale and Barthel Index in Brazil: The role of cultural adaptation and structured interviewing. *Cerebrovascular Diseases (Basel, Switzerland)*, 27(2), 119–122. <https://doi.org/10.1159/000177918>
- Colomer, C., Llorens, R., Sánchez, C., Ugart, P., Moliner, B., Navarro, M. D., Noé, E., & Ferri, J. (2023). Reliability and validity of the Spanish adaptation of the Functional Independence Measure + Functional Assessment Measure. *European Journal of Physical and Rehabilitation Medicine*, 59(4), 452–457. <https://doi.org/10.23736/S1973-9087.23.07841-3>
- Daffertshofer, M., Mielke, O., Pullwitt, A., Felsenstein, M., & Hennerici, M. (2004). Transient Ischemic Attacks Are More Than “Ministrokes”. *Stroke*, 35(11), 2453–2458. <https://doi.org/10.1161/01.STR.0000144050.90132.8e>
- Daving, Y., Andrén, E., Nordholm, L., & Grimby, G. (2001). Reliability of an interview approach to the Functional Independence Measure. *Clinical Rehabilitation*, 15(3), 301–310. <https://doi.org/10.1191/026921501669986659>

Dawson, J., Engineer, N. D., Cramer, S. C., Wolf, S. L., Ali, R., O'Dell, M. W., Pierce, D., Prudente, C. N., Redgrave, J., Feng, W., Liu, C. Y., Francisco, G. E., Brown, B. L., Dixit, A., Alexander, J., DeMark, L., Krishna, V., Kautz, S. A., Majid, A., ... Kimberley, T. J. (2023). Vagus Nerve Stimulation Paired With Rehabilitation for Upper Limb Motor Impairment and Function After Chronic Ischemic Stroke: Subgroup Analysis of the Randomized, Blinded, Pivotal, VNS-REHAB Device Trial. *Neurorehabilitation and Neural Repair*, 37(6), 367–373. <https://doi.org/10.1177/15459683221129274>

Dobkin, B. (1995). The economic impact of stroke. *Neurology*, 45(2 Suppl 1), S6-9.

Elsman, E. B. M., Mokkink, L. B., Terwee, C. B., Beaton, D., Gagnier, J. J., Tricco, A. C., Baba, A., Butcher, N. J., Smith, M., Hofstetter, C., Lee Aiyegbusi, O., Berardi, A., Farmer, J., Haywood, K. L., Krause, K. R., Markham, S., Mayo-Wilson, E., Mehdipour, A., Ricketts, J., ... Offringa, M. (2024). Guideline for reporting systematic reviews of outcome measurement instruments (OMIs): PRISMA-COSMIN for OMIs 2024. *Journal of Clinical Epidemiology*, 111422. <https://doi.org/10.1016/j.jclinepi.2024.111422>

Evers, S. M. A. A., Struijs, J. N., Ament, A. J. H. A., van Genugten, M. L. L., Jager, J. H. C., & van den Bos, G. A. M. (2004). International comparison of stroke cost studies. *Stroke*, 35(5), 1209–1215. <https://doi.org/10.1161/01.STR.0000125860.48180.48>

Feigin, V. L., Owolabi, M. O., Feigin, V. L., Abd-Allah, F., Akinyemi, R. O., Bhattacharjee, N. V., Brainin, M., Cao, J., Caso, V., Dalton, B., Davis, A., Dempsey, R., Duprey, J., Feng, W., Ford, G. A., Gall, S., Gandhi, D., Good, D. C., Hachinski, V., ... Zagożdżon, P. (2023). Pragmatic solutions to reduce the global burden of stroke: A World Stroke Organization–Lancet Neurology Commission. *The Lancet Neurology*, 22(12), 1160–1206. [https://doi.org/10.1016/S1474-4422\(23\)00277-6](https://doi.org/10.1016/S1474-4422(23)00277-6)

Galeoto, G., Formica, M. C., Mercuri, N. B., Santilli, V., Berardi, A. C., Castiglia, S. F., Mollica, R., & Servadio, A. (2019). Evaluation of the psychometric properties of the Barthel Index in an Italian ischemic stroke population in the acute phase: A cross-sectional study. *Functional Neurology*, 34(1), 29–34.

Gosman-Hedström, G., & Blomstrand, C. (2004). Evaluation of a 5-level functional independence measure in a longitudinal study of elderly stroke survivors. *Disability and Rehabilitation*, 26(7), 410–418. <https://doi.org/10.1080/09638280410001662978>

Gosman-Hedström, G., & Svensson, E. (2000). Parallel reliability of the functional independence measure and the Barthel ADL index. *Disability and Rehabilitation*, 22(16), 702–715. <https://doi.org/10.1080/09638280050191972>

- Green, J., Forster, A., & Young, J. (2001). A test-retest reliability study of the Barthel Index, the Rivermead Mobility Index, the Nottingham Extended Activities of Daily Living Scale and the Frenchay Activities Index in stroke patients. *Disability and Rehabilitation*, 23(15), 670–676. <https://doi.org/10.1080/09638280110045382>
- Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., & Schünemann, H. J. (2008). GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*, 336(7650), 924–926. <https://doi.org/10.1136/bmj.39489.470347.AD>
- Hachisuka, K., Okazaki, T., & Ogata, H. (1997)a. Self-rating Barthel index compatible with the original Barthel index and the Functional Independence Measure motor score. *Journal of UOEH*, 19(2), 107–121. <https://doi.org/10.7888/juoeh.19.107>
- Hachisuka, K., Ogata, H., Ohkuma, H., Tanaka, S., & Dozono, K. (1997)b. Test-retest and inter-method reliability of the self-rating Barthel Index. *Clinical Rehabilitation*, 11(1), 28–35. <https://doi.org/10.1177/026921559701100105>
- Heinemann, A. W., Michael Linacre, J., Wright, B. D., Hamilton, B. B., & Granger, C. (1994). Measurement characteristics of the Functional Independence Measure. *Topics in Stroke Rehabilitation*, 1(3), 1–15. <https://doi.org/10.1080/10749357.1994.11754030>
- Hsieh, Y.-W., Wang, C.-H., Wu, S.-C., Chen, P.-C., Sheu, C.-F., & Hsieh, C.-L. (2007). Establishing the minimal clinically important difference of the Barthel Index in stroke patients. *Neurorehabilitation and Neural Repair*, 21(3), 233–238. <https://doi.org/10.1177/1545968306294729>
- Hsueh, I. P., Lee, M. M., & Hsieh, C. L. (2001). Psychometric characteristics of the Barthel activities of daily living index in stroke patients. *Journal of the Formosan Medical Association = Taiwan Yi Zhi*, 100(8), 526–532.
- Hsueh, I.-P., Lin, J.-H., Jeng, J.-S., & Hsieh, C.-L. (2002). Comparison of the psychometric characteristics of the functional independence measure, 5 item Barthel index, and 10 item Barthel index in patients with stroke. *Journal of Neurology, Neurosurgery, and Psychiatry*, 73(2), 188–190. <https://doi.org/10.1136/jnnp.73.2.188>
- Hua, X., Liu, M., & Wu, S. (2023). Definition, prediction, prevention and management of patients with severe ischemic stroke and large infarction. *Chinese Medical Journal*, 136(24), 2912–2922. <https://doi.org/10.1097/CM9.0000000000002885>
- Ingram, L. A., Butler, A. A., Lord, S. R., & Gandevia, S. C. (2023). Use of a physiological profile to document upper limb motor impairment in ageing and

- in neurological conditions. *The Journal of Physiology*, 601(12), 2251–2262. <https://doi.org/10.1113/JP283703>
- Jansa, J., Pogacnik, T., & Gompertz, P. (2004). An evaluation of the Extended Barthel Index with acute ischemic stroke patients. *Neurorehabilitation and Neural Repair*, 18(1), 37–41. <https://doi.org/10.1177/0888439003262287>
- Jhong, M., Tang, N.-Y., Liu, C.-H., Huang, W.-H., Hsu, Y.-T., Liu, Y.-L., Li, T.-C., & Hsieh, C.-L. (2013). Relationship Between Chinese Medicine Pattern Types, Clinical Severity, and Prognosis in Patients With Acute Cerebral Infarct. *EXPLORE*, 9(4), 226–231. <https://doi.org/10.1016/j.explore.2013.04.002>
- Katan, M., & Luft, A. (2018). Global Burden of Stroke. *Seminars in Neurology*, 38(02), 208–211. <https://doi.org/10.1055/s-0038-1649503>
- Küçükdeveci, A. A., Yavuzer, G., Elhan, A. H., Sonel, B., & Tennant, A. (2001). Adaptation of the Functional Independence Measure for use in Turkey. *Clinical Rehabilitation*, 15(3), 311–319. <https://doi.org/10.1191/026921501676877265>
- Küçükdeveci, A. A., Yavuzer, G., Tennant, A., Süldür, N., Sonel, B., & Arasil, T. (2000). Adaptation of the modified Barthel Index for use in physical medicine and rehabilitation in Turkey. *Scandinavian Journal of Rehabilitation Medicine*, 32(2), 87–92.
- Kwakkel, G., Stinear, C., Essers, B., Munoz-Novoa, M., Branscheidt, M., Cabanas-Valdés, R., Lakičević, S., Lampropoulou, S., Luft, A. R., Marque, P., Moore, S. A., Solomon, J. M., Swinnen, E., Turolla, A., Alt Murphy, M., & Verheyden, G. (2023). Motor rehabilitation after stroke: European Stroke Organisation (ESO) consensus-based definition and guiding framework. *European Stroke Journal*, 8(4), 880–894. <https://doi.org/10.1177/23969873231191304>
- Langhammer, B., Sunnerhagen, K. S., Lundgren-Nilsson, Å., Sällström, S., Becker, F., & Stanghelle, J. K. (2017). Factors enhancing activities of daily living after stroke in specialized rehabilitation: An observational multicenter study within the Sunnaas International Network. *European Journal of Physical and Rehabilitation Medicine*, 53(5). <https://doi.org/10.23736/S1973-9087.17.04489-6>
- Lee, Y.-C., Chen, S.-S., Koh, C.-L., Hsueh, I.-P., Yao, K.-P., & Hsieh, C.-L. (2014). Development of two Barthel Index-based Supplementary Scales for patients with stroke. *PloS One*, 9(10), e110494. <https://doi.org/10.1371/journal.pone.0110494>
- Lee, Y.-C., Yu, W.-H., Hsueh, I.-P., Chen, S.-S., & Hsieh, C.-L. (2017). Test-retest reliability and responsiveness of the Barthel Index-based

- Supplementary Scales in patients with stroke. *European Journal of Physical and Rehabilitation Medicine*, 53(5), 710–718. <https://doi.org/10.23736/S1973-9087.17.04454-9>
- Leung, S. O. C., Chan, C. C. H., & Shah, S. (2007). Development of a Chinese version of the Modified Barthel Index—Validity and reliability. *Clinical Rehabilitation*, 21(10), 912–922. <https://doi.org/10.1177/0269215507077286>
- Li, S., Ghuman, J., Gonzalez-Buonomo, J., Huang, X., Malik, A., Yozbatiran, N., Francisco, G. E., Wu, H., & Frontera, W. R. (2023). Does Spasticity Correlate With Motor Impairment in the Upper and Lower Limbs in Ambulatory Chronic Stroke Survivors? *American Journal of Physical Medicine & Rehabilitation*, 102(10), 907–912. <https://doi.org/10.1097/PHM.0000000000002247>
- Loewen, S. C., & Anderson, B. A. (1988). Reliability of the Modified Motor Assessment Scale and the Barthel Index. *Physical Therapy*, 68(7), 1077–1081. <https://doi.org/10.1093/ptj/68.7.1077>
- Miki, E., Yamane, S., Yamaoka, M., Fujii, H., Ueno, H., Kawahara, T., Tanaka, K., Tamashiro, H., Inoue, E., Okamoto, T., & Kuriyama, M. (2016). Validity and reliability of the Japanese version of the FIM+FAM in patients with cerebrovascular accident. *Scandinavian Journal of Occupational Therapy*, 23(5), 398–404. <https://doi.org/10.3109/11038128.2015.1095236>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA Statement. *Open Medicine: A Peer-Reviewed, Independent, Open-Access Journal*, 3(3), e123-130.
- Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & De Vet, H. C. W. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Quality of Life Research*, 19(4), 539–549. <https://doi.org/10.1007/s11136-010-9606-8>
- Naghdi, S., Ansari, N. N., Raji, P., Shamili, A., Amini, M., & Hasson, S. (2016). Cross-cultural validation of the Persian version of the Functional Independence Measure for patients with stroke. *Disability and Rehabilitation*, 38(3), 289–298. <https://doi.org/10.3109/09638288.2015.1036173>
- Nayar, M., Vanderstay, R., Siegert, R. J., & Turner-Stokes, L. (2016). The UK Functional Assessment Measure (UK FIM+FAM): Psychometric Evaluation in Patients Undergoing Specialist Rehabilitation following a Stroke from the National UK Clinical Dataset. *PloS One*, 11(1), e0147288. <https://doi.org/10.1371/journal.pone.0147288>

- Ohura, T., Hase, K., Nakajima, Y., & Nakayama, T. (2017). Validity and reliability of a performance evaluation tool based on the modified Barthel Index for stroke patients. *BMC Medical Research Methodology*, 17(1), 131. <https://doi.org/10.1186/s12874-017-0409-2>
- Oveisgharan, S., Shirani, S., Ghorbani, A., Soltanzade, A., Baghaei, A., Hosseini, S., & Sarrafzadegan, N. (2006). Barthel index in a Middle-East country: Translation, validity and reliability. *Cerebrovascular Diseases (Basel, Switzerland)*, 22(5–6), 350–354. <https://doi.org/10.1159/000094850>
- Park, C.-S. (2018). The test-retest reliability and minimal detectable change of the short-form Barthel Index (5 items) and its associations with chronic stroke-specific impairments. *Journal of Physical Therapy Science*, 30(6), 835–839. <https://doi.org/10.1589/jpts.30.835>
- Prasad, K., Kumar, A., Misra, S., Yadav, A. K., Johri, S., Sarkar, R. S., Gorthi, S. P., Hassan, K. M., Prabhakar, S., Misra, U. K., Kumar, P., & For InveST study group. (2018). Reliability and validity of telephonic Barthel Index: An experience from multi-centric randomized control study. *Acta Neurologica Belgica*, 118(1), 53–59. <https://doi.org/10.1007/s13760-017-0843-2>
- Prinsen, C. A. C., Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., De Vet, H. C. W., & Terwee, C. B. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research*, 27(5), 1147–1157. <https://doi.org/10.1007/s11136-018-1798-3>
- Roberts, P. S., Mix, J., Rupp, K., Younan, C., Mui, W., Riggs, R. V., & Niewczyk, P. (2016). Using Functional Status in the Acute Hospital to Predict Discharge Destination for Stroke Patients. *American Journal of Physical Medicine & Rehabilitation*, 95(6), 416–424. <https://doi.org/10.1097/PHM.0000000000000412>
- Rochmah, T. N., Rahmawati, I. T., Dahlui, M., Budiarto, W., & Bilqis, N. (2021). Economic Burden of Stroke Disease: A Systematic Review. *International Journal of Environmental Research and Public Health*, 18(14), 7552. <https://doi.org/10.3390/ijerph18147552>
- Sadaria, K. S., Bohannon, R. W., Lee, N., & Maljanian, R. (2001). Ratings of physical function obtained by interview are legitimate for patients hospitalized after stroke. *Journal of Stroke and Cerebrovascular Diseases: The Official Journal of National Stroke Association*, 10(2), 79–84. <https://doi.org/10.1053/jscd.2001.24656>
- Sarker, S.-J., Rudd, A. G., Douiri, A., & Wolfe, C. D. A. (2012). Comparison of 2 extended activities of daily living scales with the Barthel Index and predictors of their outcomes: Cohort study within the South London Stroke Register

- (SLSR). *Stroke*, 43(5), 1362–1369. <https://doi.org/10.1161/STROKEAHA.111.645234>
- Shah, S., Vanclay, F., & Cooper, B. (1989). Improving the sensitivity of the Barthel Index for stroke rehabilitation. *Journal of Clinical Epidemiology*, 42(8), 703–709. [https://doi.org/10.1016/0895-4356\(89\)90065-6](https://doi.org/10.1016/0895-4356(89)90065-6)
- Shanks, M. J., & Byblow, W. D. (2024). Corticomotor pathway function and recovery after stroke: A look back and a way forward. *The Journal of Physiology*, JP285562. <https://doi.org/10.1113/JP285562>
- Shinar, D., Gross, C. R., Bronstein, K. S., Licata-Gehr, E. E., Eden, D. T., Cabrera, A. R., Fishman, I. G., Roth, A. A., Barwick, J. A., & Kunitz, S. C. (1987). Reliability of the activities of daily living scale and its use in telephone interview. *Archives of Physical Medicine and Rehabilitation*, 68(10), 723–728.
- Soleimani, M., Ghazisaeedi, M., & Heydari, S. (2024). The efficacy of virtual reality for upper limb rehabilitation in stroke patients: A systematic review and meta-analysis. *BMC Medical Informatics and Decision Making*, 24(1), 135. <https://doi.org/10.1186/s12911-024-02534-y>
- Stillman, G., Granger, C., & Niewczyk, P. (2009). Projecting function of stroke patients in rehabilitation using the AlphaFIM instrument in acute care. *PM & R: The Journal of Injury, Function, and Rehabilitation*, 1(3), 234–239. <https://doi.org/10.1016/j.pmrj.2008.10.014>
- Terwee, C. B., Jansma, E. P., Riphagen, I. I., & De Vet, H. C. W. (2009). Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Quality of Life Research*, 18(8), 1115–1123. <https://doi.org/10.1007/s11136-009-9528-5>
- Terwee, C. B., Mokkink, L. B., Knol, D. L., Ostelo, R. W. J. G., Bouter, L. M., & De Vet, H. C. W. (2012). Rating the methodological quality in systematic reviews of studies on measurement properties: A scoring system for the COSMIN checklist. *Quality of Life Research*, 21(4), 651–657. <https://doi.org/10.1007/s11136-011-9960-1>
- Vadassery, S. J., Kong, K. H., Ho, W. M. L., & Seneviratna, A. (2019). Interview Functional Independence Measure score: Self-reporting as a simpler alternative to multidisciplinary functional assessment. *Singapore Medical Journal*, 60(4), 199–201. <https://doi.org/10.11622/smedj.2018048>
- van der Putten, J. J., Hobart, J. C., Freeman, J. A., & Thompson, A. J. (1999). Measuring change in disability after inpatient rehabilitation: Comparison of the responsiveness of the Barthel index and the Functional Independence Measure. *Journal of Neurology, Neurosurgery, and Psychiatry*, 66(4), 480–484. <https://doi.org/10.1136/jnnp.66.4.480>

- van Hartingsveld, F., Lucas, C., Kwakkel, G., & Lindeboom, R. (2006). Improved interpretation of stroke trial results using empirical Barthel item weights. *Stroke*, 37(1), 162–166. <https://doi.org/10.1161/01.STR.0000195176.50830.b6>
- Wolfe, C. D., Taub, N. A., Woodrow, E. J., & Burney, P. G. (1991). Assessment of scales of disability and handicap for stroke patients. *Stroke*, 22(10), 1242–1244. <https://doi.org/10.1161/01.str.22.10.1242>
- Yang, C. M., Wang, Y.-C., Lee, C.-H., Chen, M.-H., & Hsieh, C.-L. (2022). A comparison of test-retest reliability and random measurement error of the Barthel Index and modified Barthel Index in patients with chronic stroke. *Disability and Rehabilitation*, 44(10), 2099–2103. <https://doi.org/10.1080/09638288.2020.1814429>
- Yang, H., Chen, Y., Wang, J., Wei, H., Chen, Y., & Jin, J. (2021). Activities of daily living measurement after ischemic stroke: Rasch analysis of the modified Barthel Index. *Medicine*, 100(9), e24926. <https://doi.org/10.1097/MD.00000000000024926>
- Zhao, K., He, C., Xiang, W., Zhou, Y., Zhang, Z., Li, J., & Scano, A. (2023). Evidence of synergy coordination patterns of upper-limb motor control in stroke patients with mild and moderate impairment. *Frontiers in Physiology*, 14, 1214995. <https://doi.org/10.3389/fphys.2023.1214995>
- Zhou, M., Liu, X., Zha, F., Liu, F., Zhou, J., Huang, M., Luo, W., Li, W., Chen, Y., Qu, S., Xue, K., Fu, W., & Wang, Y. (2021). Stroke outcome assessment: Optimizing cutoff scores for the Longshi Scale, modified Rankin Scale and Barthel Index. *PloS One*, 16(5), e0251103. <https://doi.org/10.1371/journal.pone.0251103>

6.1.2 FUGL-MEYER ASSESSMENT PARA MIEMBRO SUPERIOR EN PERSONAS CON ICTUS: UNA REVISIÓN SISTEMÁTICA PSICOMÉTRICA

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

FUGL-MEYER ASSESSMENT PARA MIEMBRO SUPERIOR EN PERSONAS CON ICTUS: UNA REVISIÓN SISTEMÁTICA PSICOMÉTRICA

El Artículo II ha sido publicado previa defensa de la tesis, cita:

de Blas Zamorano, P., Montagut-Martínez, P., Pérez-Cruzado, D., & Merchan-Baeza, J.-A. (2025). Fugl-Meyer Assessment for upper extremity in stroke: A psychometric systematic review. *Journal of Hand Therapy*. <https://doi.org/10.1016/j.jht.2025.04.004> .

Abstract

Background: Stroke is a very common pathology, with a high prevalence of impairment of the sensorimotor function of the upper limb. There are many scales that assess this variable. Fugl-Meyer Assessment is one of the most widely used, but there are many versions of this scale to assess this construct.

Purpose: To identify, critically appraise and summarize the different Fugl-Meyer Assessment (FMA) versions to recognize which is better to measure upper extremity sensorimotor function in stroke patients.

Study Design and Methods: A psychometric systematic review was carried out in this research. An exhaustive search was conducted during June and July 2023 in the following databases: Pubmed, Web of Science (WoS), Open Grey, Scielo, Cochrane, Dialnet and LILACS. By the most up-to-date COSMIN and PRISMA statements for a systematic review of the measurement properties of existing patient-reported outcome measures to select the most appropriate outcome measurement instrument. The protocol of this systematic review was registered in Open Science Framework (<https://osf.io/dy54p/>).

Results: 22 versions of the FMA were identified in 36 studies. The Upper Extremity subscale of the Fugl-Meyer Assessment (FMA-UE) was the version most frequently analyzed (n=13) among the included articles.

Conclusions: Fugl-Meyer Assessment (Full version, 50-Items, FMA) was the questionnaire version with the most evaluated psychometric characteristics and most translations, with the better score regarding the quality of evidence, which indicates that is the recommended FMA version to assess upper extremity sensorimotor function in stroke patients.

Keywords: Fugl-Meyer Assessment, Stroke, Sensorimotor function, Upper Limb

Resumen

Antecedentes: El ictus es una patología muy común, con una alta prevalencia de deterioro de la función sensoriomotora de las extremidades superiores. Existen numerosas escalas que evalúan esta variable. La escala de Fugl-Meyer es una de las más utilizadas, pero existen diversas versiones de esta escala para evaluar este constructo.

Objetivo: Identificar, evaluar críticamente y resumir las diferentes versiones de Fugl-Meyer Assessment (FMA) para reconocer cuál es la mejor versión para medir la función sensoriomotora de las extremidades superiores en personas con ictus.

Diseño y métodos del estudio: En esta investigación se realizó una revisión sistemática psicométrica. Se realizó una búsqueda exhaustiva durante junio y julio de 2023 en las siguientes bases de datos: Pubmed, Web of Science (WoS), Open Grey, Scielo, Cochrane, Dialnet y LILACS. Se utilizaron las directrices más actualizadas de COSMIN y PRISMA para una revisión sistemática de las propiedades de medición de las medidas de resultados existentes, informadas por los pacientes, con el fin de seleccionar el instrumento de medición de resultados más adecuado. El protocolo de esta revisión sistemática se registró en Open Science Framework. (<https://osf.io/dy54p/>).

Resultados: Se identificaron 22 versiones de FMA en 36 estudios. La subescala de extremidades superiores de la evaluación Fugl-Meyer (FMA-UE) fue la versión analizada con mayor frecuencia (n=13) entre los artículos incluidos.

Conclusiones: Fugl-Meyer Assessment (Versión completa, 50 ítems, FMA) fue la versión del cuestionario con más características psicométricas evaluadas y más traducciones, con mejor puntaje en cuanto a la calidad de la evidencia, lo que indica que es la versión del FMA recomendada para evaluar la función sensoriomotora de las extremidades superiores en personas con ictus.

Palabras clave: Fugl-Meyer Assessment, Ictus, Función Sensoriomotora, Miembro Superior

Lista de abreviaturas:

ARAT - Action Research Arm Test
BBT - Box and Block Test
CAHAI - Chedoke Arm and Hand Activity Inventory
COSMIN - Consensus-based standards for the selection of health status Measurement INstruments methodology
FMA - Fugl-Meyer Assessment
FMA-6 - Fugl-Meyer Assessment de miembro superior acortada (6 ítems)
FMA-12 - Forma corta de Fugl-Meyer (12 ítems)
FMA-13-Kinect - Fugl-Meyer Assessment de miembro superior (13 ítems) usando una cámara con sensor de profundidad (Kinect)
FMA-37 - Forma corta de Fugl-Meyer (37 ítems)
FMA-Col - Versión Español-Colombiano de Fugl-Meyer Assessment
FMA-Dan - Versión Danesa de Fugl-Meyer Assessment
FMA-Ita - Versión Italiana de Fugl-Meyer Assessment
FMA-Jap - Versión Japonesa de Fugl-Meyer Assessment
FMA-Kor - Versión Coreana de Fugl-Meyer Assessment
FMA-ML-10 - Fugl-Meyer de aprendizaje automático (10 ítems)
FMA-Remote - Evaluación remota de la subescala de miembro superior de Fugl-Meyer Assessment
FMA-Rom - Versión Rumana de Fugl-Meyer Assessment
FMA-S - Subescala sensitiva de Fugl-Meyer Assessment
FMA-UE-Dan: Versión Danesa de la subescala de miembro superior de Fugl-Meyer Assessment
FMA-UE-Ita - Versión Italiana de la subescala de miembro superior de Fugl-Meyer Assessment
FMA-UE-Rom - Versión Rumana de subescala de miembro superior de Fugl-Meyer Assessment
FMA-UE-SE - Subescalas de hombro y codo de Fugl-Meyer Assessment
FMA-UE - Subescala de miembro superior de Fugl-Meyer Assessment
FMA-UE-w/h - Subescalas de muñeca y mano de Fugl-Meyer Assessment
FMA-Urd - Versión Urdu de Fugl-Meyer Assessment
ICC - Intraclass Correlation coefficient / Coeficiente de Correlación Intraclass
MAL - Motor Activity Log
MAS - Motor Assessment Scale
NHPT - Nine Hole Peg Test
NIHSS - National Institute of Health Stroke Scale
OSF- Open Science Framework
PRISMA - Preferred Reporting Items for Systematic Reviews and Meta-Analyses
WMFT - Wolf Motor Function Test
WoS - Web of Science

Puntos destacados:

- Encontrar las diferentes versiones y traducciones de la escala Fugl-Meyer Assessment.
- Identificar las propiedades psicométricas analizadas en los diversos estudios de las diferentes versiones.
- Conocer la mejor versión de la escala Fugl-Meyer Assessment para evaluar la función sensoriomotora de las extremidades superiores.
- Reconocer la escala Fugl-Meyer Assessment como la escala de referencia para la evaluación de la función sensoriomotora de las extremidades superiores en la población con ictus.

1.INTRODUCCIÓN

El ictus es la segunda causa principal de muerte¹ y la primera causa principal de discapacidad en todo el mundo². Quince millones de personas sufren un ictus cada año en todo el mundo³. La carga de discapacidad después de un ictus es significativa y está aumentando muy rápidamente¹. En la fase aguda después de un ictus, aproximadamente entre el 60 y el 80 % de los sobrevivientes presentan déficits motores en las extremidades superiores o inferiores⁴.

Las limitaciones físicas son secuelas comunes después de un ictus, que se manifiestan en muchos dominios fisiológicos y aspectos motores como la espasticidad, las sinergias musculares, la coordinación, la fuerza, el rango de movimiento, la destreza, la sensibilidad de la piel, la propiocepción y la estabilidad del brazo⁵⁻⁷. Todas ellas provocan una incapacidad para utilizar la extremidad superior para alcanzar, agarrar y manipular objetos, lo que afecta el desempeño de actividades de la vida diaria como vestirse, alimentarse o bañarse^{5,8}, y un impacto perjudicial en la calidad de vida⁸.

Con estos datos, es necesario disponer de herramientas para evaluar los diferentes déficits físicos causados por el ictus. Existen numerosas herramientas para evaluar el constructo sensoriomotor y funcional de las personas con ictus. Algunas de ellas son Fugl-Meyer Assessment (FMA), Action Research Arm Test (ARAT), Wolf Motor Function Test (WMFT), Box and Block Test (BBT), etc.^{4, 9, 10}.

Elegir el instrumento más apropiado para evaluar la función sensoriomotora de las extremidades superiores en personas con ictus es un desafío, y la elección debe depender de las propiedades de medición y la viabilidad del instrumento¹¹.

Algunas revisiones sistemáticas demuestran que FMA es la más utilizada para la práctica clínica y la investigación y se ha visto un aumento en su uso en los últimos años^{9, 12, 13}.

FMA es conocida internacionalmente, con diferentes versiones y traducciones. Ha sido traducida a diferentes idiomas y culturas^{4, 9, 10, 13-15}.

La escala FMA es el "estándar de oro" para evaluar la función sensoriomotora de la extremidad superior, debido a su capacidad para evaluar aspectos como el movimiento dentro de sinergias, mezcla de sinergias, reflejos, muñeca, mano, agarre, coordinación y movimientos de velocidad; proporcionando así una gran cantidad de información que resulta de gran utilidad para comprender la capacidad sensoriomotora de la extremidad superior afectada tras un ictus^{9, 10, 12}.

A pesar de su amplio uso en la práctica clínica en rehabilitación, nunca antes se había realizado una revisión con la metodología COSMIN que evaluara todas las versiones existentes de esta escala. Esto nos permite alcanzar un consenso claro sobre cuál de las múltiples versiones existentes de FMA es la más recomendable por sus propiedades psicométricas, por ser la versión con mayor validez y, por lo tanto, la más recomendable para su uso tanto en investigación como en la práctica clínica^{17,18}.

Por todas estas razones, se ha considerado pertinente y necesario realizar una revisión sistemática para aclarar cuál de las diferentes versiones debería ser la más adecuada para la evaluación sensoriomotora del miembro superior afecto en personas con ictus. Por ello, esta revisión sistemática tuvo como objetivo identificar, evaluar críticamente y resumir las diferentes versiones de FMA para

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

identificar cuál mide mejor la función sensoriomotora de las extremidades superiores en personas con ictus.

2.MATERIAL Y MÉTODOS

2.1 Diseño

Se realizó una revisión sistemática de las propiedades de medición de las diferentes versiones de FMA que evalúan la función sensoriomotora de la extremidad superior en personas con ictus.

Se utilizó la lista de verificación COSMIN con una escala de 4 puntos para evaluar las propiedades de medición de las medidas de resultados informadas por los usuarios para seleccionar el instrumento de medición de resultados más apropiado^{17, 18}.

Se consultó la Declaraciones de elementos de informe preferidos para revisiones sistemáticas y metaanálisis (PRISMA)¹⁶⁻¹⁸ en la web de Equator y se utilizó para mejorar la presentación del informe de revision sistemática y metanálisis.

El protocolo de este proceso, con la información y la metodología utilizada, fue registrado en Open Science Framework (OSF) (<https://osf.io/dy54p/>).

2.2 Método de búsqueda

Se realizó una búsqueda exhaustiva durante junio y julio de 2023 en diferentes bases de datos: Medline (Pubmed), Web of Science (WoS), Open Grey, Scielo, Cochrane, Dialnet y LILACS, por parte de cuatro revisores independientes. La estrategia de búsqueda incluyó una combinación de los siguientes cinco aspectos: 1) búsqueda por constructo AND, 2) búsqueda por población AND, 3) búsqueda por instrumento AND, 4) filtro para propiedades de medición AND, y 5) filtro de exclusión "NOT"¹⁹. Para localizar la mayor cantidad de herramientas y sus diferentes versiones, no se limitó por fecha. Para realizar esta búsqueda, se utilizaron los siguientes términos clave: "Fugl-Meyer", "Fugl-Meyer Assessment", "FMA", "Stroke", "Cerebral Stroke", "Cerebrovascular Accident", "assessment", "tool", "instrument", "test", "measure" y "scale", adaptando esta estrategia a cada sistema de búsqueda de las diferentes bases de datos. La búsqueda se limitó a artículos publicados en inglés y español. Además, se revisó la lista de referencias de los artículos incluidos para identificar otros documentos relevantes [Tabla 1].

Tabla 1. Estrategia de búsqueda

Fórmula de búsqueda para Pubmed

- 1 ("Fugl" OR "Fugl Meyer" OR "Fugl Meyer Assessment" OR "FMA")
- 2 ("Stroke" OR "Cerebral Stroke" OR "Cerebrovascular Accident" OR "Cerebral Strokes" OR "Strokes" OR "Cerebrovascular Stroke" OR "CVA" OR "Cerebrovascular" OR "Cerebrovascular Accidents" OR "Brain Accident" OR "Brain Accidents")
- 3 ("Instrument" OR "instruments" OR "measure" OR "measures" OR "questionnaire" OR "questionnaires" OR "scale" OR "scales" OR "tool" OR "tools" OR "survey" OR "test")
- 4 ("Instrumentation" OR "methods" OR "Validation Studies" OR "Comparative Study" OR "psychometrics" OR "psychometr*" OR "clinimetr*" OR "clinometr*" OR "outcome assessment" OR "outcome assessment" OR "outcome measure*" OR "observer variation" OR "observer variation" OR "Health Status Indicators" OR "reproducibility of results" OR "reproducib*" OR "discriminant analysis" OR "reliab*" OR "unreliab*" OR "valid*" OR "coefficient" OR "homogeneity" OR "homogeneous" OR "internal consistency" OR "cronbach*" OR "correlation*" OR "selection*" OR "reduction*" OR "agreement" OR "precision" OR "imprecision" OR "precise values" OR "test-retest" OR "reliab*" OR "stability" OR "interrater" OR "inter-rater" OR "intrarater" OR "intra-rater" OR "intertester" OR "inter-tester" OR "intratester" OR "intra-tester" OR "interobserver" OR "inter-observer" OR "intraobserver" OR "intraobserver" OR "intertechician" OR "inter-technician" OR "intratechnician" OR "intra-technician" OR "interexaminer" OR "inter-examiner" OR "intraexaminer" OR "intra-examiner" OR "interassay" OR "inter-assay" OR "intraassay" OR "intra-assay" OR "interindividual" OR "inter-individual" OR "intraindividual" OR "intra-individual" OR "interparticipant" OR "inter-participant" OR "intraparticipant" OR "intra-participant" OR "kappa" OR "kappa's" OR "kappas" OR "repeatab*" OR "generaliza*" OR "generalisa*" OR "concordance" OR "intraclass correlation*" OR "discriminative" OR "known group" OR "factor analysis" OR "factor analyses" OR "dimension*" OR "subscale*" OR "item discriminant" OR "interscale correlation*" OR "error" OR "errors" OR "individual variability" OR "variability" OR "uncertainty" OR "standard error of measurement" OR "sensitiv*" OR "responsive*" OR "meaningful change" OR "ceiling effect" OR "floor effect" OR "Item response model" OR "IRT" OR "Rasch" OR "Differential item functioning" OR "DIF" OR "computer adaptive testing" OR "item bank" OR "cross-cultural equivalence")
- 5 #1 AND #2 AND #3 AND #4
- 6 ("Protocol" OR "addresses" OR "biography" OR "case reports" OR "comment" OR "editorial" OR "congresses" OR "consensus development conference" OR

"consensus development conference" OR "practice guideline")

7 #5 NOT #6

8 Filtro: Idioma (Inglés y Español)

Fórmula de búsqueda para Web of Science (WoS)

- 1 TI=("Fugl" OR "Fugl Meyer" OR "Fugl Meyer Assessment" OR "FMA")
 - 2 TS=("Stroke" OR "Cerebral Stroke" OR "Cerebrovascular Accident" OR "Cerebral Strokes" OR "Strokes" OR "Cerebrovascular Stroke" OR "CVA" OR "Cerebrovascular" OR "Cerebrovascular Accidents" OR "Brain Accident" OR "Brain Accidents")
 - 3 TS=("Instrument" OR "instruments" OR "measure" OR "measures" OR "questionnaire" OR "questionnaires" OR "scale" OR "scales" OR "tool" OR "tools" OR "survey" OR "test")
 - 4 TS=("Instrumentation" OR "methods" OR "Validation Studies" OR "Comparative Study" OR "psychometrics" OR "psychometr*" OR "clinimetr*" OR "clinometr*" OR "outcome assessment (health care)" OR "outcome assessment" OR "outcome measure*" OR "observer variation" OR "observer variation" OR "Health Status Indicators" OR "reproducibility of results" OR "reproducib*" OR "discriminant analysis" OR "reliab*" OR "unreliab*" OR "valid*" OR "coefficient" OR "homogeneity" OR "homogeneous" OR "internal consistency" OR ("cronbach*" AND ("alpha" OR "alphas")) OR ("item" AND ("correlation*" OR "selection*" OR "reduction*")) OR "agreement" OR "precision" OR "imprecision" OR "precise values" OR "test-retest" OR ("test" AND "retest") OR ("reliab*" AND ("test" OR "retest")) OR "stability" OR "interrater" OR "inter-rater" OR "intrarater" OR "intra-rater" OR "intertester" OR "inter-tester" OR "intratester" OR "intra-tester" OR "interobserver" OR "inter-observer" OR "intraobserver" OR "intraobserver" OR "intertechnician" OR "inter-technician" OR "intratechnician" OR "intra-technician" OR "interexaminer" OR "inter-examiner" OR "intraexaminer" OR "intra-examiner" OR "interassay" OR "inter-assay" OR "intraassay" OR "intra-assay" OR "interindividual" OR "inter-individual" OR "intraindividual" OR "intra-individual" OR "interparticipant" OR "inter-participant" OR "intraparticipant" OR "intra-participant" OR "kappa" OR "kappa's" OR "kappas" OR "repeatab*" OR ("replicab*" OR "repeated") AND ("measure" OR "measures" OR "findings" OR "result" OR "results" OR "test" OR "tests")) OR "generaliza*" OR "generalisa*" OR "concordance" OR ("intraclass" AND "correlation*") OR "discriminative" OR "known group" OR "factor analysis" OR "factor analyses" OR "dimension*" OR "subscale*" OR ("multitrait" AND "scaling" AND ("analysis" OR "analyses")) OR "item discriminant" OR "interscale correlation" OR "error" OR "errors" OR "individual variability" OR ("variability" AND
-

("analysis" OR "values")) OR ("uncertainty" AND ("measurement" OR "measuring")) OR "standard error of measurement" OR "sensitiv*" OR "responsive*" OR (("minimal" OR "minimally" OR "clinical" OR "clinically") AND ("important" OR "significant" OR "detectable") AND ("change" OR "difference")) OR ("small*" AND ("real" OR "detectable") AND ("change" OR "difference")) OR "meaningful change" OR "ceiling effect" OR "floor effect" OR "Item response model" OR "IRT" OR "Rasch" OR "Differential item functioning" OR "DIF" OR "computer adaptive testing" OR "item bank" OR "cross-cultural equivalence")

5 #1 AND #2 AND #3 AND #4

6 ("Protocol" OR "addresses" OR "biography" OR "case reports" OR "comment" OR "editorial" OR "congresses" OR "consensus development conference" OR "consensus development conference" OR "practice guideline")

7 #5 NOT #6

8 Filtro: Idioma (Inglés y Español)

Fórmula de búsqueda para for Open Grey

1 ("Fugl" OR "Fugl Meyer" OR "Fugl Meyer Assessment" OR "FMA")

2 ("Stroke" OR "Cerebral Stroke" OR "Cerebrovascular Accident" OR "Cerebral Strokes" OR "Strokes" OR "Cerebrovascular Stroke" OR "CVA" OR "Cerebrovascular" OR "Cerebrovascular Accidents" OR "Brain Accident" OR "Brain Accidents")

3 ("Instrument" OR "instruments" OR "measure" OR "measures" OR "questionnaire" OR "questionnaires" OR "scale" OR "scales" OR "tool" OR "tools" OR "survey" OR "test")

4 ("Instrumentation" OR "methods" OR "Validation Studies" OR "Comparative Study" OR "psychometrics" OR "psychometr*" OR "clinimetr*" OR "clinometr*" OR "outcome assessment (health care)" OR "outcome assessment" OR "outcome measure*" OR "observer variation" OR "observer variation" OR "Health Status Indicators" OR "reproducibility of results" OR "reproducib*" OR "discriminant analysis" OR "reliab*" OR "unreliab*" OR "valid*" OR "coefficient" OR "homogeneity" OR "homogeneous" OR "internal consistency" OR ("cronbach*" AND ("alpha" OR "alphas")) OR ("item" AND ("correlation*" OR "selection*" OR "reduction*")) OR "agreement" OR "precision" OR "imprecision" OR "precise values" OR "test-retest" OR ("test" AND "retest") OR ("reliab*" AND ("test" OR "retest")) OR "stability" OR "interrater" OR "inter-rater" OR "intra-rater" OR "intertester" OR "inter-tester" OR "intratester" OR "intra-tester" OR "interobserver" OR "inter-observer" OR "intraobserver" OR "intraobserver" OR "intertechician" OR "inter-technician")

OR "intratechnician" OR "intra-technician" OR "interexaminer" OR "inter-examiner" OR "intraexaminer" OR "intra-examiner" OR "interassay" OR "inter-assay" OR "intraassay" OR "intra-assay" OR "interindividual" OR "inter-individual" OR "intraindividual" OR "intra-individual" OR "interparticipant" OR "inter-participant" OR "intraparticipant" OR "intra-participant" OR "kappa" OR "kappa's" OR "kappas" OR "repeatab*" OR ("replicab*" OR "repeated") AND ("measure" OR "measures" OR "findings" OR "result" OR "results" OR "test" OR "tests")) OR "generaliza*" OR "generalisa*" OR "concordance" OR ("intraclass" AND "correlation*") OR "discriminative" OR "known group" OR "factor analysis" OR "factor analyses" OR "dimension*" OR "subscale*" OR ("multitrait" AND "scaling" AND ("analysis" OR "analyses")) OR "item discriminant" OR "interscale correlation" OR "error" OR "errors" OR "individual variability" OR ("variability" AND ("analysis" OR "values")) OR ("uncertainty" AND ("measurement" OR "measuring")) OR "standard error of measurement" OR "sensitiv*" OR "responsive*" OR ("minimal" OR "minimally" OR "clinical" OR "clinically") AND ("important" OR "significant" OR "detectable") AND ("change" OR "difference")) OR ("small*" AND ("real" OR "detectable") AND ("change" OR "difference")) OR "meaningful change" OR "ceiling effect" OR "floor effect" OR "Item response model" OR "IRT" OR "Rasch" OR "Differential item functioning" OR "DIF" OR "computer adaptive testing" OR "item bank" OR "cross-cultural equivalence")

5 #1 AND #2 AND #3 AND #4

6 ("Protocol" OR "addresses" OR "biography" OR "case reports" OR "comment" OR "editorial" OR "congresses" OR "consensus development conference" OR "consensus development conference" OR "practice guideline")

7 #5 NOT #6

8 Filtro: Idioma (Inglés y Español)

Fórmula de búsqueda para Scielo

1 (Fugl Meyer)

2 (Stroke)

3 #1 AND #2

4 Filtro: Idioma (Inglés y Español)

Fórmula de búsqueda para Cochrane

1 ("Fugl*":ti,ab OR "Fugl Meyer*":ti,ab OR "Fugl Meyer Assessment*":ti,ab OR "FMA*":ti,ab)

-
- 2 ("Stroke*":ti,ab OR "Cerebral Stroke*":ti,ab OR "Cerebrovascular Accident*":ti,ab OR "CVA*":ti,ab OR "Cerebrovascular*":ti,ab OR "Brain Accident*":ti,ab OR "Cerebrovascular Stroke*":ti,ab)
 - 3 ("Instrument":ti,ab OR "instruments":ti,ab OR "measure":ti,ab OR "measures":ti,ab OR "questionnaire":ti,ab OR "questionnaires":ti,ab OR "scale":ti,ab OR "scales":ti,ab OR "tool":ti,ab OR "tools":ti,ab OR "survey":ti,ab OR "test":ti,ab)
 - 4 AB ("Instrumentation" OR methods" OR "Validation Studies" OR "Comparative Study" OR "psychometrics" OR "psychometr*" OR "clinimetr*" OR "clinometr*" OR "outcome assessment (health care)" OR "outcome assessment" OR "outcome measure*" OR "observer variation" OR "observer variation" OR "Health Status Indicators" OR "reproducibility of results" OR "reproducib*" OR "discriminant analysis" OR "reliab*" OR "unreliab*" OR "valid*" OR "coefficient" OR "homogeneity" OR "homogeneous" OR "internal consistency" OR ("cronbach*" AND ("alpha" OR "alphas")) OR ("item" AND ("correlation*" OR "selection*" OR "reduction*")) OR "agreement" OR "precision" OR "imprecision" OR "precise values" OR "test-retest" OR ("test" AND "retest") OR ("reliab*" AND ("test" OR "retest")) OR "stability" OR "interrater" OR "inter-rater" OR "intra-rater" OR "intertester" OR "inter-tester" OR "intratester" OR "intra-tester" OR "interobserver" OR "inter-observer" OR "intraobserver" OR "intraobserver" OR "intertechician" OR "inter-technician" OR "intratechnician" OR "intra-technician" OR "interexaminer" OR "inter-examiner" OR "intraexaminer" OR "intra-examiner" OR "interassay" OR "inter-assay" OR "intraassay" OR "intra-assay" OR "interindividual" OR "inter-individual" OR "intraindividual" OR "intra-individual" OR "interparticipant" OR "inter-participant" OR "intraparticipant" OR "intra-participant" OR "kappa" OR "kappa's" OR "kappas" OR "repeatab*" OR (("replicab*" OR "repeated") AND ("measure" OR "measures" OR "findings" OR "result" OR "results" OR "test" OR "tests")) OR "generaliza*" OR "generalisa*" OR concordance OR (intraclass AND correlation*) OR discriminative OR "known group" OR "factor analysis" OR "factor analyses" OR dimension* OR subscale* OR (multitrait AND scaling AND (analysis OR analyses)) OR "item discriminant" OR "interscale correlation" OR error OR errors OR "individual variability" OR (variability AND (analysis OR values)) OR (uncertainty AND (measurement OR measuring)) OR "standard error of measurement" OR sensitiv* OR responsive* OR ((minimal OR minimally OR clinical OR clinically) AND (important OR significant OR detectable) AND (change OR difference)) OR (small*AND (real OR detectable) AND (change OR difference)) OR "meaningful change" OR "ceiling effect" OR "floor effect" OR "Item response model" OR IRT OR Rasch OR "Differential item functioning" OR DIF OR "computer adaptive testing" OR "item bank" OR "cross-cultural equivalence")
 - 5 #1 AND #2 AND #3 AND #4
-

6 ("Protocol" OR "addresses" OR "biography" OR "case reports" OR "comment" OR "editorial" OR "congresses" OR "consensus development conference" OR "consensus development conference" OR "practice guideline")

7 #5 NOT #6

8 Filtro: Idioma (Inglés y Español)

Fórmula de búsqueda para Dialnet

1 (Fugl Meyer Assessment)

2 (Stroke)

3 #1 AND #2

4 (Tesis)

5 #3 NOT #4

6 Filtro: Idioma (Inglés y Español)

Fórmula de búsqueda para LILACS

1 ((Fugl) OR (Fugl Meyer) OR (Fugl Meyer Assessment) OR (FMA))

2 (Stroke)

3 ((Psychometr*) OR (clinimetr*))

4 #1 AND #2 AND #3

5 Filtro: Idioma (Inglés y Español)

Fuente: elaboración propia

2.3 Criterios de elegibilidad

Los criterios de inclusión para este estudio fueron: a) personas que habían sufrido un ictus, b) el resultado debía ser la función sensoriomotora de la extremidad superior evaluada con cualquier versión de FMA, c) los estudios debían incluir una versión de FMA y evaluar al menos una propiedad psicométrica de la lista de verificación COSMIN.

Los criterios de exclusión fueron: a) instrumento o artículo no disponible en idioma inglés o español, b) editoriales, discursos, resúmenes de congresos, biografías, comentarios, reportes de casos, tesis doctorales, guías de práctica, artículos de opinión o artículos en los que no se utilizó una metodología selectiva.

Cuatro revisores examinaron de forma independiente los títulos y resúmenes para comprobar si cumplían los criterios de inclusión. Cuando fue imposible tomar decisiones basándose únicamente en el resumen y el título, los cuatro revisores realizaron el análisis del texto completo. Los conflictos se resolvieron compartiendo los diferentes aspectos problemáticos y discutiéndolos entre los cuatro revisores hasta llegar a un acuerdo.

2.4 Evaluación de la calidad

Los estudios incluidos en esta revisión fueron evaluados por cuatro investigadores, de acuerdo con la lista de verificación de riesgo de sesgo COSMIN más actualizada con una escala de 4 puntos¹⁸ (disponible en <https://www.cosmin.nl/>). Contiene nueve casillas con entre 5 y 18 ítems, y cada casilla proporciona una puntuación de calidad metodológica para la validez de contenido, validez estructural, consistencia interna, validez transcultural, error de medición, fiabilidad, validez de criterio, prueba de hipótesis y sensibilidad. El contenido evaluado en cada casilla puede calificarse como "muy bueno", "adecuado", "dudoso" e "inadecuado". La puntuación de cada propiedad psicométrica se obtiene tomando la puntuación más baja de cualquier ítem de cada propiedad¹⁸. Dado que el volumen de artículos no ha sido excesivo, no se han establecido puntos de corte para descartar algún artículo por calidad metodológica.

2.5 Extracción de datos

Se revisaron los títulos y resúmenes de todos los artículos identificados en esta revisión. Cuatro expertos revisaron de forma independiente los artículos completos y los conflictos se resolvieron por consenso. La información de cada artículo se extrajo con base en los siguientes aspectos: nombre de la herramienta y versión utilizada, población de estudio, características de la herramienta, análisis de propiedades psicométricas, puntuación COSMIN y resultados estadísticos de la investigación.

2.6 Síntesis

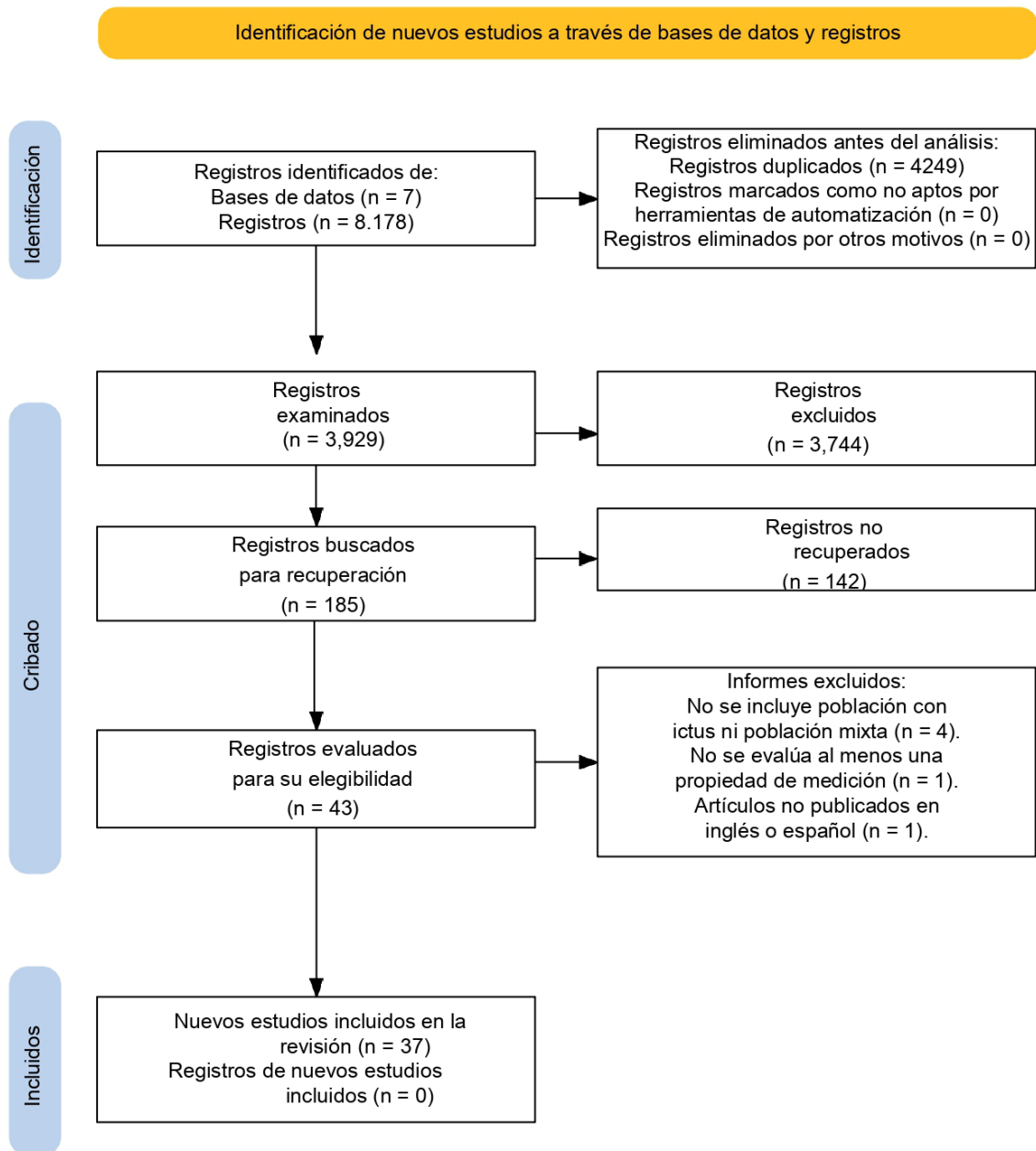
Los datos sobre las propiedades psicométricas de los artículos que cumplieron los criterios de inclusión se calificaron según la lista de verificación COSMIN¹⁸. Cuatro revisores independientes establecieron las puntuaciones de las propiedades psicométricas de cada artículo, y resolvieron las diferencias mediante discusión y consenso. Posteriormente, el resultado de cada una de las propiedades de medición de cada estudio se calificó según los criterios de Terwee¹¹. Clasificando cada resultado como "positivo" (+), "negativo" (-) o "indeterminado" (?). Finalmente, tras determinar la calidad metodológica de cada estudio individual y evaluar las propiedades de medición, los resultados se resumieron cualitativamente para determinar tanto la calidad de la evidencia como la solidez de las recomendaciones, calificándolas como "alta", "moderada", "baja" o "muy baja" según el riesgo de sesgo, inconsistencia, imprecisión e indirección, utilizando el sistema GRADE²⁰. Cuatro revisores independientes asignaron puntuaciones para las propiedades psicométricas de cada artículo y las diferencias se resolvieron mediante discusión y consenso.

3.RESULTADOS

3.1 Resultados de la búsqueda

Se identificaron 8.178 artículos, de los cuales, 36 fueron finalmente incluidos en esta revisión sistemática, hasta julio de 2023. El proceso de selección se refleja en el diagrama de flujo, que se muestra en la Figura 1.

Figura 1. Diagrama de flujo para la selección de los artículos



Fuente: elaboración propia

3.2 Características de las versiones incluidas de Fugl-Meyer Assessment

Se identificaron veintidós versiones de FMA diseñadas para evaluar la función sensoriomotora de las extremidades superiores en personas con ictus. Las características de las versiones del FMA incluidas para la medición de la función sensoriomotora de las extremidades superiores y los estudios se presentan en la Tabla 2. En todos los casos, el idioma original utilizado fue el inglés. Sin embargo, se encontraron varios de estos cuestionarios en los que se realizó una validación transcultural: Colombiano²⁷, Danés^{28, 49}, Italiano^{29, 50}, Japonés³⁰, Coreano³¹, Urdu³², Rumano³³ y Portugués brasileño³⁷.

La mayoría de los estudios se realizaron en hospitales y centros de rehabilitación neurológica de Asia (n=16, 40%), América (n=13, 32,5%), Europa (n=10, 25%) y Oriente Medio (n=1, 2,5%). Los lugares donde se desarrollaron los estudios con mayor frecuencia fueron Estados Unidos (n=7)^{25, 40, 41, 44, 53-55}, Taiwán (n=7)^{14, 23, 26, 34, 35, 42, 43} y Japón (n=4)^{15, 30, 36, 52}.

En cuanto al tamaño de la muestra de los estudios, osciló entre 10 participantes incluidos en algunas traducciones y adaptaciones transculturales de la FMA^{27-29, 32, 33} a 512 participantes en la Subescala de miembro superior de Fugl-Meyer Assessment (FMA-UE)⁴¹, con amplia heterogeneidad en términos de tamaño de muestra.

En cuanto al número de ítems, las versiones oscilaron entre 6 ítems en el FMA-6³⁶ hasta 50 ítems en la versión completa de FMA^{14, 21-26}.

Para el tipo de ítem utilizado en cada versión, todas mantuvieron el formato de puntuación original en una escala ordinal de 3 puntos (0-2).

3.2.1 Fugl Meyer Assessment (versión completa, 50 ítems, FMA)

La versión completa inicial de FMA estuvo presente en siete estudios^{14, 21-26}, y se evaluaron todas las propiedades psicométricas. La validez transcultural y la invariancia de la medición se evaluaron en las siete versiones traducidas (FMA-Col²⁷, FMA-Dan²⁸, FMA-Ita²⁹, FMA-Japón³⁰, FMA-Kor³¹, FMA-Urd³², FM-Rom³³), siendo la versión más traducida.

3.2.2 Versiones cortas de FMA

En esta revisión, se han detectado tres versiones cortas de la escala FMA, con la excepción de las versiones de subescalas, como FMA-UE, FMA-UE-SE, FMA-UE-w/h o FMA-S. Estas versiones son:

3.2.2.1 FMA-37

La forma corta de la FMA con 37 ítems se evaluó en tres estudios^{14, 26, 34}. Para todas las propiedades psicométricas, excepto la validez de contenido, la consistencia interna y la validez transcultural. La validez estructural reveló una estructura unidimensional^{14,26}. Los valores de fiabilidad test-retest fueron $\geq 0,92$. La validez de criterio se situó entre $r=0,01^{26}$ y $r=0,093^{14}$. Las pruebas de hipótesis oscilaron entre $r=0,73$ y $r=0,75^{14}$.

3.2.2.2 FMA-12

Este instrumento fue incluido en cinco estudios^{14, 24, 26, 34, 35}. Se evalúan las mismas propiedades psicométricas que la versión FMA-37 (excepto la validez de contenido, la consistencia interna y la validez transcultural). La validez estructural muestra un constructo unidimensional^{14,26}. El coeficiente de correlación osciló entre $r=0,01$ ²⁶ hasta $r=0,99$ ²⁴. La prueba de hipótesis mostró una correlación entre $r=0,39$ ³⁵ y $r=0,75$ ¹⁴.

3.2.2.3 FMA-UE-6

La FMA acortada de miembro superior de 6 ítems se incluyó en un solo estudio³⁶, evaluando cinco propiedades psicométricas: la consistencia interna (alfa de Cronbach= $0,924$), la fiabilidad (ICC= $0,994$), la validez de criterio ($r=0,97$) y la prueba de hipótesis ($r=0,91-0,94$).

3.2.3 Versiones tecnológicas

3.2.3.1 FMA-ML-10

El Fugl Meyer de aprendizaje automático de 10 ítems se evaluó en un estudio¹⁴, evaluando las siguientes propiedades psicométricas: validez estructural (estructura unidimensional), error de medición, fiabilidad (ICC= $0,92$), validez de criterio ($r=0,99$) y prueba de hipótesis ($r=0,67-0,68$).

3.2.3.2 FMA-13-Kinect

La FMA de extremidades superiores (13 ítems) evaluada mediante cámara con sensor de profundidad (Kinect) se incluyó en un solo estudio²⁷, con un coeficiente de correlación de validez de criterio de $r=0,873$ y un valor r de prueba de hipótesis de $0,387$.

3.2.3.3 FMA-Remoto

La evaluación remota de las extremidades superiores con FMA solo se incluyó en un estudio⁵², midiendo sólo una propiedad psicométrica, la fiabilidad test-retest, con ICC= $0,998$.

3.2.4 Subescala de extremidades superiores de la evaluación Fugl-Meyer (FMA-UE)

FMA-UE es la versión más evaluada en esta revisión sistemática, siendo evaluada en 13 estudios^{15,36,38-48}. Se evaluaron prácticamente todas las propiedades psicométricas, excepto la validez de contenido. Esta versión se ha traducido a tres idiomas: Danés (FMA-UE-Dan⁴⁹), Italiano (FMA-UE-Ita⁵⁰) y Rumano (FMA-UE-Rom⁵¹).

Tabla 2. Características de los estudios y descripción de Fugl Meyer Assessment en personas con ictus.

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Fugl-Meyer Assessment (Versión completa, 50-Ítems) (FMA) Sanford et al. (1993)	12 pacientes con ictus.	Ontario, Canada. Chedoke-McMaster Rehabilitation Centre. Los pacientes fueron evaluados por 3 terapeutas en ocasiones separadas.	Fugl-Meyer Assessment es una escala de evaluación de déficit físico que consta de 155 ítems, calificados en una escala ordinal de 3 puntos (0-2). 33 ítems corresponden a extremidades superiores, 17 ítems a extremidades inferiores, 7 ítems a equilibrio, 12 ítems a sensibilidad y 22 ítems a movilidad articular pasiva y dolor articular. La FMA incluye tres categorías: (1) función motora, (2) equilibrio y (3) sensibilidad. La puntuación máxima para cada una de estas	1. Fiabilidad 2. Error de medición	1. Dudoso 2. Dudoso	1. El ICC (CI del 95 %) fue de 0,96 para la puntuación total de Fugl-Meyer Assessment. Las subsecciones variaron de 0,61 para las mediciones de dolor a 0,97 para la extremidad superior.

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
			categorías y subcategorías fue: función motora para las extremidades superiores (66 puntos) e inferiores (34 puntos); equilibrio en sedestación (6 puntos) y en bipedestación (8 puntos); tacto (24 puntos) y propiocepción (16 puntos). La parte motora de la escala incluye 50 ítems de función motora de extremidades superiores e inferiores, con una puntuación posible de 0 a 100 puntos (66 de la subescala de extremidades superiores y 34 de la subescala de extremidades inferiores).			

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Malouin et al. (1994)	32 pacientes con ictus.	Quebec, Canada. Centro de Rehabilitación Francois-Charron.	Fugl-Meyer Assessment como la de anterior.	1. Validez de contenido 2. Prueba de hipótesis 3. Sensibilidad	1. Dudoso 2. Muy buena 3. Muy buena	2. La correlación entre las puntuaciones totales de la MAS y la FMA (incluyendo la función motora, el equilibrio y la sensibilidad) fue de $r = 0,96$; esta correlación disminuyó a $0,91$ al eliminar la sensibilidad. Las relaciones más sólidas se observaron para la función motora de las extremidades superiores ($r = 0,93$). Todas las correlaciones fueron positivas, fuertes y significativas ($p < 0,005$), excepto para la puntuación de equilibrio en sedestación de la FMA ($r = -0,10$).

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Lin et al. (2004)	176 pacientes con ictus.	Taiwán. La muestra del estudio se reclutó del registro del Estudio de Calidad de Vida después de un Ictus en Taiwán, del 1 de diciembre de 1999 al 7 de diciembre de 2001. Este estudio es un estudio de cohorte prospectivo de pacientes con ictus ingresados en el Hospital Universitario Nacional de Taiwán.	Fugl-Meyer Assessment como la anterior.	1. Consistencia interna 2. Fiabilidad 3. Validez de criterio 4. Prueba de hipótesis 5. Sensibilidad	1. Muy buena 2. Inadecuada 3. Muy buena 4. Muy buena 5. Muy buena	1. Los alfa de Cronbach de la FMA-S en cuatro momentos tras el ictus oscilaron entre 0,94 y 0,98, lo que indica una excelente consistencia interna. 2. La concordancia interevaluador de la puntuación total de la FMA-S fue excelente, con un ICC (CI del 95 %) de 0,93 (0,85-0,96). 3. Las puntuaciones de la FMA-S mostraron una correlación entre débil y moderada con las puntuaciones de la FMA-M (rho de Spearman entre 0,31 y 0,44). 4. Las puntuaciones de la FMA-S mostraron una correlación entre débil y moderada con el BI (rho de Spearman entre 0,38 y 0,53).

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Hsueh et al. (2008)	50 pacientes con ictus para la primera parte del estudio (validez y capacidad de respuesta) y 60 pacientes con ictus para la segunda parte del estudio (fiabilidad).	China. Departamento de rehabilitación del Hospital Universitario Kaohsiung.	Fugl-Meyer Assessment como la anterior.	1. Fiabilidad 2. Error de medición 3. Validez de criterio 4. Prueba de hipótesis 5. Sensibilidad	1. Dudoso 2. Dudoso 3. Muy buena 4. Muy buena 5. Muy buena	1. La fiabilidad test-retest fue alta para las subescalas UE y LE, así como para las puntuaciones motoras totales de las 4 escalas motoras (todos los ICC $\geq 0,93$ y límites inferiores del IC del 95 % $\geq 0,89$). 3. Las correlaciones fueron altas ($p = 0,91-0,99$) entre las 4 escalas motoras clínicas y sus subescalas al ingreso y al alta, lo que indica una alta validez concurrente. 4. Las puntuaciones de las 4 escalas motoras y sus subescalas al ingreso mostraron una correlación moderada con las puntuaciones del BI al alta ($p \geq 0,66$), lo que indica una validez predictiva aceptable.

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
See et al. (2013)	31 pacientes con ictus.	Los datos de los pacientes proceden de dos ensayos clínicos robóticos de terapia motora del brazo realizados en Estados Unidos, Canadá y Alemania.	Fugl-Meyer Assessment como la anterior.	1. Fiabilidad 2. Prueba de hipótesis 3. Sensibilidad	1. Dudoso 2. Muy buena 3. Muy buena	1. Para la fiabilidad intraevaluador, el ICC fue de 0,99 para la puntuación total de FMA, la subpuntuación proximal de FMA y la subpuntuación de FMA-p/h. Para la fiabilidad interevaluador, el ICC fue de 0,99 para la puntuación total de FMA y de 0,98 para las subpuntuaciones proximal y w/h de FMA. 2. La correlación de la puntuación total de FMA con ARAT fue de 0,93; con fuerza de pinza fue de 0,88; con fuerza de agarre fue de 0,74; con BBT fue de 0,86; con NHPT fue de 0,75 y con SIS fue de 0,86.

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Chen et al. (2014)	301 pacientes con ictus a los 14 días de la evaluación del ictus; 226 a los 90 días y 213 participant es evaluados a los 180 días posteriores al ictus.	Taiwán. Los datos se obtuvieron de un estudio previo. Los pacientes fueron evaluados con la escala motora de Fugl-Meyer de 50 ítems cuatro veces, a los 14, 30, 90 y 180 días después del inicio del ictus. Todas las evaluaciones fueron realizadas por un terapeuta ocupacional entrenado.	Fugl-Meyer Assessment como la anterior.	1. Validez estructural 2. Validez de criterio 3. Sensibilidad	1. Dudoso 2. Muy buena 3. Muy buena	1. El análisis paralelo latente y el análisis factorial ordinal indicaron que todos los ítems de W/H-UE-FM representan un único constructo unidimensional: la capacidad motora de muñeca y mano. 2. La puntuación FM-50 frente a FM-37 fue de 0,61 a 0,80. La puntuación FM-50 frente a FM-12 fue de 0,80 a 0,00. La puntuación FM-37 frente a FM-12 fue de 0,01 a 0,03.

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Lin et al. (2021)	267 pacientes con ictus.	Taiwán. Pacientes de datos secundarios de otros estudios.	Fugl-Meyer Assessment como la anterior.	1. Validez estructural 2. Fiabilidad 3. Error de medición 4. Prueba de hipótesis 5. Sensibilidad	1. Dudosa 2. Inadecuada 3. Inadecuada 4. Dudosa 5. Dudosa	1. Estructura unidimensional. 2. El ICC para FM fue de 0,98. 4. La correlación de Pearson con el BI para FMA fue de 0,67. La correlación de Pearson con el PASS para FMA fue de 0,69.
Versión Español-Colombiano de Fugl-Meyer Assessment (FMA-Col)						
Barbosa et al. (2019)	10 pacientes con ictus.	Bogota, Colombia. Hospital Militar Central de Colombia.	Los autores realizaron una traducción de la versión completa de la Evaluación Fugl-Meyer mencionada anteriormente, a una versión	1. Validez de contenido 2. Validez transcultural	1. Dudoso 2. Inadecuada	

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
			colombiana utilizando un protocolo estandarizado de traducción hacia adelante y hacia atrás.			
Versión Danesa de Fugl-Meyer Assessment (FMA-Dan) Busk et al. (2022)	10	Dinamarca. Centro Regional de Neurorrehabilitación del Hospital Nästved.	Los autores realizaron una traducción de la versión completa de la evaluación Fugl-Meyer mencionada anteriormente, a una versión danesa utilizando un procedimiento de ocho pasos.	1. Validez transcultural	1. Inadecuada	

Versión Italiana de Fugl-Meyer Assessment (FMA-Ita)

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Cecchi et al. (2021)	10 pacientes con ictus.	Floencia, Italia. Unidad de rehabilitación de la Fundación Don Carlo Gnocchi, Instituto Científico.	Un equipo multidisciplinar de expertos tradujo la versión completa de la Evaluación Fugl-Meyer de la Universidad de Gotemburgo, como se mencionó anteriormente, a la versión italiana de la misma, siguiendo un proceso estandarizado de traducciones, tanto directas como inversas, revisado por un comité de expertos. Posteriormente, realizaron la validación clínica mediante un estudio piloto con 10 sujetos con ictus.	1. Validez de contenido 2. Validez transcultural	1. Dudosa 2. Inadecuada	
Versión Japonesa de Fugl-Meyer Assessment (FMA-Jap)						
Amano et al. (2018)a	30 pacientes con ictus.	Japón. Hospital de la Facultad de Medicina de Hyogo.	Los autores realizaron una traducción de la versión completa de la evaluación Fugl-	1. Consistencia interna 2. Validez	1. Muy buena 2.	1. El alfa de Cronbach de todos los dominios del FMA (número de ítems: 63) fue de 0,973.

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
		En este estudio, se realizó una traducción del manual de la FMA y un análisis de las propiedades psicométricas.	Meyer mencionada anteriormente, a una versión japonesa.	transcultural 3. Fiabilidad 4. Error de medición 5. Prueba de hipótesis 6. Sensibilidad	Inadecuada 3. Inadecuada 4. Inadecuada 5. Muy buena 6. Muy buena	3. En cuanto a la fiabilidad interevaluador simultánea y no simultánea, el ICC osciló entre 0,809 y 0,983 (intervalo de confianza [CI] del 95 %: 0,639-0,994; p < 0,001) y entre 0,991 y 0,999 (CI del 95 %: 0,981-0,999; p < 0,001), respectivamente. En cuanto a la fiabilidad intraevaluador simultánea, el ICC osciló entre 0,994 y 0,999 (CI del 95 %: 0,988-1,000; p < 0,001). 5. La sección motora de la FMA presentó los valores más altos (entre 0,926 y 0,950), y la sección sensitiva los más bajos (entre 0,268 y 0,370). La mayor correlación se observa con ARAT (0,950 para la sección motora de la FMA y 0,947 para la FMA total).

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Versión Coreana de Fugl-Meyer Assessment (FMA-Kor) Kim et al. (2021)	31 pacientes con ictus.	Corea. Instituto Nacional de Investigación de Rehabilitación de Corea. Centro Nacional de Rehabilitación.	En primer lugar, tres traductores y un comité de traducción realizaron la traducción de la versión completa del FMA, como se mencionó anteriormente, a la versión coreana. En segundo lugar, analizaron las propiedades psicométricas del FMA-Kor.	1. Validez transcultural 2. Fiabilidad 3. Prueba de hipótesis 4. Sensibilidad	1. Inadecuada 2. Inadecuada 3. Muy buena 4. Muy buena	2. En cuanto a la fiabilidad intraevaluador, todos los dominios presentaron un ICC >0,72, excepto el ROM pasivo para LE y el dolor articular para LE. La función motora de UE presentó el ICC más alto (0,977) y el ROM pasivo de LE presentó el ICC más bajo (0,512). En cuanto a la fiabilidad interevaluador, todos los dominios presentaron un ICC >0,72, excepto el ROM pasivo para LE. La sensibilidad táctil para UE presentó el ICC más alto (0,995) y el ROM pasivo

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
						para LE presentó el ICC más bajo (0,498).
						3. La FMA-Kor se correlacionó significativamente con MI, que presentó la correlación más alta (r = 0,788). También se correlacionó significativamente con la MAS, con un coeficiente de correlación más bajo (r = 0,496). No se correlacionó significativamente con la puntuación de la BBS (r = 0,22).
Versión Urdu de Fugl-Meyer Assessment (FMA-Urd)						
Ikram et al. (2022)	10 pacientes con ictus para prueba	Pakistán. Se realizó una traducción estandarizada, paso a paso, de la FMA al	Los autores realizaron una traducción de la versión completa de la Evaluación Fugl-Meyer mencionada anteriormente, a una versión en	1. Validez de contenido 2. Validez estructural 3. Consistencia	1. Dudosa 2. Inadecuada 3. Muy buena	2. El análisis factorial y el análisis de diagrama de sedimentación subsiguientes mostraron que la varianza total se describió principalmente

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
	piloto y 50 pacientes con ictus crónico para analizar las propiedades psicométricas.	urdu con la ayuda de traductores bilingües, expertos de la Universidad Internacional Riphah y expertos familiarizados con la versión completa de la escala FMA para garantizar la equivalencia conceptual.	urdu.	interna 4. Validez transcultural 5. Fiabilidad 6. Prueba de hipótesis 7. Sensibilidad	4. Inadecuada 5. Inadecuada 6. Muy buena 7. Muy buena	mediante dos componentes. El componente 1 se asoció con la extremidad superior. El componente 2 se asoció con la extremidad inferior. 3. El alfa de Cronbach de las evaluaciones de las extremidades superiores e inferiores varió de 0,80 a 0,87. 5. Los coeficientes kappa ponderados para la fiabilidad intra e interevaluador variaron de 0,75 a 0,99. 6. El análisis de la FMA en urdu mostró una correlación moderada con la FIM ($r = 0,70$) y una correlación negativa moderada con la mRS ($r = -0,69$) y la NIHSS ($r = -0,79$).

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Versión Rumana de Fugl-Meyer Assessment (FMA-Rom) Onose et al. (2023)	10	Bucarest, Rumania División de Clínica Neuromuscular del Hospital Universitario de Urgencias “Bagdasar-Arseni”, en el Instituto Nacional de Neurología y Enfermedades Neurovasculares.	Los autores realizaron una traducción de la versión completa de la evaluación Fugl-Meyer mencionada anteriormente, a una versión rumana.	1. Validez de contenido 2. Validez transcultural 3. Fiabilidad 4. Prueba de hipótesis 5. Sensibilidad	1. Dudosa 2. Inadecuada 3. Inadecuada 4. Inadecuada 5. Inadecuada	3. PA ≥ 40% 4. Los valores de correlación obtenidos fueron muy buenos para BI y cercanos al límite para mRS. FMA-BI r = 0,917. FMA-mRS r = -0,59.
Forma corta de Fugl-Meyer (37-item) (FMA-37)						

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Hsieh et al. (2007)	279 pacientes con ictus.	Países Bajos y Taiwán. Hospital Universitario Nacional de Taiwán.	Escala motora Fugl-Meyer de 37 ítems (26 ítems UE y 11 ítems LE), puntuada con una escala de 3 puntos, que va de 0 a 2, con una puntuación máxima de 74 puntos.	1. Fiabilidad 2. Validez de criterio 3. Sensibilidad	1. Inadecuada 2. Muy buena 3. Muy buena	1. La fiabilidad de Rasch para ambas escalas S-FM fue buena (coeficientes $\geq 0,92$). 2. Ambas subescalas de S-FM y FM en 3 puntos temporales presentaron una validez predictiva moderada con la función integral de las actividades de la vida diaria ($r = 0,49$ a $0,59$ para S-FM y $0,48$ a $0,53$ para FM).
Chen et al. (2014)	301 pacientes con ictus a los 14 días de la evaluación del ictus; 276 sujetos a los 30 días	Taiwán. Los datos se obtuvieron de un estudio previo. Los pacientes fueron evaluados con la escala motora de Fugl-Meyer de 50 ítems cuatro veces, a los 14, 30, 90 y 180 días	Fugl-Meyer 37 ítems como la anterior.	1. Validez estructural 2. Validez de criterio 3. Sensibilidad	1. Dudoso 2. Muy buena 3. Muy buena	1. El análisis paralelo latente y el análisis factorial ordinal indicaron que todos los ítems de FMA de la UE de W/H representan un único constructo unidimensional: la capacidad motora de la muñeca y la mano. 2. La relación FM-50 vs. FM-37 fue de $0,61$ a $0,80$. FM-50 vs. FM-12 fue de $0,80$ a

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
	posteriores al ictus; 226 a los 90 días y 213 participant es evaluados a los 180 días posteriores al ictus.	después del inicio del ictus. Todas las evaluaciones fueron realizadas por un terapeuta ocupacional entrenado.				0,00. FM-37 vs. FM-12 fue de 0,01 a 0,03.
Lin et al. (2021)	267 pacientes con ictus.	Taiwán. Pacientes de datos secundarios de otros estudios.	Fugl-Meyer 37 items como la anterior.	1. Validez estructural 2. Fiabilidad 3. Error de medición 4. Validez de criterio 5. Prueba de hipótesis	1. Dudosa 2. Inadecuada 3. Inadecuada 4. Muy buena 5. Dudosa 6. Dudosa	1. Estructura unidimensional. 2. El ICC para la UE FM-37 fue de 0,94. 4. La r de Pearson para FM-37 fue de 0,93. 5. La correlación r de Pearson con el BI para FM-37 fue de 0,73. La correlación r de Pearson con

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
				6. Sensibilidad		el PASS para FM-37 fue de 0,75.
Forma corta de Fugl-Meyer (12-ítem) (FMA-12) Hsieh et al. (2007)	279 pacientes con ictus.	Países Bajos y Taiwán. Hospital Universitario Nacional de Taiwán.	La versión corta de 12 ítems de FMA tiene 12 ítems (flexión de hombro de 90 a 180°, agarre (pulgar en aducción), pronación/supinación de codo de 90°, flexión/extensión de muñeca de codo de 90°, extensión de codo y elevación de hombro) calificados en una escala de 3 puntos de 0 a 2 puntos con un puntaje máximo de 24 puntos.	1. Fiabilidad 2. Validez de criterio 3. Sensibilidad	1. Inadecuada 2. Muy buena 3. Muy buena	1. La fiabilidad de Rasch para ambas escalas S-FM fue buena (coeficientes $\geq 0,92$). 2. Ambas subescalas de S-FM y FM en 3 puntos temporales presentaron una validez predictiva moderada con la función integral de las actividades de la vida diaria ($r = 0,49$ a $0,59$ para S-FM y $0,48$ a $0,53$ para FM).

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Hsueh et al. (2008)	50 pacientes con ictus.	China. Departamento de Rehabilitación del Hospital Universitario Médico de Kaohsiung.	Fugl-Meyer 12 items como la anterior.	1. Fiabilidad 2. Error de medición 3. Validez de criterio 4. Prueba de hipótesis 5. Sensibilidad	1. Dudosa 2. Dudosa 3. Muy buena 4. Muy buena 5. Muy buena	1. La fiabilidad test-retest fue alta para las subescalas UE y LE, así como para las puntuaciones motoras totales de las 4 escalas motoras (todos los ICC $\geq 0,93$ y límites inferiores del CI del 95 % $\geq 0,89$). 3. Las correlaciones fueron altas ($r = 0,91-0,99$) entre las 4 escalas motoras clínicas y sus subescalas al ingreso y al alta, lo que indica una alta validez concurrente. 4. Las puntuaciones de las 4 escalas motoras y sus subescalas al ingreso mostraron una correlación moderada con las puntuaciones del BI al alta ($r \geq 0,66$), lo que indica una validez predictiva aceptable.

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Fu et al. (2012)	51 pacientes con ictus.	Taiwan. 3 hospitales taiwaneses diferentes.	Fugl-Meyer 12 items como la anterior.	1. Prueba de hipótesis 2. Sensibilidad	1. Muy buena 2. Muy buena	1. Las correlaciones ($r \geq 0,57$) entre Fugl-Meyer Assessment simplificada y las subescalas de función manual de Stroke Impact Scale indican una buena validez concurrente tanto al inicio como después del tratamiento. Wolf Motor Function Test simplificada presentó una validez concurrente de regular a buena ($r \geq 0,39$). Fugl-Meyer Assessment simplificada presentó una buena validez predictiva ($r = 0,68$).

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Chen et al. (2014)	301 pacientes con ictus a los 14 días de la evaluación del ictus; 226 a los 90 días y 213 participant es evaluados a los 180 días posteriores al ictus.	Taiwán. Los datos se obtuvieron de un estudio previo. Los pacientes fueron evaluados con la escala motora de Fugl-Meyer de 50 ítems sujetos a cuatro veces, a los 14, 30, 90 y 180 días después del inicio del ictus. Todas las evaluaciones fueron realizadas por un terapeuta ocupacional entrenado.	Fugl-Meyer 12 ítems como la anterior.	1. Validez estructural 2. Validez de criterio 3. Sensibilidad	1. Dudosa 2. Muy buena 3. Muy buena	1. El análisis paralelo latente y el análisis factorial ordinal indicaron que todos los ítems de FM de la UE de W/H representan un único constructo unidimensional: la capacidad motora de la muñeca y la mano. 2. La relación FM-50 vs. FM-37 fue de 0,61 a 0,80. FM-50 vs. FM-12 fue de 0,80 a 0,00. FM-37 vs. FM-12 fue de 0,01 a 0,03.

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Lin et al. (2021)	267 pacientes con ictus.	Taiwán. Pacientes de datos secundarios de otros estudios.	Fugl-Meyer 12 items como la anterior.	1. Validez estructural 2. Fiabilidad 3. Error de medición 4. Validez de criterio 5. Prueba de hipótesis 6. Sensibilidad	1. Dudosa 2. Inadecuada 3. Inadecuada 4. Muy buena 5. Dudosa 6. Dudosa	1. Estructura unidimensional. 2. El ICC para UE FM-12 fue de 0,93. 4. La r de Pearson para FM-12 fue de 0,93. 5. La correlación r de Pearson con BI para FM-12 fue de 0,73. La correlación r de Pearson con PASS para FM-12 fue de 0,76.
Fugl-Meyer Assessment de miembro superior						

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
acortada (6-items) (FMA-6)						
Amano et al. (2020)	30 pacientes con ictus.	Nishinomiya, Japón. Hospital Universitario de Medicina de Hyogo. Evaluaciones de video de muestra grabadas durante un estudio previo.	La versión acortada de FMA de miembro superior tiene 6 ítems puntuados en una escala de 3 puntos de 0 a 2 puntos con una puntuación máxima de 12 puntos.	1. Consistencia interna 2. Fiabilidad 3. Validez de criterio 4. Prueba de hipótesis 5. Sensibilidad	1. Muy buena 2. Inadecuada 3. Muy buena 4. Muy buena 5. Muy buena	1. El alfa de Cronbach para la FMA abreviada (6 ítems) fue de 0,924 para cada evaluador. 2. La fiabilidad interevaluador de la puntuación total de la FMA abreviada (ICC) fue de 0,994 (intervalo de confianza [CI] del 95 %: 0,988-0,997; $p < 0,001$). 3. La correlación entre la versión abreviada de la FMA y la FMA-UE fue de $r = 0,97$. 4. La correlación de la versión abreviada de la FMA con ARAT fue de 0,94; con BBT fue de 0,92; y con MAL fue de 0,91 para ambas subescalas (cantidad de uso y calidad del movimiento).

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Fugl-Meyer de aprendizaje automático (10 ítems) (FMA-ML-10)						
Lin et al. (2021)	267 pacientes con ictus.	Taiwán. Pacientes de datos secundarios de otros estudios.	En este artículo, los autores desarrollan una versión de la evaluación Fugl-Meyer basada en aprendizaje automático (FM-ML), que consta de 10 ítems de la escala original, puntuados con una escala de 3 puntos, que van de 0 a 2.	1. Validez estructural 2. Fiabilidad 3. Error de medición 4. Validez de criterio 5. Prueba de hipótesis 6. Sensibilidad	1. Dudosa 2. Inadecuada 3. Inadecuada 4. Muy buena 5. Dudosa 6. Dudosa	1. Estructura unidimensional. 2. El ICC para UE FM-ML fue de 0,92. 4. La r de Pearson para FM-ML fue de 0,99. 5. La correlación r de Pearson con BI para FM-ML fue de 0,67. La correlación r de Pearson con PASS para FM-ML fue de 0,68.
Traducción y adaptación del manual de administración de la FMA al portugués brasileño						

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Michaelsen et al. (2011)	18 pacientes con ictus.	Brasil. Diez pacientes participaron en la primera parte del estudio (traducción y adaptación del manual y evaluación de la fiabilidad interevaluadores), y ocho pacientes participaron en la segunda parte (evaluación de la fiabilidad interevaluadores de la prueba).	La versión francocanadiense del manual de la FMA fue traducida al brasileño por un fisioterapeuta.	1. Fiabilidad	1. Inadecuada	1. Parte 1: El ICC entre los dos evaluadores para la puntuación total y la puntuación de las extremidades superiores fue de 0,99 ($p < 0,001$) para ambos y de 0,88 ($p = 0,006$) para las extremidades inferiores. Parte 2: El ICC entre los dos examinadores para la puntuación total de la función motora fue de 0,98 ($p < 0,001$). La comparación entre los ICC de las partes 1 y 2 para los ítems de sensibilidad táctil, sentido del movimiento y ROM muestra que la fiabilidad se volvió excelente tras la adaptación del manual. La fiabilidad de la evaluación del dolor articular para las extremidades superiores fue moderada, y para las extremidades inferiores fue baja.

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Subescala de miembro superior de Fugl-Meyer Assessment (FMA-UE)						
De Weerd et al. (1985)	53 pacientes con ictus.	Nottingham, Reino Unido.	La Evaluación Fugl-Meyer de las Extremidades Superiores (FMA-UE) examina la actividad refleja y los movimientos voluntarios, tanto internos como externos, e independientemente de las sinergias. Incluye 33 ítems divididos en 4 subescalas: hombro/codo (A, 18 ítems), muñeca (B, 5 ítems), mano (C, 7 ítems) y coordinación/velocidad (D, 3 ítems). Cada ítem se puntúa en una escala ordinal de 3 puntos que va de 0 a 2. La puntuación total oscila entre 0 y	1. Prueba de hipótesis 2. Sensibilidad	1. Muy buena 2. Muy buena	1. La correlación de Spearman entre FM y ARAT fue de 0,91 a las dos semanas y de 0,94 a las ocho semanas. Estos coeficientes de correlación son significativos al nivel de significancia de 0,01.

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento			Propiedades de medición	Puntuación COSMIN	Valores de medición
			66.					
Van der Lee et al. (2001)	22 pacientes con ictus crónico.	Países Bajos. En este ensayo clínico aleatorizado se admitieron personas con ictus crónico que cumplían los criterios. Se obtuvo el consentimiento informado y se realizaron dos mediciones basales antes de la intervención.	Fugl-Meyer	Assessment	de la	1. Fiabilidad 2. Prueba de hipótesis 3. Sensibilidad	1. Inadecuada 2. Muy buena 3. Muy buena	1. La diferencia entre dos mediciones debe ser del 10,0 % para la FMA. 2. Considerando la mejora media (1,7 puntos en ARAT) durante la intervención de referencia, se consideró más válido el uso del período basal (con una mejora media de 0,3 puntos en ARAT) para definir la estabilidad.
Rabadi et al. (2006)	104 pacientes con ictus agudo.	Nueva York, Estados Unidos. Hospital de Rehabilitación Burke.	Fugl-Meyer	Assessment	de la	1. Prueba de hipótesis 2. Sensibilidad	1. Muy buena 2. Muy buena	1. Las 2 escalas motoras de las extremidades superiores (ARAT, puntuación motora FMA) se correlacionaron altamente entre sí, tanto al ingreso ($r = 0,77$, $p < 0,001$) como al alta ($r = 0,87$, $p < 0,001$)

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Woodbury et al. (2007)	512 pacientes con ictus subagudo.	Aprobado por la Junta de Revisión Institucional de la Universidad de Florida. 459 sujetos participaron en el Estudio de Ictus de Kansas City. Los investigadores de este estudio utilizaron un análisis secundario de datos agrupados de dos conjuntos de datos existentes: un ensayo clínico aleatorizado de ejercicio terapéutico y un estudio longitudinal de cohorte sobre la recuperación del ictus.	Fugl-Meyer Assessment de miembro superior como la anterior.	1. Validez estructural 2. Consistencia interna	1. Muy buena 2. Muy buena	0,001). 1. El PCA retuvo 4 factores con valores propios superiores a 1,0. Un solo componente explicó el 68 % de la varianza de los datos. Otros tres factores explicaron un 5 %, un 4 % y un 3 % adicionales de la varianza. 2. El índice de fiabilidad de Pearson, análogo al coeficiente, fue de 0,96.

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Hsieh et al. (2009)	57 pacientes con ictus.		Fugl-Meyer Assessment de miembro superior como la anterior.	1. Prueba de hipótesis 2. Sensibilidad	1. Muy buena 2. Muy buena	1. Las correlaciones entre la FMA y otras medidas de resultado antes del tratamiento fueron relativamente altas (Spearman: 0,71-0,76). Después del tratamiento, la FMA presentó correlaciones moderadas a buenas con las demás medidas de resultado (Spearman: 0,51-0,74).
Lin et al. (2009)	53 pacientes con ictus iniciaron el estudio. 11 de ellos sufrieron ictus recurrente	Taiwán. Personas con ictus que ingresaron consecutivamente en el Departamento de Neurología del Hospital Universitario Médico de Kaohsiung entre el 1 de	Fugl-Meyer Assessment de miembro superior como la anterior.	1. Fiabilidad 2. Error de medición 3. Prueba de hipótesis 4. Sensibilidad	1. 2. Inadecuada 3. Muy buena 4. Muy buena	1. La fiabilidad interevaluador para las 4 medidas fue alta (ICC $\geq 0,92$, CI del 95 %; FM = 0,96). La fiabilidad test-retest de las 4 medidas fue alta (ICC $\geq 0,97$, CI del 95 %; FM = 0,99). 3. La correlación de FM con las otras 3 medidas fue alta (p de Spearman = 0,85-0,86).

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
	s y 6 pacientes se perdieron en el seguimient o. Un total de 35 pacientes con ictus completaro n todas las evaluacion es.	septiembre de 2006 y el 31 de agosto de 2007 y que cumplieron con los criterios de inclusión y exclusión.				
Sullivan et al. (2011)	15 pacientes con ictus.	Los Ángeles, Estados Unidos. Se realizaron 17 evaluaciones en total, ya que dos terapeutas diferentes midieron a	Fugl-Meyer Assessment de miembro superior como la anterior.	1. Fiabilidad	1. Inadecuada	1. La concordancia interevaluador para la evaluación de la respuesta motora (ER) fue alta para las evaluaciones motora (motora total: 0,99; UE: 0,95; LE: 0,99) y sensorial (sensible total: 0,96).

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
		dos sujetos. En este artículo, se compara a un evaluador experto con evaluadores terapéuticos entrenados.				La concordancia interevaluador entre los evaluadores expertos y terapeutas fue alta para la puntuación motora total (ICC: 0,98; CI del 95 %: 0,93-0,99) y la subpuntuación motora de UE (ICC: 0,99; CI del 95 %: 0,97-1,0), y se mantuvo entre moderada y alta para la subpuntuación motora de LE (ICC: 0,91; CI del 95 %: 0,69-0,97). La concordancia interevaluador fue alta para la puntuación sensorial total (ICC: 0,93; CI del 95 %: 0,83-0,98).

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento			Propiedades de medición	Puntuación COSMIN	Valores de medición
Wei et al. (2011)	27 pacientes con ictus.	Hong Kong. Universidad Politécnica de Hong Kong.	Fugl-Meyer	Assessment	de miembro superior como la anterior.	1. Validez de criterio 2. Prueba de hipótesis 3. Sensibilidad	1. Muy buena 2. Muy buena 3. Muy buena	1. La correlación entre FMA y FMA (SE) fue de 0,91 y 0,94; y con FMA (WH) fue de 0,92 y 0,94; puntuaciones en el preentrenamiento y postentrenamiento respectivamente. La correlación entre FMA (SE) y FMA (WH) fue de 0,69 y 0,74. 2. La correlación de FMA fue alta con MSS y ARAT (0,93 para ambos en el pretratamiento y 0,91 y 0,92 para el postratamiento respectivamente). Con MSS (SE) fue de 0,89 en el pretratamiento y 0,86 en el postratamiento; con MMSS (WH) la correlación fue de 0,9 tanto para el preentrenamiento como para el postentrenamiento. Para MAS (E), la correlación no fue

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
						significativa en el preentrenamiento, pero fue de 0,62 en el postentrenamiento. Con MAS (W), la correlación fue de 0,42 en el preentrenamiento y no significativa para el postentrenamiento. Con FIM, la correlación fue de 0,49 antes del tratamiento y de 0,56 después del tratamiento.
Kim et al. (2012)	50 pacientes con ictus para estudio de Fiabilidad y validez, y 16 pacientes con ictus para	Tres hospitales de rehabilitación en Corea del Sur.	Fugl-Meyer Assessment de miembro superior como la anterior.	1. Fiabilidad 2. Error de medición 3. Validez de criterio 4. Prueba de hipótesis 5. Sensibilidad	1. Dudosa 2. Dudosa 3. Muy buena 4. Muy buena 5. Muy buena	1. Para la fiabilidad interevaluador de la función motora, el ICC osciló entre 0,930 y 1,000. Para la fiabilidad test-retest con un intervalo de dos semanas de la función motora, el ICC fue de 0,834 a 0,972. 3. La FMA-UE comparada con la FMA-LE fue de 0,723 ($p<0,05$). 4. La relación entre la función

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
	estudio de Sensibilidad					motora de las extremidades superiores de la FMA y la función manual de Jebsen-Taylor fue de 0,757 ($p < 0,05$), y la relación entre la función motora de las extremidades inferiores de la FMA y la MAS fue de 0,725 ($p < 0,05$).
Kim et al. (2016)	41 pacientes con ictus.	Seúl, Corea del Sur. Hospital Bundang de la Universidad Nacional de Seúl.	Fugl-Meyer Assessment de miembro superior como la anterior.	1. Validez de criterio 2. Prueba de hipótesis 3. Sensibilidad	1. Muy buena 2. Muy buena 3. Muy buena	1. La correlación entre los 13 ítems de la FMA con Kinect y la FMA real fue de 0,873 ($p < 0,0001$). 2. Correlación entre el logaritmo (puntuación espasmódica) y el estadio del brazo de Brunnstrom (3 a 6) en la extremidad superior hemipléjica ($n = 27$) (coeficiente de correlación de Spearman = -0,387 ($p = 0,046$)).

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Amano et al. (2020)	30 pacientes con ictus.	Nishinomiya, Japón. Hospital Universitario de Medicina de Hyogo. Evaluaciones de video de muestra grabadas durante un estudio previo.	Fugl-Meyer Assessment de miembro superior como la anterior.	1. Consistencia interna 2. Fiabilidad 3. Validez de criterio 4. Prueba de hipótesis 5. Sensibilidad	1. Muy buena 2. Inadecuada 3. Muy buena 4. Muy buena 5. Muy buena	1. El alfa de Cronbach para la FMA abreviada (6 ítems) fue de 0,924 para cada evaluador. 2. La fiabilidad interevaluador de la puntuación total de la FMA abreviada (ICC) fue de 0,994 (CI del 95 %: 0,988-0,997; $p < 0,001$). 3. La correlación entre la versión abreviada de la FMA y la FMA-UE fue de $r = 0,97$. 4. La correlación de la versión abreviada de la FMA con ARAT fue de 0,94; con BBT fue de 0,92; y con MAL fue de 0,91 para ambas subescalas (cantidad de uso y calidad del movimiento).

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Hernández et al. (2019)	60 pacientes con ictus.	Bogotá, Colombia. 60 pacientes consecutivos, incluyendo un período de 17 meses en el Hospital Militar Central de Colombia.	Fugl-Meyer Assessment de miembro superior como la anterior.	1. Fiabilidad	1. Inadecuada	1. Para la fiabilidad intraevaluador, la PA entre las ocasiones de prueba 1 y 2 dentro de cada evaluador fue del 79 % anterior para todos los ítems evaluados. Para la fiabilidad interevaluador, la PA fue del 90 % anterior para todos los ítems entre los evaluadores.
Tauchi et al. (2022)	268 pacientes con ictus.	Japón. Pacientes de 22 hospitales japoneses.	Fugl-Meyer Assessment de miembro superior como la anterior.	1. Validez estructural	1. Muy buena	1. Como resultado del análisis de un factor utilizando CFA.

Versión Danesa de la subescala de miembro superior de Fugl-Meyer Assessment (FMA-

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
UE-Dan) Lundquist et al. (2017)	50 pacientes con ictus.	Dinamarca. Hospital Regional de Skive. Traducción transcultural de la subescala de miembro superior de Fugl-Meyer Assessment al danés y evaluación de sus propiedades psicométricas, utilizando una muestra de 50 pacientes.	Los autores realizaron una traducción de la evaluación Fugl-Meyer de extremidades superiores como la anterior a una versión danesa.	1. Validez transcultural 2. Fiabilidad 3. Prueba de hipótesis 4. Sensibilidad	1. Inadecuada 2. Dudosa 3. Dudosa 4. Dudosa	2. La fiabilidad interevaluador fue alta (ICC: 0,95; CI del 95 %, 0,93-0,98). 3. La correlación de rangos de Spearman entre la FMA-UE y la MAS-UE fue de 0,95 ($p < 0,001$) al inicio y de 0,94 ($p < 0,001$) en el seguimiento.

Versión Italiana de la subescala de miembro superior de Fugl-Meyer Assessment (FMA-

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
UE-Ita)						
Hochleitner et al. (2023)	60 pacientes con ictus subagudo.	Florenia, Italia. Unidad de Rehabilitación Neurológica de la Fundación Don Carlo Gnocchi, Instituto Científico.	Los autores realizaron una traducción de Fugl-Meyer Assessment de extremidades superiores como la anterior a una versión italiana.	1. Fiabilidad	1. Inadecuada	1. Fiabilidad interevaluador a nivel de ítem: el porcentaje de acuerdo fue satisfactorio o excelente para todos los ítems (PA >70%). Fiabilidad intraevaluador a nivel de ítem: el porcentaje de acuerdo fue satisfactorio o excelente (PA >70%).
Versión Rumana de la subescala de miembro superior de Fugl-Meyer Assessment (FMA-UE-Rom)						
Roman et al. (2020)	64 pacientes con ictus.	Rumanía. Hospital Clínico de Psiquiatría y Neurología de Braşov,	Los autores hicieron una traducción de Fugl-Meyer Assessment de miembro superior como la anterior, a una versión	1. Validez estructural 2. Consistencia interna	1. Inadecuada 2. Muy buena	1. Tras realizar el análisis factorial exploratorio, se extrajo un único factor. 2. Alfa de Cronbach = 0,981.

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Fugl-Meyer Assessment de miembro superior (13-Items) usando una cámara con sensor de profundidad (Kinect) (FMA-13-Kinect)	41 pacientes con ictus.	Rumanía.	rumana.	3. Validez transcultural	3. Inadecuada	4. CCI de las medidas promedio = 0,992.
				4. Fiabilidad	4. Inadecuada	5. La correlación con FIM fue de 0,789 (p < 0,001), con MRS fue de -0,787 (p < 0,001) y con SRM fue de 1,117.
				5. Prueba de hipótesis	5. Muy buena	
				6. Sensibilidad	6. Muy buena	
Kim et al. (2016)	41	Seúl, Corea del Sur.	En este estudio, los investigadores comparan el método tradicional de observación para evaluar la FMA y el registro de datos de	1. Validez de criterio	1. Muy buena	1. La correlación entre los ítems de la FMA 13 con Kinect y los ítems de la FMA 13 real fue de 0,873 (p < 0,0001).
		Instituto del Hospital Bundang de la Universidad Nacional de Seúl.		2. Prueba de hipótesis	2. Muy buena	2. Correlación entre el logaritmo
				3. Sensibilidad	3. Muy buena	

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
			movimiento de la cámara de profundidad Kinect para evaluar la FMA. - FMA-UE 33 ítems - FMA-UE 13 ítems (sinergia flexora, sinergia extensora, movimiento voluntario que combina sinergia dinámica flexora y extensora, y movimiento voluntario con poca o ninguna dependencia de la sinergia).		buena	(puntuación de sacudidas) y el estadio del brazo de Brunnstrom (3 a 6) en la extremidad superior hemipléjica (n = 27) (coeficiente de correlación de Spearman = - 0,387 (p = 0,046)).
Evaluación remota de la subescala de miembro superior de Fugl-Meyer Assessment (FMA-Remote)						
Amano et al. (2018)b	30 pacientes	Japón. Hospital de la Facultad de Medicina	Fugl-Meyer Assessment de miembro superior como la	1. Fiabilidad	1. Dudosa	1. ICC = 0.998

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
	con ictus.	de Hyogo.	anterior, evaluado a través de una evaluación remota			
Subescalas de hombro y codo de Fugl-Meyer Assessment (FMA-UE-SE)						
Wei et al. (2011)	27 pacientes con ictus.	Hong Kong. Universidad Politécnica de Hong Kong.	La Evaluación Fugl-Meyer de Hombro y Codo de la Extremidad Superior (SE-UE-FM) incluye los ítems más proximales de la UE-FM. Cada ítem se califica en una escala ordinal de 3 puntos que va de 0 a 2.	1. Validez de criterio 2. Prueba de hipótesis 3. Sensibilidad	1. Muy buena 2. Muy buena 3. Muy buena	1. La correlación entre FMA y FMA (SE) fue de 0,91 y 0,94; y con FMA (WH) fue de 0,92 y 0,94; puntuaciones preentrenamiento y postentrenamiento, respectivamente. La correlación entre FMA (SE) y FMA (WH) fue de 0,69 y 0,74. 2. La correlación de FMA con MSS fue de 0,79; con MSS (SE) fue de 0,79; con MMSS (WH) la correlación fue de 0,75. Con ARAT fue de 0,85. Para MAS (E)

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
						la correlación fue de 0,54. Con MAS (W) la correlación fue de 0,4; y con FIM, la correlación fue de 0,52.
See et al. (2013)	31 pacientes con ictus.	Datos de pacientes de dos ensayos clínicos de terapia motora robótica de brazos en accidentes cerebrovasculares de Estados Unidos, Canadá y Alemania.	Subescala de hombro-codo de la extremidad superior de la evaluación de Fugl-Meyer como la anterior.	1. Fiabilidad 2. Prueba de hipótesis 3. Sensibilidad	1. Dudosa 2. Muy buena 3. Muy buena	1. Para la fiabilidad intraevaluador, el ICC fue de 0,99 para la puntuación total de FMA, la subpuntuación proximal de FMA y la subpuntuación de FMA-p/h. Para la fiabilidad interevaluador, el ICC fue de 0,99 para la puntuación total de FMA y de 0,98 para las subpuntuaciones proximal y p/h de FMA. 2. La correlación de la puntuación de FMA-SE con ARAT fue de 0,8; con fuerza de pinza fue de 0,85; con fuerza de agarre fue de 0,73; con BBT fue

Programa de Doctorado en Ciencias de la Salud

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
						la correlación fue de 0,56. Con MAS (W) la correlación no fue significativa; y con FIM, la correlación fue de 0,49.
Page et al. (2012)	29 pacientes con ictus.	Medio Oeste de Estados Unidos.	Fugl-Meyer Assessment de la extremidad superior de la muñeca y la mano como la anterior.	1. Consistencia interna 2. Fiabilidad 3. Prueba de hipótesis 4. Sensibilidad	1. Muy buena 2. Adequate 3. Muy buena 4. Muy buena	1. El alfa de Cronbach para FMA de muñeca y mano fue de 0,90 para la prueba previa 1 y de 0,88 para la prueba previa 2. 2. El ICC para FMA de muñeca y mano fue de 0,97. 3. La puntuación de correlación entre FMA de muñeca y mano y el ARAT fue de 0,72 ($p<0,001$).

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
See et al. (2013)	31 pacientes con ictus.	Datos de pacientes de dos ensayos clínicos de terapia motora robótica de brazos en accidentes cerebrovasculares de Estados Unidos, Canadá y Alemania.	Fugl-Meyer Assessment de la extremidad superior de la muñeca y la mano como la anterior.	1. Fiabilidad 2. Prueba de hipótesis 3. Sensibilidad	1. Dudosa 2. Muy buena 3. Muy buena	1. Para la fiabilidad intraevaluador, el ICC fue de 0,99 para la puntuación total de FMA, la subpuntuación de FMA-proximal y la subpuntuación de FMA de muñeca y mano. Para la fiabilidad interevaluador, el ICC fue de 0,99 para la puntuación total de FMA y de 0,98 para las subpuntuaciones de FMA-proximal y FMA de muñeca y mano. 2. La correlación de la puntuación de FMA de muñeca y mano con ARAT fue de 0,89; con fuerza de pinza fue de 0,85; con fuerza de agarre fue de 0,73; con BBT fue de 0,88; con NHPT fue de 0,8 y con SIS fue de 0,88.

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Page et al. (2015)	32 pacientes con ictus.	Medio Oeste de Estados Unidos. Centro Nacional de Medicina Complementaria y Alternativa.	Fugl-Meyer Assessment de la extremidad superior de la muñeca y la mano como la anterior.	1. Consistencia interna 2. Fiabilidad 3. Prueba de hipótesis 4. Sensibilidad	1. Muy buena 2. Dudosa 3. Muy buena 4. Muy buena	1. El alfa de Cronbach para la FMA de muñeca y mano fue de 0,82 en la prueba previa 1. 2. El ICC para FMA de muñeca y mano fue de 0,95. 3. El r de Spearman para la correlación entre ARAT y la FMA de muñeca y mano fue de 0,74 (p<0,001) en la prueba previa 1.
Persch et al. (2015)	150 pacientes con ictus.	Pacientes del análisis secundario de datos obtenidos durante ensayos controlados aleatorizados ambulatorios aprobados por el Programa de Protección de la Investigación Humana de la Universidad Estatal de Ohio. Los	Fugl-Meyer Assessment de la extremidad superior de la muñeca y la mano como la anterior.	1. Validez estructural 2. Error de medición	1. Adequate 2. Dudosa	1. El EFA resultó en un análisis de 12 ítems con valores entre 0,477 y 0,861.

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
		participantes fueron reclutados para los ensayos en todo el Medio Oeste de Estados Unidos.				
		Subescala sensitiva de Fugl-Meyer Assessment (FMA-S)				
Lin et al. (2004)	176 pacientes con ictus.	Taiwán. La muestra del estudio se reclutó del registro del Estudio de Calidad de Vida después de un ictus en Taiwán, del 1 de diciembre de 1999 al 7 de diciembre de 2001. El Estudio de Calidad de Vida después de un ictus es un estudio de cohorte	La subescala sensitiva de Fugl-Meyer Assessment (FMA-S) consta de 12 ítems de tres puntos, con una puntuación de 0 a 2. Cuatro ítems corresponden al tacto ligero y ocho al sentido de la posición. La puntuación total oscila entre 0 y 24.	1. Consistencia interna 2. Fiabilidad 3. Validez de criterio 4. Prueba de hipótesis 5. Sensibilidad	1. Muy buena 2. Inadecuada 3. Muy buena 4. Muy buena 5. Muy buena	1. Los alfa de Cronbach de la FMA-S en cuatro momentos tras el ictus oscilaron entre 0,94 y 0,98, lo que indica una excelente consistencia interna. 2. La concordancia interevaluador de la puntuación total de la FMA-S fue excelente, con un ICC (CI del 95 %) de 0,93 (0,85-0,96). 3. Las puntuaciones de la FMA-S mostraron una correlación

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
		prospectivo de pacientes con ictus ingresados en el Hospital Universitario Nacional de Taiwán.				entre débil y moderada con las puntuaciones de la FMA-Motora (rho de Spearman entre 0,31 y 0,44). 4. Las puntuaciones de la FMA-S mostraron una correlación entre débil y moderada con el BI (rho de Spearman entre 0,38 y 0,53).

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Estudio (autor y año)	Población	Características	Descripción del instrumento	Propiedades de medición	Puntuación COSMIN	Valores de medición
Sullivan et al. (2011)	15 pacientes con ictus.	Los Ángeles, California. Se realizaron 17 evaluaciones en total, ya que dos terapeutas diferentes midieron a dos sujetos. En este artículo, se compara a un evaluador experto con evaluadores terapéuticos capacitados.	Subescala sensitiva de Fugl-Meyer Assessment como la anterior.	1. Fiabilidad	1. Inadecuada	1. La concordancia interevaluador para la evaluación de la respuesta motora (ER) fue alta para las evaluaciones motora (motora total: 0,99; UE: 0,95; LE: 0,99) y sensorial (sensible total: 0,96). La concordancia interevaluador entre los evaluadores expertos y terapeutas fue alta para la puntuación motora total (ICC: 0,98; CI del 95 %: 0,93-0,99) y la subpuntuación motora de UE (ICC: 0,99; CI del 95 %: 0,97-1,0), y se mantuvo entre moderada y alta para la subpuntuación motora de LE (ICC: 0,91; CI del 95 %: 0,69-0,97). La concordancia interevaluador fue alta para la puntuación sensorial total (ICC: 0,93; CI del 95 %: 0,83-0,98).

Abreviaturas: (ARAT) Action Research Arm Test; (BBS) Berg Balance Scale; (BBT) Box & Block Test; (BI) Barthel Index; (CI) Confidence Interval; (CFA) Confirmatory Factor Analysis; (EFA) Exploratory Factor Analysis; (ER) Expert Rater; (FIM) Functional Independence Measure; (FM)/(FMA) Fugl-Meyer Assessment; (FM-ML) Fugl-Meyer Machine-Learning; (FMA-Col) Fugl-Meyer Assessment Colombian; (FMA-Dan) Fugl-Meyer Assessment Danish; (FMA-Ita) Fugl-Meyer Assessment Italian; (FMA-Jap) Fugl-Meyer Assessment Japanese; (FMA-Kor) Fugl-Meyer Assessment Korean; (FMA-LE) Fugl-Meyer Assessment Lower Extremity; (FMA-LE-Ita) Fugl-Meyer Assessment Lower Extremity Italian; (FMA-M) Fugl-Meyer Assessment Motor; (FMA-S) Fugl-Meyer Assessment Sensory; (FMA-SE) Fugl-Meyer Assessment Shoulder/Elbow; (FMA-UE) Fugl-Meyer Assessment Upper Extremity; (FMA-UE-Dan) Fugl-Meyer Assessment Upper Extremity Danish; (FMA-UE-Ita) Fugl-Meyer Assessment Upper Extremity Italian; (FMA-UE-Rom) Fugl-Meyer Assessment Upper Extremity Romain; (FMA-UE-SE)/(FMA-SE) Fugl-Meyer Assessment Shoulder-Elbow; (FMA-UE-w/h)/(FMA-W/H) Fugl-Meyer Assessment Upper Extremity Wrist/Hand; (FMA-Urd) Fugl-Meyer Assessment Urdu; (ICC) Intraclass correlation coefficient; (LE) Lower Extremity; (LE-SFM) Lower-Extremity Short-Fugl-Meyer; (MAL) Motor Activity Log; (MAS) Motor Assessment Scale; (MAS-UE) Motor Assessment Scale-Upper Extremity; (MRS) Modified Rankin Scale; (MSS) Motor Status Scale; (NHPT) Nine-Hole Peg Test; (NIHSS) National Institutes of Health Stroke Scale; (PA) Percentage of Agreement; (PASS) Postural Assessment Scale for Pacientes con ictus; (PCA) Principal Components Analysis; (ROM) Range of Movement; (S-FM) Short-Fugl-Meyer; (SIS) Stroke Impact Scale; (UE) Upper Extremity; (UE-FM-ML) Upper-Extremity Fugl-Meyer Machine-Learning; (UE-SFM) Upper-Extremity Short-Fugl-Meyer.

Fuente: elaboración propia

3.3 Propiedades psicométricas

Las principales propiedades evaluadas fueron la sensibilidad ($n=39$)^{14, 22-26, 30-36, 38-40, 42, 43, 45-47, 49, 51, 53, 54}, prueba de hipótesis ($n=34$)^{14, 22-25, 30-33, 35, 36, 38-40, 42, 43, 45-47, 49, 51, 53, 54}, fiabilidad ($n=33$)^{14, 21, 23-25, 30-34, 36, 37, 39, 43, 44, 46, 48-54} y validez de criterio ($n=20$)^{14, 23, 24, 26, 34, 36, 45-47}. Los resultados de la lista de verificación COSMIN de 4 puntos de las versiones de FMA para la medición de la función sensoriomotora de las extremidades superiores en personas con ictus se muestran en la Tabla 3. En ella se observa que las dos versiones con mayor número de propiedades psicométricas evaluadas y mayor número de estudios son la versión completa de la FMA y la FMA-UE. Entre los resultados de calidad metodológica, se observa que la versión con mayor número de calificaciones como "muy buena" es la FMA-UE ($n=26$), seguida de la versión completa de la FMA ($n=15$); mientras que la FMA-UE es también la versión con mayor número de puntuaciones "inadecuadas" ($n=6$). Esto indica que es la versión con mejor calidad metodológica, ya que se ha estudiado en múltiples artículos, pero también se ha analizado en estudios de menor rigor. Los estudios con mayor calidad metodológica, es decir, aquellos con mayor número de puntuaciones "muy buenas", fueron Lin et al. (2004)²³ y Amano et al. (2020)³⁶, con 4 cada una. En cuanto a la calidad de los resultados, las versiones con mayor puntuación positiva fueron FMA-UE ($n=10$) y FMA versión completa ($n=9$); mientras que las versiones con mayor número de puntuaciones negativas («-») fueron FMA versión completa y FMA-12, con 15 «-» cada una. Sin embargo, FMA-UE, a pesar de presentar más «+» y menos «-», ha sido la versión con mayor número de puntuaciones indeterminadas («?») ($n=23$). Los estudios con mejor calidad de resultados fueron Amano et al. (2020)³⁶, Lin et al. (2021)¹⁴, Roman et al. (2020)⁵¹, Page et al. (2012)⁵³ y Page et al. (2015)⁵⁴. La Tabla 4 muestra la calidad de la evidencia analizada en los 36 artículos de las 22 versiones de FMA, analizadas mediante la metodología COSMIN. Las versiones con mayor nivel de evidencia fueron FMA-UE y FMA-S, con 4 puntuaciones «Altas» cada una; mientras que las versiones con menor nivel de evidencia fueron FMA-Rom, con 5 puntuaciones «Muy Bajas», y las versiones FMA-Urd y FMA-UE-Dan, con 4 puntuaciones «Muy Bajas» cada una.

La calidad metodológica de los estudios para cada una de las propiedades psicométricas evaluadas no fue la esperada, resultando en "inadecuada" o "dudosa" en prácticamente todas las versiones debido a la falta de evaluación de contenido, y en el caso de la validez de contenido, a la relevancia del ítem en la población de estudio. En el caso de la validez estructural, debido a que no se calculó para cada dimensión por separado o a la falta de descripción del porcentaje de gestión de valores faltantes. En la propiedad psicométrica de validez transcultural, la baja calidad metodológica se debió al pequeño tamaño de la muestra y a la falta de comparación entre los diferentes grupos tras el análisis transcultural. En el caso de la fiabilidad y la validez de criterio, esta baja calidad metodológica se debió a la falta de descripción de los porcentajes y la gestión de los valores faltantes, y a que, en la fase posterior al ictus, la mayoría de los estudios se realizaron en personas en fase aguda post-ictus que no

presentaban estabilidad neurológica en el momento de la evaluación. Finalmente, la baja calidad metodológica en la prueba de hipótesis se debe a que las hipótesis no fueron formuladas a priori en los estudios, por lo que no fue posible conocer lo que se esperaba encontrar en dichos estudios.

Quince versiones (FMA-Col²⁷, FMA-Dan²⁸, FMA-Ita²⁹, FMA-Jap³⁰, FMA-Kor³¹, FMA-Urd³², FMA-Rom³³, FMA-6³⁶, FMA-ML-10¹⁴, la Traducción y adaptación del manual de administración de FMA al portugués brasileño³⁷, FMA-UE-Dan⁴⁹, FMA-UE-Ita⁵⁰, FMA-UE-Rom⁵¹, FMA-13-Kinect⁴⁷ y el FMA-Remote⁵²) solo se evaluaron en un estudio. Por lo tanto, se requieren estudios adicionales para evaluar las propiedades psicométricas de estas versiones ya traducidas y comprender mejor estas variantes de FMA, de modo que puedan utilizarse con mayor rigor en el ámbito clínico y de investigación, y para aprovechar sus ventajas en comparación con otras versiones.

Sólo hubo dos estudios de la misma versión de FMA, la FMA-UE-w/h, en los que la calidad metodológica fue calificada como "adecuada", para la validez estructural⁵⁵, gracias a un análisis factorial desarrollado, un análisis de Rasch y un buen tamaño de muestra. El otro estudio⁵³ se calificó como "adecuada" en cuanto a fiabilidad debido a la estabilidad de la muestra, el intervalo de tiempo apropiado y las condiciones de prueba similares.

La calidad de la evidencia fue "baja" o "muy baja" para la mayoría de las propiedades de medición analizadas en todas las versiones incluidas en esta revisión sistemática. Teniendo en cuenta que la validez de contenido, la validez estructural y la consistencia interna se consideran las propiedades de medición más importantes^{11, 18}. La validez de contenido estuvo presente en sólo cinco estudios de cinco versiones diferentes: Versión completa de 50 ítems de FMA²², FMA-Col²⁷, FMA-Ita²⁹, FMA-Rom³³ y FMA-Urd³². Para las cinco versiones, Versión completa de 50 ítems de FMA²², FMA-Col²⁷, FMA-Ita²⁹, FMA-Rom³³ y FMA-Urd³², la calidad de la evidencia se calificó como "dudosa" debido a la probabilidad de que la propiedad de medición real se aleje de la estimación, aun cuando existe la posibilidad de que sea sustancialmente diferente. Además, en cuatro versiones que evalúan la validez de contenido, la calidad de los resultados se calificó como "-"^{22, 27, 29, 32}, y el otro como "?"³³, indicando que no se ha evaluado la pertinencia de los ítems, la comprensibilidad y la exhaustividad de los ítems de las diferentes versiones.

La validez estructural se evaluó en nueve estudios: Versión completa de 50 ítems de FMA^{14, 23, 26}, FMA-Urd³², FMA-37^{14, 26}, FMA-12^{14, 26}, FMA-ML-10¹⁴, FMA-UE^{15, 41}, FMA-UE-Rom⁵¹, FMA-UE-w/h⁵⁵ y FMA-S²³. Sólo una única versión FMA-UE^{15, 41} obtuvo una calidad de evidencia calificada como "muy buena". Un estudio se calificó como "adecuado" (FMA-UE-w/h⁵⁵). El resto de las versiones y estudios fueron calificados como "dudosos" (la mayoría) e "inadecuados" (FMA-Urd³² y FMA-UE-Rom⁵¹).

Finalmente, para la consistencia interna, ocho estudios evaluaron esta propiedad psicométrica: Versión completa de 50 ítems de FMA²³, FMA-Jap³⁰, FMA-Urd³², FMA-6³⁶, FMA-UE^{36, 41}, FMA-UE-Rom⁵¹, FMA-UE-w/h^{53, 54} y FMA-S²³. Todas las versiones y estudios que evalúan esta propiedad psicométrica han obtenido una calificación de «muy buena»; versión completa de 50 ítems de FMA.²³, FMA-

Jap³⁰, FMA-Urd³², FMA-6³⁶, FMA-UE^{36, 41}, FMA-UE-Rom⁵¹, FMA-UE-w/h^{53, 54} y FMA-S²³. Y todos mostraron una calificación de “+”, lo que indica que el alfa de Cronbach es $\geq 0,70$, excepto para FMA-Urd³².

Tabla 3. Calidad metodológica de los estudios y calidad de los resultados reportados por propiedad de medición, instrumento y estudio utilizando la metodología COSMIN y GRADE.

Instrumento	Artículos	Validez de contenido		Validez estructural		Consistencia interna		Validez transcultural/Invarianza de medición		Error de medición		Fiabilidad		Validez criterio		Prueba de hipótesis		Sensibilidad	
		M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q
Fugl-Meyer Assessment (Versión completa, 50-Items) (FMA)	Sanford et al. (1993)									D	?	D	+						
	Malouin et al. (1994)	D	-													V	?	V	?
	Lin et al. (2004)					V	+					I	+	V	-	V	?	V	?
	Hsueh et al. (2008)									D	-	D	+	V	+	V	?	V	?
	See et al. (2013)											D	+			V	+	V	+
	Chen et al. (2014)			D	?									V	-			V	-

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Artículos	Validez de contenido		Validez estructural		Consistencia interna		Validez transcultural/Invarianza de medición		Error de medición		Fiabilidad		Validez de criterio		Prueba de hipótesis		Sensibilidad	
		M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q
	Lin et al. (2021)			D	?					I	?	I	+			D	?	D	?
Versión Española-Colombiana de Fugl-Meyer Assessment (FMA-Col)	Barbosa et al. (2019)	D	-					I	-										
Versión Danesa de Fugl-Meyer Assessment (FMA-Dan)	Busk et al. (2022)							I	-										
Versión Italiana de Fugl-Meyer Assessment (FMA-Ita)	Cecchi et al. (2021)	D	-					I	-										
Versión Japonesa de Fugl-Meyer	Amano et al. (2018)a					V	+	I	-	I	?	I	+			V	?	V	?

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Artículos	Validez de contenido		Validez estructural		Consistencia interna		Validez transcultural/Invarianza de medición		Error de medición		Fiabilidad		Validez de criterio		Prueba de hipótesis		Sensibilidad	
		M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q
Assessment (FMA-Jap)																			
Versión Coreana de Fugl-Meyer Assessment (FMA-Kor)	Kim et al. (2021)							I	-			I	+			V	?	V	?
Versión Urdu de Fugl-Meyer Assessment (FMA-Urd)	Ikram et al. (2022)	D	-	I	-	V	-	I	-			I	+			V	?	V	?
Versión Rumana de Fugl-Meyer Assessment (FMA-Rom)	Onose et al. (2023)	D	?					I	?			I	?			I	?	I	?
Forma corta de Fugl-Meyer (37-item) (FMA-37)	Hsieh et al. (2007)											I	?	V	-			V	-

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Artículos	Validez de contenido		Validez estructural		Consistencia interna		Validez transcultural/Invarianza de medición		Error de medición		Fiabilidad		Validez criterio		Prueba de hipótesis		Sensibilidad	
		M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q
Forma corta de Fugl-Meyer (12-item) (FMA-12)	Chen et al. (2014)			D	?									V	-			V	-
	Lin et al. (2021)			D	?					I	?	I	+	V	+	D	?	D	?
	Hsieh et al. (2007)											I	?	V	-			V	-
	Hsueh et al. (2008)									D	-	D	+	V	+	V	?	V	?
	Fu et al. (2012)															V	?	V	?
	Chen et al. (2014)			D	?									V	-			V	-
Fugl-Meyer Assessment de miembro superior	Lin et al. (2021)			D	?					I	?	I	+	V	+	D	?	D	?
	Amano et al. (2020)					V	+					I	+	V	+	V	?	V	?

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Artículos	Validez de contenido		Validez estructural		Consistencia interna		Validez transcultural/Invarianza de medición		Error de medición		Fiabilidad		Validez de criterio		Prueba de hipótesis		Sensibilidad	
		M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q
acortada (6-items) (FMA-6)																			
Fugl-Meyer de aprendizaje automático (10 items) (FMA-ML-10)	Lin et al. (2021)			D	?					I	?	I	+	V	+	D	?	D	?
Traducción y adaptación del manual de administración de FMA al portugués brasileño	Michaelson et al. (2011)											I	+						
Subescala de miembro superior de Fugl-Meyer Assessment	De Weerd et al. (1985)															V	?	V	?

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Artículos	Validez de contenido		Validez estructural		Consistencia interna		Validez transcultural/Invarianza de medición		Error de medición		Fiabilidad		Validez criterio		Prueba de hipótesis		Sensibilidad	
		M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q
t (FMA-UE)	Van der Lee et al. (2001)											I	?			V	?	V	?
	Rabadi et al. (2006)															V	?	V	?
	Woodbury et al. (2007)			V	-	V	?												
	Hsieh et al. (2009)															V	?	V	?
	Lin et al. (2009)									I	?	I	+			V	?	V	?
	Sullivan et al. (2011)											I	+						
	Wei et al. (2011)													V	+	V	?	V	?
	Kim et al. (2012)									D	?	D	+	V	+	V	?	V	?
	Kim et al. (2016)													V	+	V	?	V	?

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Artículos	Validez de contenido		Validez estructural		Consistencia interna		Validez transcultural/Invarianza de medición		Error de medición		Fiabilidad		Validez criterio		Prueba de hipótesis		Sensibilidad	
		M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q
	Hernández et al. (2019)											I	?						
	Amano et al. (2020)					V	+					I	+	V	+	V	?	V	?
	Tauchi et al. (2022)			V	+														
Versión Danesa de la subescala de miembro superior de Fugl-Meyer Assessment (FMA-UE-Dan)	Lundquist et al. (2017)							I	?			D	+			D	+	D	+
Versión Italiana de la subescala de miembro superior de Fugl-Meyer Assessment	Hochleitner et al. (2023)											I	?						

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Artículos	Validez de contenido		Validez estructural		Consistencia interna		Validez transcultural/Invarianza de medición		Error de medición		Fiabilidad		Validez de criterio		Prueba de hipótesis		Sensibilidad	
		M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q
(FMA-UE-Ita)																			
Versión rumana de la subescala de miembro superior de Fugl-Meyer Assessment (FMA-UE-Rom)	Roman et al. (2020)			I	-	V	+	I	?			I	+			V	?	V	?
Fugl-Meyer Assessment de miembro superior (13-Items) usando una cámara con sensor de profundidad (Kinect) (FMA-13-Kinect)	Kim et al. (2016)													V	+	V	?	V	?

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Artículos	Validez de contenido		Validez estructural		Consistencia interna		Validez transcultural/Invarianza de medición		Error de medición		Fiabilidad		Validez criterio		Prueba de hipótesis		Sensibilidad	
		M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q
Evaluación remota de la subescala de miembro superior de Fugl-Meyer Assessment (FMA-Remote)	Amano et al. (2018) ^b											D	+						
Subescalas de hombro y codo de Fugl-Meyer Assessment (FMA-UE-SE)	Wei et al. (2011)													V	+	V	?	V	?
	See et al. (2013)											D	+			V	+	V	+
Subescalas de muñeca y mano de Fugl-Meyer Assessment (FMA-UE-	Wei et al. (2011)													V	+	V	?	V	?

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Artículos	Validez de contenido		Validez estructural		Consistencia interna		Validez transcultural/Invarianza de medición		Error de medición		Fiabilidad		Validez de criterio		Prueba de hipótesis		Sensibilidad	
		M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q	M	Q
w/h)	Page et al. (2012)					V	+					A	+			V	?	V	?
	See et al. (2013)											D	+			V	+	V	+
	Page et al. (2015)					V	+					D	+			V	?	V	?
	Persch et al. (2015)			A	?					D	?								
	Lin et al. (2004)					V	+					I	+	V	-	V	?	V	?
Subescala sensitiva de Fugl-Meyer Assessment (FMA-S)	Sullivan et al. (2011)											I	+						

Abreviaturas: A = Adecuada, D = Dudosa, Casillas vacías = no reportado, I = Inadecuada, M: Calidad metodológica, Q: Calidad de los resultados, V = Muy buena, + = puntuación positiva, ? = puntuación indeterminada, -= puntuación negativa.

Fuente: elaboración propia

Tabla 4. Calidad de la evidencia reportada por instrumento usando la metodología COSMIN.

Instrumento	Validez de contenido	Validez estructural	Consistencia interna	Validez transcultural/invarianza de medición	Error de medición	Fiabilidad	Validez de criterio	Prueba de hipótesis	Sensibilidad
Fugl-Meyer Assessment (Versión completa, 50-Items) (FMA) ^{14,21-26}	VL	M	H		VL	L	H	M	L
Versión Español-Colombiano de Fugl-Meyer Assessment (FMA-Col) (50 items) ²⁷	VL			VL					
Versión Danesa de Fugl-Meyer Assessment (FMA-Dan) (50 items)				VL					

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Validez de contenido	Validez estructural	Consistencia interna	Validez transcultural/invarianza de medición	Error de medición	Fiabilidad	Validez de criterio	Prueba de hipótesis	Sensibilidad
items) ²⁸									
Versión Italiana de Fugl-Meyer Assessment (FMA-Ita) (50 items) ²⁹	VL			VL					
Versión Japonesa de Fugl-Meyer Assessment (FMA-Jap) (50 items) ³⁰			L	VL	VL	VL		L	L
Versión Coreana de Fugl-Meyer Assessment (FMA-Kor) (50 items) ³¹				VL		VL		L	L

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Validez de contenido	Validez estructural	Consistencia interna	Validez transcultural/invarianza de medición	Error de medición	Fiabilidad	Validez de criterio	Prueba de hipótesis	Sensibilidad
Versión Urdu de Fugl-Meyer Assessment (FMA-Urd) (50 items) ³²	VL	VL	L	VL		VL		L	L
Versión Rumana de Fugl-Meyer Assessment (FMA-Rom) (50 items) ³³	VL			VL		VL		VL	VL
Forma corta de Fugl-Meyer (37-item) (FMA-37) ^{14,26,34}		M			VL	L	H	L	M
Forma corta de Fugl-Meyer (12-item) (FMA-12)		M			VL	VL	H	M	M

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Validez de contenido	Validez estructural	Consistencia interna	Validez transcultural/invarianza de medición	Error de medición	Fiabilidad	Validez de criterio	Prueba de hipótesis	Sensibilidad
12) ^{14,24,26,34,35}									
Fugl-Meyer Assessment de miembro superior acortada (6-items) (FMA-6) ³⁶			L			VL	L	L	L
Fugl-Meyer de aprendizaje automático (10 items) (FMA-ML-10) ¹⁴		L			VL	VL	H	L	L
Traducción y adaptación del manual de administración de FMA al portugués brasileño (50						VL			

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Validez de contenido	Validez estructural	Consistencia interna	Validez transcultural/invarianza de medición	Error de medición	Fiabilidad	Validez de criterio	Prueba de hipótesis	Sensibilidad
ítems) ³⁷									
Subescala de miembro superior de Fugl-Meyer Assessment (FMA-UE) (33 ítems) ^{15,36,38-48}		H	L		VL	VL	H	H	H
Versión Danesa de la subescala de miembro superior de Fugl-Meyer Assessment (FMA-UE-Dan) (33 ítems) ⁴⁹				VL		VL		VL	VL
Versión Italiana de la subescala de miembro superior						VL			

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Validez de contenido	Validez estructural	Consistencia interna	Validez transcultural/invarianza de medición	Error de medición	Fiabilidad	Validez de criterio	Prueba de hipótesis	Sensibilidad
de Fugl-Meyer Assessment (FMA-UE-Ita) (33 items) ⁵⁰									
Versión Rumanade la subescala de miembro superior de Fugl-Meyer Assessment (FMA-UE-Rom) (33 items) ⁵¹		VL	L	VL		VL		L	L
Fugl-Meyer Assessment de miembro superior (13-Items) usando una cámara con sensor de profundidad							L	L	L

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Validez de contenido	Validez estructural	Consistencia interna	Validez transcultural/invarianza de medición	Error de medición	Fiabilidad	Validez de criterio	Prueba de hipótesis	Sensibilidad
(Kinect) (FMA-13-Kinect) ⁴⁷									
Evaluación remota de la subescala de miembro superior de Fugl-Meyer Assessment (FMA-Remote) (33 items) ⁵²						VL			
Subescalas de hombro y codo de Fugl-Meyer Assessment (FMA-UE-SE) (18 items) ^{25,45}						VL	L	M	M

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

Instrumento	Validez de contenido	Validez estructural	Consistencia interna	Validez transcultural/invarianza de medición	Error de medición	Fiabilidad	Validez de criterio	Prueba de hipótesis	Sensibilidad
Subescalas de muñeca y mano de Fugl-Meyer Assessment (FMA-UE-w/h) (12 items) ^{25,45,53-55}		L	M		L	L	L	H	H
Subescala sensitiva de Fugl-Meyer Assessment (FMA-S) (12 items) ^{23,44}			H			VL	H	H	H

Abreviaturas: H=Alto, L=Bajo, M=Moderado, QE=Calidad de la evidencia, VL=Muy baja.

Fuente: elaboración propia.

4.DISCUSIÓN

Esta es la primera revisión sistemática para identificar las versiones de FMA para medir la función sensoriomotora de las extremidades superiores en personas con ictus y evaluar las propiedades psicométricas y la calidad metodológica de las versiones, basado en la metodología COSMIN más actualizada^{17, 18}. Se identificaron veintidós versiones de FMA diseñadas para evaluar la función sensoriomotora de las extremidades superiores en personas con ictus mediante la inclusión de 36 estudios.

Los profesionales clínicos e investigadores deben considerar diversos factores, como el entorno y la población donde se utilizará el instrumento, sus dimensiones, el número de ítems y la evidencia demostrada en la evaluación de cada propiedad psicométrica, para decidir qué instrumento es el más adecuado para medir este constructo. Además, se debe considerar el idioma y la cultura de las diferentes versiones al seleccionar la versión a utilizar, con qué población, país, entorno e idioma.

Se incluyeron 11 traducciones en esta revisión sistemática, mostrando heterogeneidad entre ellas, tanto en la población incluida como en el número de propiedades psicométricas evaluadas. Sin embargo, mostraron resultados similares en cuanto a las diferentes propiedades psicométricas. Además, los resultados fueron muy positivos, lo que indica que las diferentes traducciones se realizaron con métodos adecuados y que su versión final puede utilizarse en el país de destino, sin afectar la fiabilidad ni la sensibilidad de la prueba. La traducción más completa es la de FMA-Urd³², que evalúa 7 propiedades psicométricas y presenta puntuaciones psicométricas positivas. No existía una versión de FMA para medir la función sensoriomotora de las extremidades superiores, para la cual se evaluaron todas las propiedades psicométricas.

La versión completa de 50 ítems de FMA y FMA-UE han sido las 2 versiones con mayor volumen de estudios, con más propiedades psicométricas analizadas y las más traducidas; además de ser las 2 versiones más antiguas de FMA. Esto las convierte en las versiones de FMA más extendidas y utilizadas por clínicos e investigadores para evaluar la función sensoriomotora de la extremidad superior en personas post-ictus. Esto hace que estas dos FMA (Fugl-Meyer Assessment (Versión completa, 50-ítems, FMA) y Subescala de miembro superior de Fugl-Meyer Assessment (FMA-UE)) sean las más ampliamente validadas y las más sugeridas para su uso en la evaluación de la función sensoriomotora de la extremidad superior en personas con ictus tanto en la práctica clínica como en la investigación. Destacando la versión completa inicial de Fugl-Meyer Assessment (Versión completa, 50-ítems, FMA) como la más válida y recomendada porque, con base en los resultados de la presente investigación, es la versión más utilizada con más propiedades psicométricas analizadas y ha demostrado mejores puntuaciones. Esto hace que esta versión de FMA sea la más ampliamente validada y recomendada como la mejor versión y el "estándar de oro" para evaluar la función sensoriomotora de las extremidades superiores en personas con ictus.

Esto no significa que no se deban utilizar las otras 20 versiones, ya que son muy útiles en ciertos contextos y entornos específicos. Por ejemplo, en Italia se

pueden utilizar tanto FMA-Ita como FMA-EU-Ita; el evaluador debe elegir cuál seleccionar en función del constructo que desee evaluar o del tiempo que desee dedicar. Al elegir entre FMA-Ita y FMA-EU-Ita, será útil consultar la Tabla 2, ya que se recopilan las puntuaciones de las propiedades psicométricas de ambas versiones.

Otra posibilidad que ofrecen las versiones encontradas en esta revisión es poder utilizar una versión reducida de FMA, que permite una evaluación rápida y fiable, pudiendo elegir entre FMA-37 y FMA-12 si se desea una evaluación general de FMA; o por categorías de la escala, como las partes proximales del miembro superior (FMA-UE-SE), las partes distales del miembro superior (FMA-UE-w/h), la subescala sensitiva (FMA-S), o la subescala de miembro superior de 33 ítems (FMA-UE), que es un método rápido de evaluación exclusiva de la función motora del miembro superior, y que cuenta con un gran respaldo de múltiples estudios y propiedades psicométricas evaluadas. Si un terapeuta dispone de muy poco tiempo, se recomienda usar la versión FMA-12, ya que realiza una evaluación completa con FMA, no solo una subescala o una porción corporal. Además, se ha evaluado en cinco estudios con 948 participantes, donde se evaluaron las propiedades psicométricas de esta versión y las compararon con otras versiones y escalas de evaluación, con las que presenta correlaciones positivas. Además, esta versión muestra buenos niveles de fiabilidad y sensibilidad.

A lo largo de los diferentes estudios incluidos en esta revisión, se ha evaluado un tamaño muestral amplio, con aspectos muy diversos, como fase aguda/crónica, secuelas leves y graves, evaluación del progreso con evaluación inicial y seguimientos con evaluaciones periódicas; y la FMA se ha mostrado útil para la evaluación del constructo en las diferentes situaciones que se pueden presentar con la variabilidad de estos aspectos; por tanto, es una escala válida y fiable para aplicar en las diferentes situaciones que se pueden presentar en la población post-ictus; aunque con mayor preferencia por el uso de la versión completa de 50 ítems de FMA; y FMA-EU.

La principal implicación de esta revisión es ofrecer una especie de lista de las diferentes versiones de la FMA, con información sobre sus características, como el número de ítems, las partes que evalúa, el material necesario, el idioma, las propiedades psicométricas evaluadas para cada una de estas propiedades y sus puntuaciones correspondientes. Esto permite a los profesionales clínicos seleccionar la escala que mejor se adapte a sus recursos y necesidades de atención médica, y consultar rápidamente el artículo original para resolver cualquier duda. Asimismo, permite a un terapeuta clínico que dude entre usar dos versiones, poder compararlas rápidamente consultando el presente estudio. Al mismo tiempo, permite a los investigadores disponer de la versión de referencia de la FMA y basarse en esta investigación para justificar su uso en futuras investigaciones; así como consultar las propiedades psicométricas de cualquier versión de FMA, lo que ahorrará mucho tiempo y permitirá un uso mucho más válido y justificado de dicha herramienta.

Tras realizar esta investigación, se recomienda comparar la versión gold standard de FMA de forma similar a la realizada en esta investigación, de ser

posible utilizando la metodología COSMIN, con otras escalas de evaluación que evalúen el mismo constructo, para encontrar la herramienta más válida y ajustada para evaluar la función sensoriomotora del miembro superior en la población con ictus. También se recomienda analizar y comparar las diferentes versiones de FMA en otras poblaciones como Parkinson, Esclerosis Múltiple, Lesión Medular, etc. O extender esta investigación a la población neurológica general. Finalmente, se recomienda realizar traducciones y validaciones a otros idiomas, como español o chino, para poder utilizar esta escala de forma segura en estas poblaciones; además, continuar estudiando las propiedades psicométricas de las versiones que ya han sido traducidas.

4.1 Fortalezas y limitaciones

Este estudio es la revisión sistemática más reciente que identifica, evalúa y resume la evidencia sobre las versiones de FMA para medir la función sensoriomotora de las extremidades superiores en personas con ictus. La revisión se realizó según los estándares COSMIN y PRISMA más actualizados; por lo tanto, se utilizó un método adecuado, siguiendo las recomendaciones de expertos^{11, 16, 17}.

Aunque se realizó una búsqueda exhaustiva, es posible que no se hayan identificado todas las versiones disponibles de FMA, debido a que solo se seleccionaron artículos en inglés y español, por lo que es posible que no se hayan identificado artículos desarrollados y utilizados en otros idiomas.

A pesar de que la FMA ha demostrado ser una escala altamente válida y confiable, y ser un “gold standard”, presenta limitaciones como un rango de puntuación de 3 puntos, que puede limitar la detección de pequeños cambios en usuarios que se recuperan poco; así como aspectos funcionales de la función sensoriomotora del miembro superior, como su uso en actividades de la vida diaria.

5.CONCLUSIONES

Se han desarrollado diferentes versiones de la FMA y traducciones transculturales para evaluar la función sensoriomotora de las extremidades superiores en personas con ictus. La evaluación de este constructo en personas con ictus permitirá mejorar la rehabilitación física en esta población. Las versiones de la FMA para evaluar la función sensoriomotora de las extremidades superiores pueden facilitar la detección, el establecimiento de objetivos, la monitorización del progreso y la toma de decisiones. La evaluación de este aspecto es muy útil en la práctica clínica y la investigación, lo que requiere nuevos conocimientos en esta área. La evaluación de esta variable es crucial, ya que permite identificar y evaluar la intervención centrada en la mejora de la función sensoriomotora de las extremidades superiores en personas con ictus. Con base en la evidencia actual de todas las propiedades de medición de las versiones que evalúan el nivel de función sensoriomotora de las extremidades superiores en personas con ictus, Fugl-Meyer Assessment (versión completa, 50 ítems, FMA) parece ser la versión más recomendada según los resultados obtenidos y las directrices COSMIN, pero esto requiere una evaluación adicional. Se necesitan más estudios de alta calidad metodológica que evalúen las propiedades de medición de las versiones de FMA para medir la función sensoriomotora de las extremidades superiores en personas con ictus.

6.DECLARACIÓN DE CONFLICTO DE INTERESES

Los autores han declarado no tener ningún conflicto de intereses.

7.CONTRIBUCIONES DEL AUTOR

Todos los autores han acordado la versión final y cumplen al menos uno de los siguientes criterios (recomendados por el ICMJE): contribuciones sustanciales a la concepción y el diseño, adquisición de datos o análisis e interpretación de datos; redacción del artículo; o revisión crítica del mismo para determinar su contenido intelectual importante.

8.FINANCIACIÓN

Esta investigación no recibió ninguna subvención específica de agencias de financiación de los sectores público, comercial o sin fines de lucro.

9.REFERENCIAS

[1]

Feigin VL, Owolabi MO, Feigin VL, Abd-Allah F, Akinyemi RO, Bhattacharjee NV, et al. Pragmatic solutions to reduce the global burden of stroke: a World Stroke Organization–Lancet Neurology Commission. *The Lancet Neurology* 2023;22:1160–206. [https://doi.org/10.1016/S1474-4422\(23\)00277-6](https://doi.org/10.1016/S1474-4422(23)00277-6).

[2]

Hua X, Liu M, Wu S. Definition, prediction, prevention and management of patients with severe ischemic stroke and large infarction. *Chinese Medical Journal* 2023;136:2912–22. <https://doi.org/10.1097/CM9.0000000000002885>.

[3]

Katan M, Luft A. Global Burden of Stroke. *Semin Neurol* 2018;38:208–11. <https://doi.org/10.1055/s-0038-1649503>.

[4]

Kwakkel G, Stinear C, Essers B, Munoz-Novoa M, Branscheidt M, Cabanas-Valdés R, et al. Motor rehabilitation after stroke: European Stroke Organisation (ESO) consensus-based definition and guiding framework. *European Stroke Journal* 2023;8:880–94. <https://doi.org/10.1177/23969873231191304>.

[5]

Ingram LA, Butler AA, Lord SR, Gandevia SC. Use a physiological profile to document upper limb motor impairment in ageing and neurological conditions. *The Journal of Physiology* 2023;601:2251–62. <https://doi.org/10.1113/JP283703>.

[6]

Li S, Ghuman J, Gonzalez-Buonomo J, Huang X, Malik A, Yozbatiran N, et al. Does Spasticity Correlate With Motor Impairment in the Upper and Lower Limbs in Ambulatory Chronic Stroke Survivors? *Am J Phys Med Rehabil* 2023;102:907–12. <https://doi.org/10.1097/PHM.0000000000002247>.

[7]

Zhao K, He C, Xiang W, Zhou Y, Zhang Z, Li J, et al. Evidence of synergy coordination patterns of upper-limb motor control in stroke patients with mild and moderate impairment. *Front Physiol* 2023;14:1214995. <https://doi.org/10.3389/fphys.2023.1214995>.

[8]

Dawson J, Engineer ND, Cramer SC, Wolf SL, Ali R, O'Dell MW, et al. Vagus Nerve Stimulation Paired With Rehabilitation for Upper Limb Motor Impairment and Function After Chronic Ischemic Stroke: Subgroup Analysis

of the Randomized, Blinded, Pivotal, VNS-REHAB Device Trial. *Neurorehabil Neural Repair* 2023;37:367–73.
<https://doi.org/10.1177/15459683221129274>.

[9]

Bushnell C, Bettger JP, Cockcroft KM, Cramer SC, Edelen MO, Hanley D, et al. Chronic Stroke Outcome Measures for Motor Function Intervention Trials: Expert Panel Recommendations. *Circ: Cardiovascular Quality and Outcomes* 2015;8. <https://doi.org/10.1161/CIRCOUTCOMES.115.002098>.

[10]

Gunduz ME, Bucak B, Keser Z. Advances in Stroke Neurorehabilitation. *JCM* 2023;12:6734. <https://doi.org/10.3390/jcm12216734>.

[11]

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;27:1147–57. <https://doi.org/10.1007/s11136-018-1798-3>.

[12]

Santisteban L, Térémetz M, Bleton J-P, Baron J-C, Maier MA, Lindberg PG. Upper Limb Outcome Measures Used in Stroke Rehabilitation Studies: A Systematic Literature Review. *PLoS ONE* 2016;11:e0154792. <https://doi.org/10.1371/journal.pone.0154792>.

[13]

Yu B, Kaku A, Liu K, Parnandi A, Fokas E, Venkatesan A, et al. Quantifying Impairment and Disease Severity Using AI Models Trained on Healthy Subjects. *ArXiv* 2023:arXiv:2311.12781v1.

[14]

Lin G-H, Huang C-Y, Lee S-C, Chen K-L, Lien J-JJ, Chen M-H, et al. A 10-item Fugl-Meyer Motor Scale Based on Machine Learning. *Physical Therapy* 2021;101:pzab036. <https://doi.org/10.1093/ptj/pzab036>.

[15]

Tauchi Y, Kyougoku M, Takahashi K, Okita Y, Takebayashi T. Dimensionality and item-difficulty hierarchy of the Fugl-Meyer assessment of the upper extremity among Japanese patients who have experienced stroke. *Topics in Stroke Rehabilitation* 2022;29:579–87. <https://doi.org/10.1080/10749357.2021.1965797>.

[16]

Moher D, Liberati A, Tetzlaff J, Altman DG, for the PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;339:b2535–b2535. <https://doi.org/10.1136/bmj.b2535>.

[17]

Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res* 2010;19:539–49. <https://doi.org/10.1007/s11136-010-9606-8>.

[18]

Terwee CB, Mokkink LB, Knol DL, Ostelo RWJG, Bouter LM, De Vet HCW. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res* 2012;21:651–7. <https://doi.org/10.1007/s11136-011-9960-1>.

[19]

Terwee CB, Jansma EP, Riphagen II, De Vet HCW. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Qual Life Res* 2009;18:1115–23. <https://doi.org/10.1007/s11136-009-9528-5>.

[20]

Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924–6. <https://doi.org/10.1136/bmj.39489.470347.AD>.

[21]

Sanford J, Moreland J, Swanson LR, Stratford PW, Gowland C. Reliability of the Fugl-Meyer Assessment for Testing Motor Performance in Patients Following Stroke. *Physical Therapy* 1993;73:447–54. <https://doi.org/10.1093/ptj/73.7.447>.

[22]

Malouin F, Pichard L, Bonneau C, Durand A, Corriveau D. Evaluating motor recovery early after stroke: Comparison of the Fugl-Meyer Assessment and the Motor Assessment Scale. *Archives of Physical Medicine and Rehabilitation* 1994;75:1206–12. [https://doi.org/10.1016/0003-9993\(94\)90006-X](https://doi.org/10.1016/0003-9993(94)90006-X).

[23]

Lin J-H, Hsueh I-P, Sheu C-F, Hsieh C-L. Psychometric properties of the sensory scale of the Fugl-Meyer Assessment in stroke patients. *Clin Rehabil* 2004;18:391–7. <https://doi.org/10.1191/0269215504cr737oa>.

[24]

Hsueh I-P, Hsu M-J, Sheu C-F, Lee S, Hsieh C-L, Lin J-H. Psychometric Comparisons of 2 Versions of the Fugl-Meyer Motor Scale and 2 Versions of the Stroke Rehabilitation Assessment of Movement. *Neurorehabil Neural Repair* 2008;22:737–44. <https://doi.org/10.1177/1545968308315999>.

[25]

See J, Dodakian L, Chou C, Chan V, McKenzie A, Reinkensmeyer DJ, et al. A Standardized Approach to the Fugl-Meyer Assessment and Its Implications for Clinical Trials. *Neurorehabil Neural Repair* 2013;27:732–41. <https://doi.org/10.1177/1545968313491000>.

[26]

Chen K-L, Chen C-T, Chou Y-T, Shih C-L, Koh C-L, Hsieh C-L. Is the Long Form of the Fugl-Meyer Motor Scale More Responsive Than the Short Form in Patients With Stroke? *Archives of Physical Medicine and Rehabilitation* 2014;95:941–9. <https://doi.org/10.1016/j.apmr.2014.01.014>.

[27]

Barbosa NE, Forero SM, Galeano CP, Hernández ED, Landinez NS, Sunnerhagen KS, et al. Translation and cultural validation of clinical observational scales – the Fugl-Meyer assessment for post-stroke sensorimotor function in Colombian Spanish. *Disability and Rehabilitation* 2019;41:2317–23. <https://doi.org/10.1080/09638288.2018.1464604>.

[28]

Busk H, Alt Murphy M, Korsman R, Skou ST, Wienecke T. Cross-cultural translation and adaptation of the Danish version of the Fugl-Meyer assessment for post-stroke sensorimotor function. *Disability and Rehabilitation* 2022;44:4888–95. <https://doi.org/10.1080/09638288.2021.1919215>.

[29]

Cecchi F, Carrabba C, Bertolucci F, Castagnoli C, Falsini C, Gnetti B, et al. Transcultural Translation and validation of Fugl-Meyer assessment to Italian. *Disability and Rehabilitation* 2021;43:3717–22. <https://doi.org/10.1080/09638288.2020.1746844>.

[30]

Amano S, Umeji A, Uchita A, Hashimoto Y, Takebayashi T, Takahashi K, et al. Clinimetric properties of the Fugl-Meyer assessment with adapted guidelines for the assessment of arm function in hemiparetic patients after stroke. *Topics in Stroke Rehabilitation* 2018a;25:500–8. <https://doi.org/10.1080/10749357.2018.1484987>.

[31]

Kim T, Hwang SH, Lee WJ, Hwang JW, Cho I, Kim E-H, et al. The Korean Version of the Fugl-Meyer Assessment: Reliability and Validity Evaluation. *Ann Rehabil Med* 2021;45:83–98. <https://doi.org/10.5535/arm.20225>.

[32]

Ikram M, Rehman SSU, Sunnerhagen KS, Alt Murphy M. Urdu translation and cross-cultural validation of the Fugl-Meyer assessment in people with stroke. *Disability and Rehabilitation* 2022;44:8048–53. <https://doi.org/10.1080/09638288.2021.2003449>.

[33]

Onose G, Anghelescu A, Ionescu A, Tataranu LG, Spînu A, Bumbea AM, et al. Translation of the Fugl-Meyer assessment into Romanian: Transcultural and semantic-linguistic adaptations and clinical validation. *Front Neurol* 2023;13:1022546. <https://doi.org/10.3389/fneur.2022.1022546>.

[34]

Hsieh Y-W, Hsueh I-P, Chou Y-T, Sheu C-F, Hsieh C-L, Kwakkel G. Development and Validation of a Short Form of the Fugl-Meyer Motor Scale in Patients With Stroke. *Stroke* 2007;38:3052–4. <https://doi.org/10.1161/STROKEAHA.107.490730>.

[35]

Fu TS-T, Wu C, Lin K, Hsieh C, Liu J, Wang T, et al. Psychometric comparison of the shortened Fugl-Meyer Assessment and the streamlined Wolf Motor Function Test in stroke rehabilitation. *Clin Rehabil* 2012;26:1043–7. <https://doi.org/10.1177/0269215511431474>.

[36]

Amano S, Umeji A, Takebayashi T, Takahashi K, Uchiyama Y, Domen K. Clinimetric properties of the shortened Fugl-Meyer Assessment for the assessment of arm motor function in hemiparetic patients after stroke. *Topics in Stroke Rehabilitation* 2020;27:290–5. <https://doi.org/10.1080/10749357.2019.1701176>.

[37]

Michaelsen SM, Rocha AS, Knabben RJ, Rodrigues LP, Fernandes CGC. Translation, adaptation and inter-rater reliability of the administration manual for the Fugl-Meyer assessment. *Rev Bras Fisioter* 2011;15:80–8.

[38]

De Weerdt WJG. Measuring recovery of arm-hand function in stroke patients: A comparison of the Brunnstrom-Fugl-Meyer test and the Action Research Arm test. *Physiotherapy Canada* 1985;37:65–70. <https://doi.org/10.3138/ptc.37.2.065>.

[39]

H. Van Der Lee, Heleen Beckerman, G J. The responsiveness of the Action Research Arm Test and the Fugl-Meyer Assessment scale in chronic stroke patients. *Journal of Rehabilitation Medicine* 2001;33:110–3. <https://doi.org/10.1080/165019701750165916>.

[40]

Rabadi MH, Rabadi FM. Comparison of the Action Research Arm Test and the Fugl-Meyer Assessment as Measures of Upper-Extremity Motor Weakness After Stroke. *Archives of Physical Medicine and Rehabilitation* 2006;87:962–6. <https://doi.org/10.1016/j.apmr.2006.02.036>.

[41]

Woodbury ML, Velozo CA, Richards LG, Duncan PW, Studenski S, Lai S-M. Dimensionality and Construct Validity of the Fugl-Meyer Assessment of the Upper Extremity. *Archives of Physical Medicine and Rehabilitation* 2007;88:715–23. <https://doi.org/10.1016/j.apmr.2007.02.036>.

[42]

Hsieh Y, Wu C, Lin K, Chang Y, Chen C, Liu J. Responsiveness and Validity of Three Outcome Measures of Motor Function After Stroke Rehabilitation. *Stroke* 2009;40:1386–91. <https://doi.org/10.1161/STROKEAHA.108.530584>.

[43]

Lin J-H, Hsu M-J, Sheu C-F, Wu T-S, Lin R-T, Chen C-H, et al. Psychometric Comparisons of 4 Measures for Assessing Upper-Extremity Function in People With Stroke. *Physical Therapy* 2009;89:840–50. <https://doi.org/10.2522/ptj.20080285>.

[44]

Sullivan KJ, Tilson JK, Cen SY, Rose DK, Hershberg J, Correa A, et al. Fugl-Meyer Assessment of Sensorimotor Function After Stroke: Standardized Training Procedure for Clinical Practice and Clinical Trials. *Stroke* 2011;42:427–32. <https://doi.org/10.1161/STROKEAHA.110.592766>.

[45]

Wei X-J, Tong K, Hu X. The responsiveness and correlation between Fugl-Meyer Assessment, Motor Status Scale, and the Action Research Arm Test in chronic stroke with upper-extremity rehabilitation robotic training. *International Journal of Rehabilitation Research* 2011;34:349–56. <https://doi.org/10.1097/MRR.0b013e32834d330a>.

[46]

Kim H, Her J, Ko J, Park DS, Woo JH, You Y, et al. Reliability, concurrent validity, and responsiveness of the Fugl-Meyer Assessment (FMA) for hemiplegic patients. JPTS 2012;24:893–9.

[47]

Kim W-S, Cho S, Baek D, Bang H, Paik N-J. Upper Extremity Functional Evaluation by Fugl-Meyer Assessment Scoring Using Depth-Sensing Camera in Hemiplegic Stroke Patients. PLoS ONE 2016;11:e0158640. <https://doi.org/10.1371/journal.pone.0158640>.

[48]

Hernández E, Galeano C, Barbosa N, Forero S, Nordin , Sunnerhagen K, et al. Intra- and inter-rater reliability of Fugl-Meyer Assessment of Upper Extremity in stroke. J Rehabil Med 2019;51:652–9. <https://doi.org/10.2340/16501977-2590>.

[49]

Lundquist CB, Maribo T. The Fugl–Meyer assessment of the upper extremity: reliability, responsiveness and validity of the Danish version. Disability and Rehabilitation 2017;39:934–9. <https://doi.org/10.3109/09638288.2016.1163422>.

[50]

Hochleitner I, Pellicciari L, Castagnoli C, Paperini A, Politi AM, Campagnini S, et al. Intra- and inter-rater reliability of the Italian Fugl-Meyer assessment of upper and lower extremity. Disability and Rehabilitation 2023;45:2989–99. <https://doi.org/10.1080/09638288.2022.2114553>.

[51]

Roman N, Miclaus R, Repanovici A, Nicolau C. Equal Opportunities for Stroke Survivors' Rehabilitation: A Study on the Validity of the Upper Extremity Fugl-Meyer Assessment Scale Translated and Adapted into Romanian. Medicina 2020;56:409. <https://doi.org/10.3390/medicina56080409>.

[52]

Amano S, Umeji A, Uchita A, Hashimoto Y, Takebayashi T, Kanata Y, et al. Reliability of remote evaluation for the Fugl–Meyer assessment and the action research arm test in hemiparetic patients after stroke. Topics in Stroke Rehabilitation 2018b;1–6. <https://doi.org/10.1080/10749357.2018.1481569>.

[53]

Page SJ, Levine P, Hade E. Psychometric Properties and Administration of the Wrist/Hand Subscales of the Fugl-Meyer Assessment in Minimally Impaired Upper Extremity Hemiparesis in Stroke. Archives of Physical Medicine and Rehabilitation 2012;93:2373–2376.e5. <https://doi.org/10.1016/j.apmr.2012.06.017>.

[54]

Page SJ, Hade E, Persch A. Psychometrics of the Wrist Stability and Hand Mobility Subscales of the Fugl-Meyer Assessment in Moderately Impaired Stroke. *Physical Therapy* 2015;95:103–8.
<https://doi.org/10.2522/ptj.20130235>.

[55]

Persch AC, Gugiu PC, Velozo CA, Page SJ. Rasch Analysis of the Wrist and Hand Fugl-Meyer: Dimensionality and Item-Level Characteristics. *Journal of Neurologic Physical Therapy* 2015;39:185–92.
<https://doi.org/10.1097/NPT.0000000000000096>.

6.2 INVESTIGACIÓN CLÍNICA

6.2.1 ACTUALIZACIÓN, TRADUCCIÓN, ADAPTACIÓN TRANSCULTURAL Y VALIDACIÓN DE NOTTINGHAM STROKE DRESSING ASSESSMENT AL ESPAÑOL.

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

ACTUALIZACIÓN, TRADUCCIÓN, ADAPTACIÓN TRANSCULTURAL Y VALIDACIÓN DE NOTTINGHAM STROKE DRESSING ASSESSMENT AL ESPAÑOL.

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

El Artículo III se encuentra bajo revisión y pendiente para su publicación en la revista “American Journal of Occupational Therapy”.

ABSTRACT

Background and purpose: The Nottingham Stroke Dressing Assessment (NSDA) is a scale for assessing the ability and independence to dress in people who have suffered a stroke. The objective of this study was to develop an update of the original scale and create a Spanish version of the scale.

Methods: The NSDA scale was reviewed, translated, cross-culturally adapted, and validated. A standardized multi-step translation and updating protocol was implemented to achieve optimal conceptual and semantic equivalence and validate it. The Spanish version developed was validated in 110 stroke survivors.

Results: The exploratory analysis generated 12 factors. Internal consistency was satisfactory ($\alpha = 0.921$). Test-retest reliability was high ($ICC = 0.916$). The correlation between the ENVI-R and the FMA was demonstrated and significant ($r = 0.34-0.5$).

Conclusions: The updated, translated, and culturally validated Spanish version of this scale can be used reliably and validly in research and clinical practice after demonstrating adequate and significant psychometric properties. The use of these scales will improve the quality of stroke dressing assessment and allow for reliable comparisons.

Keywords: Dressing; Stroke; ENVI-R; Translation; Validation

RESUMEN

Antecedentes y propósito: Nottingham Stroke Dressing Assessment (NSDA) es una escala de evaluación de la capacidad e independencia para el vestido en personas que han sufrido un ictus. El objetivo de este estudio fue desarrollar una actualización de la escala original y generar una versión española de la escala.

Métodos: Se realizó una revisión, traducción, adaptación transcultural y validación de la escala NSDA, para lo cual se implementó un protocolo de actualización y traducción estandarizado de varios pasos, así como para lograr una equivalencia conceptual y semántica óptima y validarlo. La versión española desarrollada fue validada en 110 personas que habían sufrido un ictus.

Resultados: El análisis exploratorio generó 12 factores. La consistencia interna fue satisfactoria ($\alpha = 0.921$). La fiabilidad test-retest fue elevada ($ICC=0.916$). La correlación encontrada entre ENVI-R y FMA fue demostrada y significativa ($r=0.34-0.5$).

Conclusiones: La versión española de esta escala, actualizada, traducida y culturalmente validada, se puede utilizar de forma fiable y válida en investigaciones y en la práctica clínica tras mostrar propiedades psicométricas adecuadas y significativas. El uso de estas escalas mejorará la calidad de la evaluación del vestido en el ictus y permitirá comparaciones fiables.

Palabras clave: Vestido; Ictus; ENVI-R; Traducción; Validación.

1.INTRODUCCIÓN

El ictus es la segunda causa de muerte¹ y la primera causa de discapacidad en todo el mundo². Quince millones de personas sufren un ictus cada año en todo el mundo; de ellas, 5 millones mueren y otros 5 millones presentan discapacidad permanentemente, lo que supone una carga para la familia y la comunidad³.

La carga de discapacidad después de un ictus es significativa y aumenta muy rápidamente, y la incidencia de ictus está aumentando en personas jóvenes y de mediana edad menores de 55 años a nivel mundial¹. Se estima que la carga de ictus aumentará un 50% en los próximos años, de 6,6 millones de casos en 2020 a 9,7 millones de casos en 2050¹.

En España, los ictus son la primera causa de discapacidad en el adulto y la segunda causa de demencia⁴. En 2023 se registraron 23.173 defunciones por ACV, de las cuales 12.817 fueron mujeres⁵. Los estudios estiman que la incidencia de ictus es de 187 casos por cada 100.000 habitantes/año, mayor que en Inglaterra, Alemania y Francia, pero menor que en otros países de nuestro entorno como Italia, Grecia, o Portugal⁵.

Las disfunciones del ictus, que se manifiestan principalmente como deterioros motores, de percepción y cognitivos, con frecuencia afectan el desempeño de las actividades de la vida diaria (AVD), las tareas ocupacionales y, en última instancia, la independencia y la funcionalidad de los sobrevivientes del ictus⁶. Tras el ictus se produce una alteración física de las extremidades y deterioro cognitivo⁷⁻⁹, que provoca la incapacidad de utilizar las extremidades superiores para alcanzar, agarrar y manipular objetos, afectando la realización de tareas de la vida diaria como vestirse, alimentarse o bañarse, provocando la pérdida de autonomía en las actividades de la vida diaria y la dependencia de otros para realizarlas^{7,10}.

La capacidad de vestirse de forma independiente es una actividad básica de la vida diaria que permite a las personas ser independientes en su día a día. Debido a las alteraciones tanto motora, como perceptiva y cognitiva tras sufrir un ictus, se altera la capacidad para desempeñar de manera funcional la AVD Básica de vestido¹¹⁻²⁰. Se calcula que alrededor del 54% de los supervivientes de ictus presentan dificultad para vestirse de forma independiente a los 6 meses del ictus y alrededor del 36% continúa presentando dependencia para esta AVD a los 2 años del ictus^{12,19}. Es por ello, que la evaluación del vestido en pacientes post-ictus es necesaria para conocer la independencia de los pacientes en esta AVD tan comúnmente afectada y demandada por los pacientes y familiares para mejorar la autonomía de estas personas, lo cual es clave en la rehabilitación postictus^{15,16}. La posibilidad de monitorizar el progreso de los pacientes empleando evaluaciones cuantitativas y cualitativas de la actividad del vestido permite a los equipos de rehabilitación definir objetivos específicos y personalizados que permitan optimizar la selección de tratamientos y la recuperación funcional y de independencia²¹. Existen multitud de herramientas de evaluación que cuentan con un subdominio de evaluación del vestido. Algunas de ellas son el Índice de Barthel (BI), Medida de Independencia Funcional (FIM), Escala de Rankin modificada (mRS), National Institutes of

Health Stroke Scale (NIHSS), Frenchay Activity Index (FAI), Motor Activity Log – 30 ítems (MAL-30), Nottingham Extended Activities of Daily Living (NEADL), etc.²²⁻²⁷ Sin embargo, la actividad de vestido implica una multitud de opciones y un despliegue de habilidades tanto motoras, como perceptivas y cognitivas, lo cual requiere de evaluaciones más específicas y exhaustivas. En este sentido, existen muy pocas escalas que permitan evaluar exclusivamente y con globalidad la actividad de vestido en personas con ictus, concretamente sólo se encuentra en la literatura científica tres: Upper-Body Dressing Scale (UBDS)²⁷⁻²⁹, Nottingham Stroke Dressing Assessment (NSDA)^{12,14-19,30} y Jacket Test³¹. La escala NSDA se lleva utilizando en la práctica clínica desde hace treinta años¹⁶, lo que la convierte en la herramienta más empleada y con mayor presencia en la bibliografía científica relacionada con problemas para el vestido en población que ha sufrido un ictus^{12,14-19,30}. La escala NSDA se ha desarrollado, empleado y estudiado en países occidentales^{12,14-19,30}, más similares a la cultura española; mientras que las escalas Jacket Test³¹ y UBDS han sido creadas, utilizadas y estudiadas en países orientales²⁷⁻²⁹, cuya cultura difiere más de la española. Nottingham Stroke Dressing Assessment (NSDA) fue originalmente publicada en inglés por Marion Walker en 1990¹⁶. Esta herramienta fue diseñada como una medida de evaluación muy completa para la AVD del vestido, valorando un gran número de prendas, descomponiendo la colocación de cada prenda en pasos y otorgando una puntuación ajustada a la independencia o necesidad de ayuda para vestirse en cada paso. A su vez, analiza si el origen de los errores del vestido son motores, cognitivos, perceptivos, motivacionales, o la combinación de ellos, lo cual no se tiene en cuenta en otras escalas. Finalmente, refleja una puntuación total de la independencia que tiene la persona que ha sufrido un ictus para vestirse de forma autónoma por completo, no solo una parte del cuerpo^{12,14-19,30}.

La importancia de adaptar los instrumentos a los entornos de investigación y práctica clínica actuales es universalmente reconocida y necesaria para garantizar que los conceptos explorados por una herramienta sean iguales entre el idioma y el contexto original y de destino. Un proceso integral de traducción lingüística y un proceso de adaptación transcultural son necesarios para garantizar la validez de constructo y la confiabilidad del instrumento adaptado³². La evaluación de la equivalencia operativa de la herramienta traducida, lo que significa que se pueden aplicar formatos, instrucciones, modos de administración y métodos de medición similares a las poblaciones objetivo, también es necesaria para adaptar un instrumento validado a un idioma, entorno y momento diferentes³³. Estos aspectos se hacen más relevantes en el caso de la Terapia Ocupacional, la cual cuenta con menor número de escalas de valoración en comparación con otras disciplinas, y en su mayor parte en idiomas distintos al castellano, por lo que, o no se pueden utilizar en España, o su uso se encuentra sesgado, por malas traducciones y sin validar. Es por ello, que la actualización, traducción, adaptación transcultural y validación de una escala del vestido, que es una actividad propia de terapia ocupacional, es de gran relevancia para esta disciplina, y más concretamente en España, ya que no cuenta con herramientas estandarizadas para valorar en profundidad la AVD del vestido, permitiendo esta

traducción que esta escala de evaluación sea utilizada de manera estandarizada por los terapeutas ocupacionales en España³⁴.

Por todo ello, el objetivo de este estudio fue desarrollar una revisión y actualización de la escala NSDA, generando una versión más novedosa denominada NSDA-Revised (NSDA-R), para su posterior traducción y adaptación transcultural de la escala NSDA-R al español, según un procedimiento estandarizado que permitió elaborar la versión denominada Escala Nottingham del Vestido en Ictus-Revisada (ENVI-R); para finalmente validar la versión oficial española (ENVI-R) en una muestra de pacientes post-ictus.

2.MATERIALES Y MÉTODOS

Diseño del estudio

Estudio de revisión y actualización de la versión original Nottingham Stroke Dressing Assessment (NSDA) a la versión revisada Nottingham Stroke Dressing Assessment-Revised (NSDA-R). Posteriormente, estudio de traducción, adaptación transcultural y validación de NSDA-R al castellano, generando la versión Escala Nottingham del Vestido en Ictus-Revisada (ENVI-R).

Proceso de revisión y actualización de la escala

El proceso de revisión y actualización de NSDA-R se llevó a cabo en una sola fase que supuso un análisis exhaustivo de la escala y varias reuniones de comité de expertos y consultas con las autoras de la versión original.

Proceso de traducción, adaptación cultural y validación

El proceso de traducción, adaptación transcultural y validación de ENVI-R fue llevado a cabo en dos fases: la traducción y adaptación transcultural de ENVI-R, y, por otra parte, la validación de la escala.

Esta escala consta de 3 documentos:

- El primer documento con las instrucciones para la administración de la escala, recomendaciones para el terapeuta, así como su sistema de puntuación por ítems y de puntuación final.
- El segundo documento es de evaluación cuantitativa. Tiene 2 versiones: femenina y masculina. Consiste en una tabla de puntuación y un apartado de observaciones. En ella aparecen las diferentes prendas de vestir y cierres que más se emplean. Para cada prenda hay un desglose de los diferentes pasos que se realizan desde que se coge la prenda hasta que la persona consigue colocarla. Teniendo en cuenta las diferentes estrategias de vestido que se pueden emplear.
- El tercer documento es de evaluación cualitativa. Es una tabla de análisis de los errores más comúnmente cometidos durante el vestido, por diferentes factores, tanto motores, como cognitivos y perceptivos.

Cada paso o cierre de prenda del segundo documento se puntúa en una escala ordinal de 3 puntos (0=necesita asistencia física para realizarlo, 1=necesita asistencia no física para realizarlo, 2=lo hace de manera independiente)¹⁵.

Fase 1. Proceso de revisión y actualización

Dado que la escala se publicó en 1990^{15,16} la tendencia de moda de vestimenta ha cambiado mucho; y tras contactar con la autora de la versión original, se consideró apropiado revisar la escala proponiendo una actualización de la misma, para facilitar y agilizar la administración de esta prueba.

Para ello, se formó un comité de expertos³⁵ formado por siete terapeutas ocupacionales investigadores, seis de ellos fueron españoles, cuatro de ellos doctores, profesores e investigadores de diferentes universidades españolas,

expertos en traducción y validación de escalas, y los otros 2 fueron estudiantes de doctorado y expertos en neurorrehabilitación. También participó una terapeuta ocupacional danesa, doctora, profesora e investigadora de la Universidad de Dinamarca, que posteriormente trabajó paralelamente en la traducción y adaptación transcultural de la escala NSDA-R al danés. Todos ellos mostraban un dominio fluido del inglés, lo que permitió mayor exactitud y rigurosidad a la hora de realizar las traducciones y las reuniones de comité. Cada uno de los expertos revisó la escala de forma individual, proponiendo las modificaciones que creían oportunas. Llevaron a cabo 7 reuniones para consensuar acerca de las modificaciones a realizar.

Se contactó con la autora de la versión original de la escala^{15,16}, para la aclaración de determinadas cuestiones que favoreciesen la actualización de la herramienta de evaluación.

El investigador principal llevaba a cabo las modificaciones y revisaba de manera exhaustiva la escala tras cada reunión, estableciendo los puntos críticos conflictivos de la revisión para abordarlos desde el comité de expertos en las posteriores reuniones. Este proceso se llevó a cabo entre enero de 2022 y julio de 2022^{36,37}.

Tras la revisión y actualización se decidió cambiar el nombre a la escala para denominarla “Nottingham Stroke Dressing Assessment-Revised” (NSDA-R), con el consentimiento y aprobación por parte de la autora de la versión original.

Fase 2. Proceso de traducción y adaptación transcultural

Se tradujeron todos los documentos que forman parte de la escala y que han sido detallados previamente. Para el proceso de traducción se siguió el procedimiento metodológico recomendado por la guía ISPOR, la cual resulta en un riguroso proceso estandarizado de diez pasos de traducciones de ida y de vuelta, supervisadas metódicamente por un amplio grupo de expertos, permitiendo determinar la equivalencia conceptual, semántica y operativa de la escala traducida³⁵⁻³⁷. Este proceso se llevó a cabo entre los meses de septiembre de 2022 y enero de 2023.

Dos investigadores independientes tradujeron la escala NSDA-R al castellano y la enviaron al comité de expertos para que revisasen las 2 versiones. Tras esto, se reunió el comité de expertos en cinco ocasiones para comparar las 2 versiones, seleccionar las mejores partes de cada una de ellas y combinarlo en una primera versión en castellano. Esa primera versión en castellano se tradujo al inglés de vuelta por una persona nativa anglosajona.

A continuación, se comparó la versión inglesa original con la nueva versión inglesa de vuelta tanto por el comité de expertos como por la autora de la versión original con el fin de unificar las 2 versiones inglesas.

Una vez generada la nueva versión inglesa, el investigador principal tradujo esa versión inglesa al castellano, produciendo la versión en castellano de vuelta. El comité de expertos se reunió de nuevo para revisar esa última versión en castellano, realizar las últimas observaciones y cambios, y dar el visto bueno a esta última versión, configurando la versión española final, denominada “Escala Nottingham del Vestido en Ictus-Revisada (ENVI-R)”.

Este comité de expertos estuvo formado por 6 terapeutas ocupacionales españoles, 4 de ellos doctores, profesores e investigadores de diferentes universidades españolas, expertos en traducción y validación de escalas; y los otros 2 eran estudiantes de doctorado y expertos en neurorrehabilitación. Todos ellos dominaban ambos idiomas, lo que permitió llevar a cabo este proceso de traducción de manera eficiente. En la figura 1 se muestra el flujograma con el proceso multipasos de traducción y adaptación transcultural.

Fase 3. Proceso de validación

Se realizó la validación de la escala, estudiando sus propiedades psicométricas en población con ictus: validez estructural, consistencia interna, fiabilidad test-retest, error de medición, validez de criterio, prueba de hipótesis y capacidad de respuesta.

Para validar la versión NSDA-R adaptada al contexto español: ENVI-R, se administró esta escala a 110 pacientes post-ictus.

Los pacientes fueron reclutados de manera continua entre los meses de febrero y septiembre de 2024, en 9 centros de neurorrehabilitación privados y concertados repartidos por el territorio español. Los pacientes incluidos en el estudio se encontraban tanto en una fase aguda como en una fase crónica, cumpliendo los siguientes criterios de inclusión y de exclusión.

Criterios de inclusión: a) Personas mayores de edad (18-99 años); b) Haber sufrido, al menos, un ictus; c) Ser capaz de mover activamente (Daniels 4/5) el hemicuerpo “sano”.

Criterios de exclusión: a) Ausencia/Disminución de respuesta al medio (Glasgow inferior a 15); b) Incapacidad para comprender las instrucciones del terapeuta; c) Estar conectado a un dispositivo de soporte vital básico; d) Agresividad o suponer un riesgo para sí mismo o los demás; e) Estar encamado.

Procedimiento

A todos los pacientes potencialmente elegibles de participar en el estudio se les informó y se les pidió que leyesen y firmasen un consentimiento informado en el que aparecía toda la información referente a la investigación y su participación. Una vez firmado, se les citaba un día concreto en su centro de rehabilitación de referencia para participar en este estudio mediante la evaluación con las escalas ENVI-R y Fugl-Meyer Assessment (FMA). Para evaluar la validez de criterio de ENVI-R, se evaluó a los mismos pacientes, y el mismo día, con la escala FMA, una escala de evaluación clínica de la funcionalidad sensoriomotora de miembros superiores e inferiores, ampliamente empleada en la práctica clínica y en la investigación con personas post-ictus. Se trata de un instrumento válido y fiable en dicha población³⁸⁻⁴⁰.

El investigador principal, terapeuta ocupacional clínico, que conoce y domina la escala ENVI-R, administró la escala y evaluó el desempeño y la independencia de los pacientes durante la actividad del vestido.

Se llevó a cabo una reevaluación por parte del mismo terapeuta de la evaluación inicial en el 59,1% de la muestra. Esta reevaluación fue realizada entre 1 y 4 meses después de la primera evaluación.

El examinador no comparó las puntuaciones durante la sesión de evaluación ni después, y las pruebas completadas permanecieron ciegas para el resto de los investigadores hasta que se completaron todas las evaluaciones.

Todas las preguntas e incertidumbres destacadas durante el ensayo fueron informadas y revisadas por el comité de expertos para llegar a un consenso sobre la versión definitiva de ENVI-R.

Análisis estadístico

Las variables sociodemográficas y clínicas fueron expresadas en medias y porcentajes. La edad, fase (aguda/crónica), ubicación geográfica y sexo fueron especificados.

El análisis de las propiedades psicométricas de la escala se llevó a cabo basándose en la metodología empleada por la normativa COSMIN⁴¹⁻⁴⁴.

La prueba Kaiser-Meyer-Olkin (KMO) de adecuación del muestreo medido y el test de esfericidad de Bartlett's se utilizaron para comprobar la redundancia entre variables; KMO >0.5 se consideró bueno, junto con valores significativos para el test de esfericidad de Bartlett's («IBM Kaiser-Meyer-Olkin medida para la matriz de correlación de identidad United States», 2011). Con el análisis factorial exploratorio (EFA) con extracción de máxima verosimilitud y rotación Varimax se estimó la estructura interna del nuevo cuestionario.

La consistencia interna del cuestionario se midió con el análisis estadístico α de Cronbach⁴⁵. Este dato fue arrojado en términos absolutos y por factores.

El coeficiente de correlación intraclase (ICC) se utilizó para medir la confiabilidad intraobservador.

El Error estándar de la medida fue usado para conocer el error de medición.

La validación de criterio convergente se refirió a la evaluación de instrumentos que evaluaban variables que medían constructos similares o que debían mostrar una relación similar. La relación de estas puntuaciones con las variables se utilizó para validar la escala. El coeficiente de correlación de Spearman se utilizó para evaluar la validez convergente entre las variables ENVI-R y el Fugl-Meyer Assessment. Para llevar a cabo las mencionadas correlaciones se utilizó en primer lugar la prueba de Kolmogorov-Smirnov para conocer si los datos seguían una distribución normal y así usar análisis paramétricos o no paramétricos. Al ser una muestra significativa no paramétrica se empleó el test de Spearman para obtener coeficientes de correlación. Como hipótesis se plantea que existirá una correlación positiva significativa entre ambas escalas.

En todos los contrastes de hipótesis se consideró un nivel de significación estadística del 0.05. El análisis estadístico fue realizado empleando el software SPSS versión 29 (IBM, USA), excepto para el análisis factorial, que se empleó el software JASP.

Consideraciones éticas

Esta investigación contó con la aprobación del Comité de Ética de la Universidad de Málaga, con número de registro 7-2024-H, y respeta los preceptos éticos fundamentales según la declaración de Helsinki y el Informe Belmont, además de la normativa legal vigente en nuestro país en materia de investigación clínica, especialmente la Ley 14/2007, de 3 de julio, de Investigación Biomédica. Se garantizó en todo momento la protección y confidencialidad de la información, ya que el tratamiento, la comunicación y la cesión de los datos de carácter personal de todos los sujetos participantes se ajustará a lo dispuesto en la Ley Orgánica 3/2018, del 5 de diciembre, de protección de datos personales y garantía de los derechos digitales, y en su reglamento de desarrollo de datos. Los datos serán tratados de forma confidencial, asignándose un código correlativo y en ningún caso, figurará el nombre y apellidos del paciente u otros datos que comprometan su intimidad. Los datos fueron tratados informáticamente y se incorporaron a un fichero automatizado de datos de carácter personal.

3.RESULTADOS

Participantes

110 pacientes que cumplieron con los criterios de inclusión formaron parte de la muestra de este estudio. No hubo pérdidas de pacientes en la recogida de datos. De los participantes del estudio, el 73.64% se encontraba en fase crónica (n=81), y el 26.36% en fase aguda (n=29). La edad media de los pacientes fue de 62 ± 13.62 años. Los hombres representaron un 60% de la muestra total (n=66) y las mujeres el 40% (n=44).

Fase 1. Proceso de revisión y actualización

El comité de expertos realizó diversas revisiones de la escala, tanto individuales como colectivas, coincidiendo de manera uniforme en los cambios a llevar a cabo.

Entre las modificaciones realizadas para la actualización se encontraron: a) Modificación de las instrucciones de aplicación y uso de la escala con el fin de facilitar la comprensión por parte de los terapeutas, y la interpretación de diferentes situaciones que se puedan dar durante la administración de la escala (documento 1); b) Unificación de las tablas de puntuación de escalas de hombre y mujer del documento 2; c) Inclusión y/o modificación de nombres de prendas en los títulos de las categorías de las prendas de vestido (documento 2), cuyos pasos de colocación coincidían entre ellos para englobarse en el mismo ítem de vestido; d) Inclusión de fotografías en los ítems de la tabla de puntuación (documento 2), para facilitar la comprensión y agilizar la administración; e) Modificaciones en la explicación y definición de los ítems del documento “Análisis de los Errores de Acción” para facilitar la comprensión y agilizar la cumplimentación (documento 3).

Por lo tanto, la escala continuó manteniendo el primer documento de “Instrucciones para los terapeutas” con algunas modificaciones que facilitaron su comprensión, implementación e interpretación.

También mantuvo la estructura y forma la “Escala de puntuación” del segundo documento, pero unificando los ítems para hombre y mujer en una sola tabla, añadiendo prendas por ítem, pero manteniendo los mismos ítems, e incluyendo fotografías de dichas prendas. Esta tabla incluye los siguientes dominios: 1. Cierres; 2. Sujetador/Top; 3. Camiseta de tirantes; 4. Camiseta, Jersey de manga corta; 5. Vestido; 6. Camisa, Chaqueta, Abrigo; 7. Camiseta, Jersey de manga larga; 8. Falda; 9. Bragas/Calzoncillos; 10. Pantalones, Leggings; 11. Medias, Pantys, Calcetines; 12. Calzado.

Dentro de cada una de estas categorías, se encuentran los ítems o pasos que deben ser puntuados. Esta puntuación permanece igual que la versión original (0, 1 y 2, en función de la ayuda necesaria). El cálculo de la puntuación final se realiza analizando la cantidad de ítems intentados, y la puntuación otorgada a cada uno de ellos. Esta puntuación final se mantiene igual que en la versión original y se sigue expresando en porcentaje.

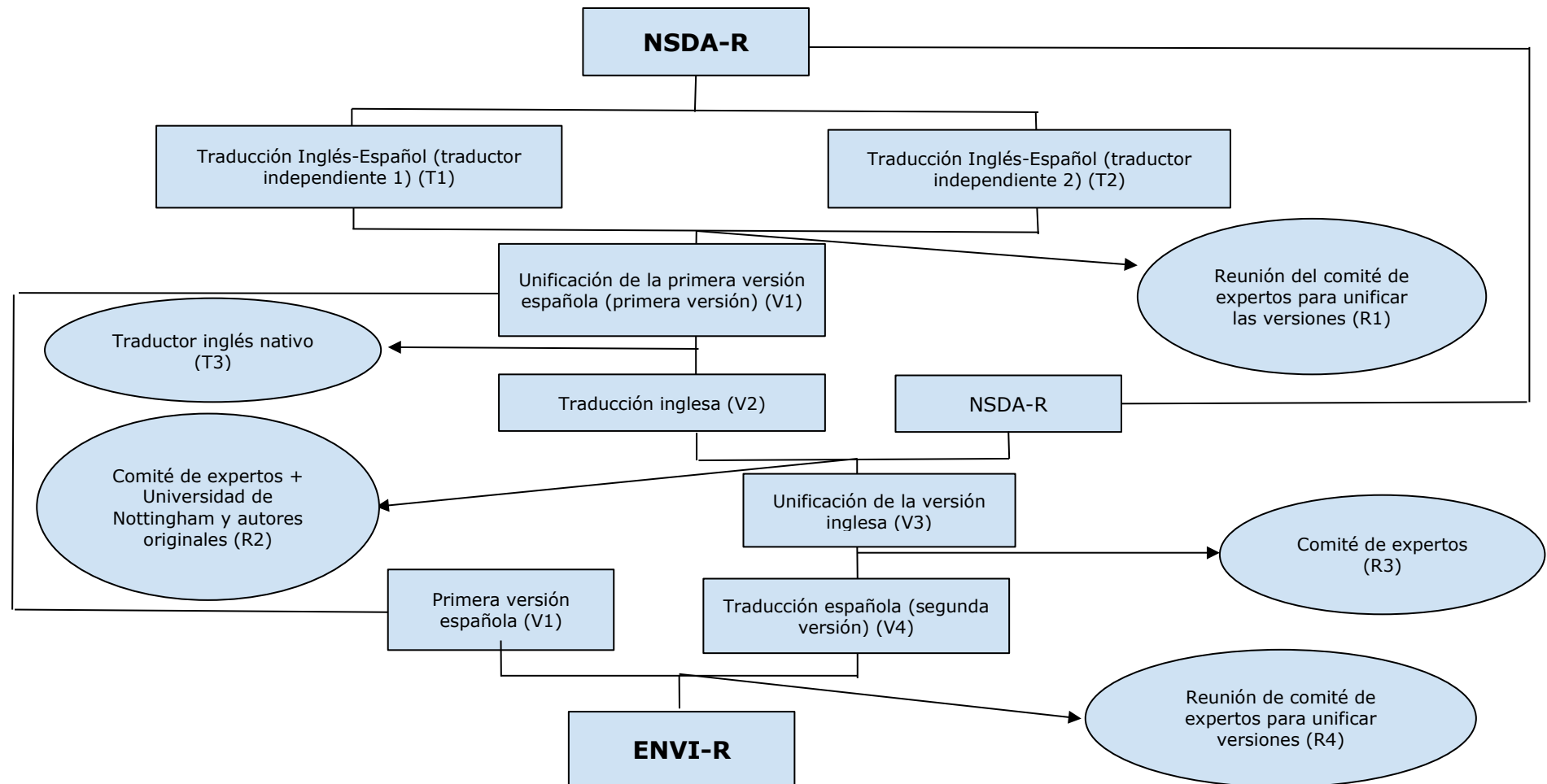
El tercer documento, “Análisis de los Errores de Acción”, presenta una explicación mejorada de cómo ese error, interfiere en cualquiera de los ítems presentados previamente, influyendo a la hora de realizar el vestido. Este documento permite conocer si está presente o ausente, en qué momento aparece, en qué prenda, etc. Aportando información cualitativa a la evaluación.

Muchas de las modificaciones realizadas consistieron principalmente en aclaraciones sobre la forma de expresar un concepto, puesto que en inglés se utilizaban oraciones más breves y concisas; mientras que en español se requerían expresiones más extensas, detalladas y elaboradas.

Los pasos para colocarse la prenda fue lo que menos modificaciones sufrió, ya que la forma de vestirse con las prendas presentadas en la escala no había variado con el paso del tiempo ni era necesario actualizarlo.

La versión actualizada de NSDA se denominó “Nottingham Stroke Dressing Assessment-Revised (NSDA-R)”.

Figura 1: Flujograma del proceso de traducción



Fuente: elaboración propia

Fase 2. Proceso de traducción y adaptación transcultural

La primera versión de NSDA-R en castellano, traducida del inglés por 2 traductores independientes (T1 y T2), mostró algunas diferencias lingüísticas. Estas diferencias fueron abordadas durante la primera revisión (R1) por parte del grupo de expertos y dio como resultado la unificación de ambas versiones en una sola (V1).

Tras la traducción al inglés de vuelta por parte de una persona nativa anglosajona (T3), se comparó con la escala en inglés previamente actualizada (NSDA-R) y, desde el comité de expertos (R2), se aclararon las incongruencias lingüísticas y terminológicas surgidas. Posteriormente, la autora de la versión original y una terapeuta ocupacional involucrada en la traducción al danés revisaron la escala (R2). Ambas resolvieron las discrepancias que presentaba este paso, al tener que comparar dos versiones diferentes de la misma escala, llegando a una unificación de las mismas y facilitando el proceso para alcanzar una versión inglesa mejorada (V3).

Se tradujo de nuevo al castellano de vuelta (V4) y el comité de expertos comparó esa versión de vuelta en castellano con la versión castellana previa (V1). Tras este último paso, se obtuvo la versión española de NSDA-R, denominada "Escala Nottingham del Vestido en Ictus-Revisada" (ENVI-R) (Figura 1).

ENVI-R contiene 12 categorías: 1. Cierres (7 ítems); 2. Sujetador/Top (8 ítems); 3. Camiseta de tirantes (3 ítems); 4. Camiseta, jersey de manga corta (3 ítems); 5. Vestido (3 ítems); 6. Camisa, chaqueta, abrigo (3 ítems); 7. Camiseta, jersey de manga larga (3 ítems); 8. Falda (5 ítems); 9. Bragas, calzoncillos (2 ítems); 10. Pantalones, leggings (2 ítems); 11. Medias, pantys, calcetines (2 ítems); 12. Calzado (3 ítems). Con un total de 44 ítems puntuables de 0 a 2, con una puntuación total posible desde 0% (dependiente total para vestirse) hasta 100% (totalmente independiente para vestirse).

Fase 3. Validación clínica

Durante el período de estudio, que duró 7 meses, se evaluó la elegibilidad de 150 pacientes con un diagnóstico de ictus que formaban parte de los centros de neurorrehabilitación colaboradores. De ellos, 120 pacientes cumplieron los criterios de inclusión y de exclusión. De estos, 10 pacientes se negaron a participar. Se registraron datos demográficos (edad, sexo, centro de rehabilitación), fase de recuperación del ictus (aguda/crónica), fecha de evaluación, puntuación de la escala ENVI-R, puntuación de Fugl-Meyer Assessment de MMSS, puntuación de Fugl-Meyer Assessment de MMII de cada persona participante en el estudio.

Durante esta fase se verificó la ausencia de dificultades en el entendimiento del instrumento de medida por parte de la población diana. ENVI-R demostró ser fácil y comprensible para los participantes, por lo que no se necesitaron cambios adicionales en su formato final.

Análisis factorial exploratorio

La escala en su versión traducida fue sometida a un análisis factorial exploratorio. La puntuación en el test de esfericidad de Bartlett's fue significativa

($X^2 = \infty$; $df = 1035.000$; $p < 0.001$); y para la prueba de Chi-cuadrado también resultó significativa (Valor = 8586.907; $df = 549$; $p < 0.001$). Por tanto, los datos fueron sometidos a un análisis factorial exploratorio con extracción de máxima verosimilitud y rotación Varimax. El Valor propio > 1 y el Screeplot (Figura 2) mostraron una solución de doce factores que explicó un 84% de la varianza total. La carga de cada uno de los ítems en cada factor fue determinada por una carga superior a 0.4. Ninguno de los ítems del cuestionario presentó carga cruzada en más de un factor. Dos ítems fueron eliminados ya que la carga factorial no fue suficiente (<0.30) (Tablas 1 and 2).

Figura 2: Screeplot/Gráfico de sedimentación

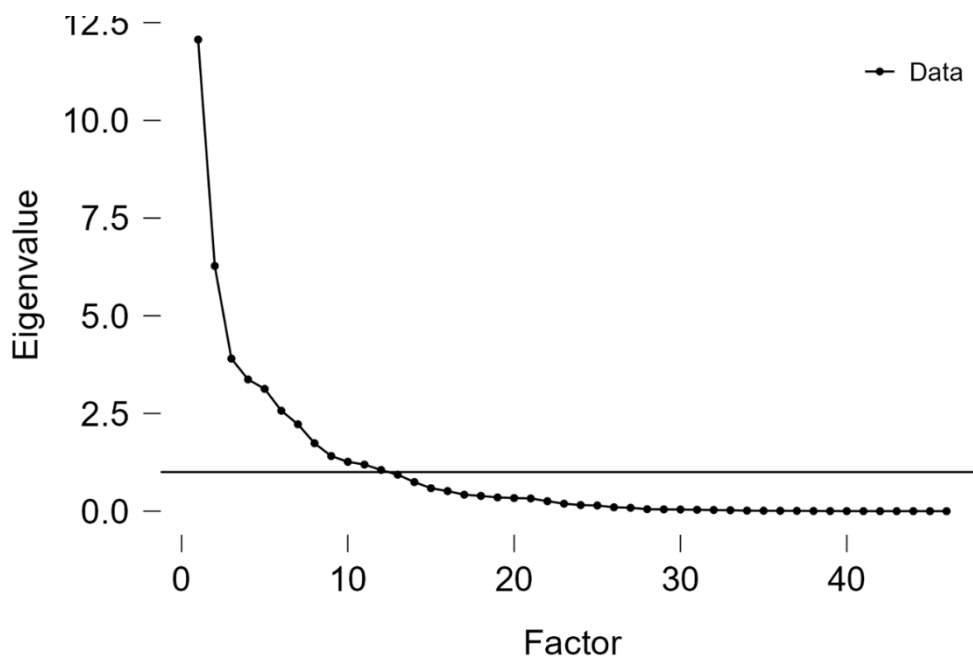


Tabla 1: Características de los factores

Solución rotada				
	Valores propios	Cargas de SumSq	Proporción var.	Acumulativo
Factor 1	12.067	5.198	0.113	0.113
Factor 2	6.275	5.118	0.111	0.224
Factor 3	3.903	3.506	0.076	0.300
Factor 4	3.371	3.447	0.075	0.375
Factor 5	3.129	3.195	0.069	0.445
Factor 6	2.570	3.086	0.067	0.512
Factor 7	2.223	3.064	0.067	0.579
Factor 8	1.737	3.057	0.066	0.645
Factor 9	1.411	2.984	0.065	0.710
Factor 10	1.262	2.431	0.053	0.763
Factor 11	1.192	1.961	0.043	0.805
Factor 12	1.053	1.608	0.035	0.840

Fuente: elaboración propia

Tabla 2: Carga de los factores

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8	Factor 9	Factor 10	Factor 11	Factor 12
ENV 9.1	0.936											
ENV 12.1	0.936											
ENV 9.2	0.906											
ENV 12.2	0.902											
ENV 10.1	0.823											
ENV 10.2	0.796											
ENV 12.3	0.557											
ENV 5.2		1.120										
ENV 5.1		1.059										
ENV 5.3		1.050										
ENV 8a.1		0.852										
ENV 8a.2		0.827										
ENV 8b.1			1.037									
ENV 8b.2			1.037									
ENV 8b.3			1.037									
ENV 1.7			0.588									
ENV 2b.2				1.036								
ENV 2b.3				1.036								
ENV 2b.1				1.036								
ENV 1.4				0.521								
ENV 6.2					1.079							
ENV 6.3					1.074							
ENV 6.1					1.057							
ENV 3.2						1.091						

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

ENV 3.3	1.043								
ENV 3.1	0.992								
ENV 4.2		1.031							
ENV 4.1		1.014							
ENV 4.3		0.949							
ENV 2c.3			0.991						
ENV 2c.2			0.989						
ENV 2c.1			0.986						
ENV 7.2				1.020					
ENV 7.3				0.983					
ENV 7.1				0.981					
ENV 1.2					0.880				
ENV 1.1					0.660				
ENV 1.6					0.540				
ENV 1.5					0.484				
ENV 1.3					0.424				
ENV 2a.1						1.044			
ENV 2a.2						0.978			
ENV 11.2							0.831		
ENV 11.1							0.770		

Fuente: elaboración propia

Consistencia interna

El cuestionario en su versión española presentó buenos valores de consistencia interna totales ($\alpha = 0.921$). En cuanto a la consistencia interna por factores la escala mostró buenos valores de consistencia interna, siendo el máximo para los ítems 2a.1 y 2a.2 ($\alpha = 0.923$) y el mínimo para los ítems 1.8, 3.1, 3.2, 3.3, 6.1 y 6.2 ($\alpha = 0.918$) (Table 3).

Table 3: Alpha de Cronbach's si el ítem es eliminado

	Alpha de Cronbach si el ítem es eliminado
ENV 1.1	0.919
ENV 1.2	0.919
ENV 1.3	0.919
ENV 1.4	0.920
ENV 1.5	0.920
ENV 1.6	0.918
ENV 1.7	0.921
ENV 2a.1	0.923
ENV 2a.2	0.923
ENV 2b.1	0.919
ENV 2b.2	0.919
ENV 2b.3	0.919
ENV 2c.1	0.922
ENV 2c.2	0.922
ENV 2c.3	0.922
ENV 3.1	0.918
ENV 3.2	0.918
ENV 3.3	0.918
ENV 4.1	0.921
ENV 4.2	0.921
ENV 4.3	0.921
ENV 5.1	0.920
ENV 5.2	0.919
ENV 5.3	0.919
ENV 6.1	0.918
ENV 6.2	0.918
ENV 6.3	0.919
ENV 7.1	0.920
ENV 7.2	0.920

ENV 7.3	0.919
ENV 8a.1	0.919
ENV 8a.2	0.919
ENV 8b.1	0.921
ENV 8b.2	0.921
ENV 8b.3	0.921
ENV 9.1	0.920
ENV 9.2	0.920
ENV 10.1	0.919
ENV 10.2	0.919
ENV 11.1	0.919
ENV 11.2	0.919
ENV 12.1	0.919
ENV 12.2	0.919
ENV 12.3	0.920

Fuente: elaboración propia

Fiabilidad test-retest

La fiabilidad test-retest fue medida a través de un índice de correlación intraclase (ICC). La versión española ENVI-R mostró una fiabilidad satisfactoria al aplicarse por segunda vez a 65 pacientes (59.1% de la muestra) al mes y medio de media después de la evaluación inicial, alcanzando un coeficiente de correlación de 0.92 (ICC 95% CI 0.90 a 0.93).

Error de medición

El error estándar de medición mostró una puntuación conveniente, con 0 valores perdidos, y una puntuación de error estándar de medida de 18.03.

Validez de criterio

ENVI-R mostró correlación al ser comparada su puntuación total con la de Fugl-Meyer Assessment tanto de miembro superior (FMA-UE), como de miembro inferior (FMA-LE). Se confirmó correlación significativa entre ENVI-R y FMA de miembro superior ($r=0.34$), de miembro inferior ($r=0.39$), y FMA completa ($r=0.50$) ($p<0.01$).

Versión final de ENVI-R

A partir de la metodología llevada a cabo durante el proceso siguiendo la guía ISPOR, el análisis estadístico de las propiedades psicométricas, y tras el consenso por parte del comité de expertos, se realizaron los cambios menores para refinar y terminar de perfilar la versión española final de ENVI-R.

De esta manera, se pudo finalizar la validación de la escala ENVI-R al castellano.

4.DISCUSIÓN

Actualmente existen tres escalas específicas de evaluación del vestido, NSDA, Jacket Test y UBDS. NSDA ha sido más ampliamente empleada y estudiada, pero solo en población con ictus^{12,14-19,30}. UBDS surgió más recientemente, y se ha estudiado tanto en ictus²⁷⁻²⁹, como en demencia vascular con hemiparesia²⁹. La escala Physical Performance Test (PPT) consta de 9 ítems; uno de ellos se extrajo para utilizarlo de forma independiente, generando la prueba Jacket Test. De esta prueba se han evaluado sus propiedades psicométricas, con el fin de conocer su fiabilidad en pacientes con ictus crónico³¹. Resultan escasas las escalas de evaluación del vestido, dada la importancia de esta AVD para la independencia y la calidad de vida^{11,12,14-16,19,46}. Añadido a esto, las tres han sido empleadas en inglés, sin existir traducciones ni validaciones de ninguna de ellas a otros idiomas.

Actualmente en España no se dispone de ninguna escala de evaluación del vestido para ninguna población en concreto, ya sea pediátrica, neurológica, psiquiátrica, traumática, neurológica, ni para una población genérica. Para conocer el desempeño en esta actividad suelen emplearse escalas como el Índice de Barthel, FIM o MAL-30, las cuales incluyen solamente un ítem que evalúa el vestido. Sin embargo, dichos instrumentos no son específicos para la evaluación del vestido, sino que valoran otros constructos como la independencia o la cantidad y calidad del uso del miembro superior en AVD²²⁻²⁷. Debido a esta ausencia de evaluaciones del vestido en español, y a la existencia de tres de ellas en inglés, se planteó como necesaria la traducción y validación de escalas que permitieran evaluar esta actividad en población española.

El objetivo del presente estudio fue desarrollar una revisión y actualización de la escala NSDA para su posterior traducción y adaptación transcultural al español (ENVI-R), y su validación en una muestra de pacientes post-ictus. Para ello, se adoptó un proceso riguroso estandarizado de revisión y actualización, y una posterior traducción, adaptación transcultural y validación clínica³⁵⁻³⁷, para desarrollar una actualización de la escala NSDA a NSDA-R, y su consiguiente traducción oficial al español de NSDA-R a ENVI-R.

A pesar de ser una escala que solo se ha empleado en Reino Unido y únicamente en población con ictus, es la herramienta de evaluación específica del vestido más estudiada y empleada a día de hoy, debido a que analiza dicha actividad al detalle, considerando muchos factores que no se tienen en cuenta en otras escalas^{12,14-19,30}.

La validación clínica de la versión española de ENVI-R fue sometida a un análisis estadístico de las propiedades psicométricas (análisis factorial exploratorio, consistencia interna, fiabilidad, error de medición, validez de criterio, prueba de hipótesis y sensibilidad), mostrando buena concordancia para la mayoría de los ítems. Dichas propiedades estadísticas fueron adecuadas y significativas⁴⁷⁻⁴⁹. La validación clínica proporcionó información para los ajustes finales realizados a la versión española oficial desarrollada de ENVI-R.

Ofreciendo una versión altamente válida y recomendada para la evaluación de la capacidad y el desempeño del vestido en personas que han sufrido un ictus.

Tras someter ENVI-R a un proceso de traducción-retrotraducción y de adaptación cultural se verificó la ausencia de dificultades para el entendimiento de los ítems de la versión española de ENVI-R, mostrándose el cuestionario como un instrumento fácil y comprensible para la población diana.

Este artículo supone, a conocimiento de los autores, el primer análisis factorial de ENVI-R, dado que no existen otras traducciones ni validaciones, y las propiedades psicométricas analizadas anteriormente fueron fiabilidad^{19,30} y nivel de acuerdo¹⁹. El EFA arrojó un modelo de doce factores, rotado con Varimax que explicó el 84% de la varianza total⁵⁰. Previa realización del EFA se comprobó que la muestra fuera adecuada. Esta varianza resulta mayor a la mostrada en la escala UBDS, en la cual se hallaron 2 componentes, con mayor carga hacia el primer componente (0.65-0.83) y un porcentaje de varianza total de 62.3%²⁸. Esto posiblemente se debe a que UBDS es una escala más reducida, que solo evalúa una prenda, por lo que ofrece menos posibilidades para que surjan varios factores entre sus ítems, mientras que NSDA cuenta con 12 categorías, teniendo en cuenta un gran número de prendas, y de pasos para la colocación de las mismas, con 44 ítems, lo que causa que, en este análisis, la carga de los ítems se distribuya en 12 factores donde se explique el reparto del 84% de los ítems. A su vez, todos los ítems mostraron una carga factorial adecuada superior a 0.3, excepto 2 ítems del apartado de cierres que fueron eliminados tras mostrar una carga inferior a 0.3, los cuales fueron “Velcro” y “Botón de presión”.

En cuanto a la consistencia interna de ENVI-R se encontraron excelentes niveles de consistencia interna⁴⁷, con una puntuación alta de Alfa de Cronbach ($\alpha = 0.921$) para la escala en general. Estos valores resultan superiores a los de la escala UBDS, ($\alpha = 0.88$). Esto se debe a que la escala ENVI-R ha seguido un proceso riguroso de traducción y validación que permite que siga midiendo el constructo original de manera precisa, tanto de forma general, como de cada uno de sus ítems. Estos ítems han presentado puntuaciones altamente significativas variando de $\alpha = 0.918$ a $\alpha = 0.923$; lo cual solo se ha analizado en este estudio de la escala ENVI-R.

La fiabilidad test-retest del cuestionario en español fue evaluada a través de un ICC, el cual mostró propiedades de fiabilidad excelentes^{51,52}, con puntuación de 0.92 (ICC 95% CI 0.90 a 0.93). Esta puntuación resulta ligeramente menor a la obtenida en NSDA (ICC=0.9988)^{19,30}, pero igualmente elevada y significativa. Al compararse con otras escalas que permiten evaluar el vestido, como UBDS, se observan valores de fiabilidad similares para la misma población (ictus), mostrando UBDS un ICC de 0.87²⁸, de 1.00 para inter-evaluador²⁷, y de 0.76 a 0.91 para la fiabilidad test-retest²⁷. Dicho rango se debe a que se evaluó por separado la subescala de tiempo de UBDS, y la subescala de puntuación de UBDS. No obstante, ambas escalas han presentado propiedades de fiabilidad excelentes. También se puede comparar con la fiabilidad mostrada por el Jacket Test, del cual se evaluaron sus propiedades psicométricas en un estudio realizado en población con ictus crónico³¹, donde se halló una puntuación de fiabilidad alta (ICC=0.78-1.00). Esto concuerda con lo mostrado previamente,

sugiriendo que cualquiera de las 3 pruebas sería fiable para evaluar el vestido. Difiere únicamente en la forma de evaluación, ya que ENVI-R analiza el vestido de todas las prendas por pasos, mientras que UBDS evalúa solo el vestido de una camisa por pasos, y el Jacket Test valora el vestido y desvestido de una chaqueta.

Los autores utilizaron la puntuación total de ENVI-R y de FMA para medir la validez de criterio, mostrando en ambos casos correlaciones positivas (Spearman rho: ENVI-R vs FMA-UE = 0.34; ENVI-R vs FMA-LE = 0.39; ENVI-R vs FMA = 0.50), de manera que las variables varían, sin implicar causalidad, y que la distribución de la variable es anormal^{153,54}. Otros estudios también evaluaron la correlación entre sus escalas del vestido y Fugl-Meyer Assessment, como en el caso de UBDS, que halló correlación moderada negativa significativa tanto con FMA-UE ($r=-0.58$) como con FMA-LE ($r=-0.18$)²⁷. En el caso de Jacket Test, la correlación con FMA fue de $r=-0.285$ para el lado afecto, y de $r=-0.75$ para el lado sano³¹. Estas puntuaciones de correlación son significativamente moderadas para las 3 escalas, debido a que FMA no evalúa el vestido, sino la función sensoriomotora del miembro superior y del miembro inferior. La diferencia de signo positivo a negativo puede deberse a la diferente forma de puntuación que hay entre las tres escalas. No se han podido establecer correlaciones entre escalas que evalúen el mismo constructo (vestido en personas tras ictus), debido a que hay muy pocas escalas que evalúen esta variable, y en el caso de ENVI-R, no se ha podido comparar por no estar validadas UBDS y Jacket Test en el idioma al que ha sido traducido, adaptado y validado, es decir, al español. En el caso de UBDS sí que se ha comparado con el ítem de vestido de la escala FIM, mostrando una correlación mayor, de $r=-0.72$; lo que muestra que la variable a evaluar (el vestido) es más similar. No obstante, esta correlación debe tomarse con cautela, ya que FIM no ha sido diseñada para evaluar el desempeño en el vestido, sino la independencia en actividades de la vida diaria.

Mediante esta correlación se pudo confirmar la hipótesis planteada en este estudio, y lograr una adecuada sensibilidad para la escala ENVI-R.

Observando los diferentes hallazgos y la comparación con otras escalas, ENVI-R presenta valores similares para algunas propiedades psicométricas como fiabilidad o validez de criterio, e incluso superiores, como es el caso de la validez de criterio y del análisis factorial. Con estos hallazgos, se puede afirmar que ENVI-R podría ser actualmente la única escala en español que evalúe de manera específica y exhaustiva el vestido en personas que han sufrido un ictus, presentando buenas propiedades psicométricas tras el proceso de traducción, adaptación transcultural y validación, por lo que se recomienda su uso en población española que haya sufrido un ictus y que presente dificultades para el vestido.

En un futuro, sería recomendable poder generalizar y validar la escala ENVI-R a personas con otras patologías que causen limitaciones en la AVD del vestido.

Limitaciones y fortalezas del estudio

Una limitación de este estudio fue que el tamaño de la muestra de la validación no fue muy amplio, y además fue obtenida a través de un muestreo por conveniencia en la que el paciente aceptaba participar en su centro de rehabilitación tras la invitación por parte del investigador para implicarse en el estudio. No realizar las evaluaciones por 2 evaluadores independientes, impidió evaluar la fiabilidad inter-evaluador⁵⁵. Otra de las limitaciones fue que no se pudo evaluar la validez de criterio de la escala ENVI-R, debido a que las escalas UBDS y Jacket Test no han sido traducidas, adaptadas transculturalmente y validadas en población española. Se necesitarían más escalas que permitan evaluar el vestido, así como la traducción, adaptación transcultural y validación de estas escalas, para permitir evaluar la validez de criterio entre escalas que evalúen el mismo constructo en la misma población. Una de las propuestas a futuro sería validar la versión actualizada de la escala original en el idioma original (NSDA-R) en población inglesa. Estas limitaciones sugieren precaución al utilizar tanto NSDA-R, como ENVI-R, antes de realizar una evaluación más extensa de las propiedades psicométricas en una población más amplia, tanto española, como inglesa.

Como principal fortaleza se destaca que este artículo es la primera actualización de la escala NSDA a NSDA-R. Al igual que resulta la primera traducción, adaptación transcultural y validación de NSDA-R a ENVI-R en población y lengua española, el cual es un lenguaje ampliamente hablado a nivel global. Además, ningún estudio previo analizó muchas de las propiedades psicométricas estudiadas en esta investigación, ni para NSDA, ni NSDA-R, ni ENVI-R.

5.CONCLUSIONES

El proceso de actualización, de traducción y adaptación transcultural, y posterior validación, demostraron que la versión final de NSDA-R y ENVI-R son herramientas que podrían ser empleadas en poblaciones inglesa y española respectivamente, para la valoración de la actividad del vestido. También se verificó que los valores obtenidos para las propiedades psicométricas fueron adecuados y significativos en el uso de ENVI-R en una muestra de personas que sufren ictus en población española. Basándose en estos resultados, se puede recomendar y defender el uso de la actualización (NSDA-R), y de la traducción oficial española (ENVI-R) en la investigación y en la práctica clínica. El uso de esta escala actualizada, traducida y validada en castellano permitirá identificar disfunciones ocupacionales en la actividad del vestido en población que ha sufrido un ictus y, por lo tanto, permitirá una comparación confiable del resultado del accidente cerebrovascular entre regiones, tanto en la práctica clínica, como en la investigación.

Agradecimientos

Los autores agradecen a todos los pacientes que participaron en este estudio y a los centros colaboradores que permitieron realizar las evaluaciones en sus instalaciones, como Integra Daño Cerebral, Dacemur, CMVCaridad Puerto Lumbreras, Asociación DISMO, CIAN Grupo 5, Neuroavanzo, Caser Salud, Neuron, y Fundación AISSE. Gracias también a la Universidad de Nottingham, así como a los autores originales y a todas las personas involucradas en esta investigación.

Declaración de conflicto de intereses

Los autores no informaron de ningún conflicto de intereses potencial.

Contribuciones de los autores

Todos los autores han acordado la versión final y cumplen al menos uno de los siguientes criterios (recomendados por el ICMJE): Contribuciones sustanciales a la concepción y el diseño, la adquisición de datos o el análisis e interpretación de datos; la redacción del artículo; o su revisión crítica para determinar su contenido intelectual importante.

Fondos

Esta investigación no recibió ninguna subvención específica de agencias de financiación de los sectores público, comercial o sin fines de lucro.

6.REFERENCIAS

1.
Feigin VL, Owolabi MO, World Stroke Organization–Lancet Neurology Commission Stroke Collaboration Group. Pragmatic solutions to reduce the global burden of stroke: a World Stroke Organization-Lancet Neurology Commission. *Lancet Neurol.* 2023 Dec;22(12):1160-206.
2.
Hua X, Liu M, Wu S. Definition, prediction, prevention and management of patients with severe ischemic stroke and large infarction. *Chin Med J (Engl).* 2023 Dec;136(24):2912-22.
3.
Katan M, Luft A. Global Burden of Stroke. *Semin Neurol.* 2018 Apr;38(2):208-11.
4.
Masjuan J, Alvarez-Sabín J, Arenillas J, Calleja S, Castillo J, Dávalos A, et al. [Stroke health care plan (ICTUS II. 2010)]. *Neurología.* 2011 Sep;26(7):383-96.
5.
Díaz-Guzmán J, Egido JA, Gabriel-Sánchez R, Barberá-Comes G, Fuentes-Gimeno B, Fernández-Pérez C, et al. Stroke and transient ischemic attack incidence rate in Spain: the IBERICTUS study. *Cerebrovasc Dis.* 2012;34(4):272-81.
6.
Soleimani M, Ghazisaeedi M, Heydari S. The efficacy of virtual reality for upper limb rehabilitation in stroke patients: a systematic review and meta-analysis. *BMC Med Inform Decis Mak.* 2024 May;24(1):135.
7.
Ingram LA, Butler AA, Lord SR, Gandevia SC. Use of a physiological profile to document upper limb motor impairment in ageing and in neurological conditions. *J Physiol.* 2023 Jun;601(12):2251-62.
8.
Li S, Ghuman J, Gonzalez-Buonomo J, Huang X, Malik A, Yozbatiran N, et al. Does Spasticity Correlate With Motor Impairment in the Upper and Lower Limbs in Ambulatory Chronic Stroke Survivors? *Am J Phys Med Rehabil.* 2023 Oct;102(10):907-12.
9.
Zhao K, He C, Xiang W, Zhou Y, Zhang Z, Li J, et al. Evidence of synergy coordination patterns of upper-limb motor control in stroke patients with mild and moderate impairment. *Front Physiol.* 2023;14:1214995.

10.

Dawson J, Engineer ND, Cramer SC, Wolf SL, Ali R, O'Dell MW, et al. Vagus Nerve Stimulation Paired With Rehabilitation for Upper Limb Motor Impairment and Function After Chronic Ischemic Stroke: Subgroup Analysis of the Randomized, Blinded, Pivotal, VNS-REHAB Device Trial. *Neurorehabil Neural Repair*. 2023 Jun;37(6):367-73.

11.

Fujita T, Sato A, Yamamoto Y, Yamane K, Otsuki K, Tsuchiya K, et al. Relationship between dressing and motor function in stroke patients: a study with partial correlation analysis. *J Phys Ther Sci*. 2015 Dec;27(12):3771-4.

12.

Walker MF, Sunderland A, Fletcher-Smith J, Drummond A, Logan P, Edmans JA, et al. The DRESS trial: a feasibility randomized controlled trial of a neuropsychological approach to dressing therapy for stroke inpatients. *Clin Rehabil*. 2012 Aug;26(8):675-85.

13.

Sunderland A, Walker CM, Walker MF. Action errors and dressing disability after stroke: an ecological approach to neuropsychological assessment and intervention. *Neuropsychol Rehabil*. 2006 Dec;16(6):666-83.

14.

Walker CM, Sunderland A, Sharma J, Walker MF. The impact of cognitive impairment on upper body dressing difficulties after stroke: a video analysis of patterns of recovery. *J Neurol Neurosurg Psychiatry*. 2004 Jan;75(1):43-8.

15.

Walker M. Dressing ability after stroke. *Nurs Times*. 1992 Jan;88(1):51.

16.

Walker MF, Lincoln NB. Factors influencing dressing performance after stroke. *J Neurol Neurosurg Psychiatry*. 1991 Aug;54(8):699-701.

17.

Walker MF, Lincoln NB. Reacquisition of dressing skills after stroke. *Int Disabil Stud*. 1990;12(1):41-3.

18.

Walker CM, Walker MF, Sunderland A. Dressing after a Stroke: A Survey of Current Occupational Therapy Practice. *British Journal of Occupational Therapy*. 2003 Jun;66(6):263-8.

19.

Fletcher-Smith JC. Recovery of dressing ability after stroke [Internet]. [Nottingham, UK]: University of Nottingham; 2011. Available from: <https://eprints.nottingham.ac.uk/11913/>

20.

Walker CM, Walker MF. Dressing Ability after Stroke: A Review of the Literature. *British Journal of Occupational Therapy*. 2001 Sep;64(9):449–54.

21.

Cecchi F, Carrabba C, Bertolucci F, Castagnoli C, Falsini C, Gnetti B, et al. Transcultural translation and validation of Fugl-Meyer assessment to Italian. *Disabil Rehabil*. 2021 Dec;43(25):3717-22.

22.

Colomer C, Llorens R, Sánchez C, Ugart P, Moliner B, Navarro MD, et al. Reliability and validity of the Spanish adaptation of the Functional Independence Measure + Functional Assessment Measure. *Eur J Phys Rehabil Med*. 2023 Aug;59(4):452-7.

23.

Jhong M chi, Tang NY, Liu CH, Huang WH, Hsu YT, Liu YL, et al. Relationship between Chinese medicine pattern types, clinical severity, and prognosis in patients with acute cerebral infarct. *Explore (NY)*. 2013;9(4):226-31.

24.

Langhammer B, Sunnerhagen KS, Lundgren-Nilsson Å, Sällström S, Becker F, Stanghelle JK. Factors enhancing activities of daily living after stroke in specialized rehabilitation: an observational multicenter study within the Sunnaas International Network. *Eur J Phys Rehabil Med*. 2017 Oct;53(5):725-34.

25.

Miki E, Yamane S, Yamaoka M, Fujii H, Ueno H, Kawahara T, et al. Validity and reliability of the Japanese version of the FIM+FAM in patients with cerebrovascular accident. *Scand J Occup Ther*. 2016 Sep;23(5):398-404.

26.

Sarker SJ, Rudd AG, Douiri A, Wolfe CDA. Comparison of 2 extended activities of daily living scales with the Barthel Index and predictors of their outcomes: cohort study within the South London Stroke Register (SLSR). *Stroke*. 2012 May;43(5):1362-9.

27.

Ng SSM, Chen P, Chang HS, Chun WK, Kong TH, Lam Y, et al. Psychometric properties of Upper-body Dressing Scale in people with stroke. *J Rehabil Med*. 2023 Apr;55:jrm00391.

28.

Suzuki M, Yamada S, Omori M, Hatakeyama M, Sugimura Y, Matsushita K, et al. Development of the upper-body dressing scale for a buttoned shirt: a preliminary correlational study. *Am J Phys Med Rehabil*. 2008 Sep;87(9):740-9.

29.

Endo A, Suzuki M, Akagi A, Chiba N, Ishizaka I, Matsunaga A, et al. Reliability and validity of the upper-body dressing scale in Japanese patients with vascular dementia with hemiparesis. *Occup Ther Int*. 2015 Mar;22(1):10-8.

30.

Fletcher-Smith J, Walker M, Sunderland A, Garvey K, Wan A, Turner H. An Interrater Reliability Study of the Nottingham Stroke Dressing Assessment. *British Journal of Occupational Therapy*. 2010 Dec;73(12):570-8.

31.

Chen P, Lai CKY, Chung RCK, Ng SSM. The Jacket Test for assessing people with chronic stroke. *Disabil Rehabil*. 2017 Dec;39(25):2577-83.

32.

Gjersing L, Caplehorn JRM, Clausen T. Cross-cultural adaptation of research instruments: language, setting, time and statistical considerations. *BMC Med Res Methodol*. 2010 Feb;10:13.

33.

Reichenheim ME, Moraes CL. [Operationalizing the cross-cultural adaptation of epidemiological measurement instruments]. *Rev Saude Publica*. 2007 Aug;41(4):665-73.

34.

Profesionales sanitarios colegiados por sexo. Series desde 1952. Terapeutas Ocupacionales colegiados por año y sexo. [Internet]. España: Instituto Nacional de Estadística (INE); 2023. Available from: <https://www.ine.es/jaxi/Datos.htm?tpx=30732>

35.

Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, et al. Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Health*. 2005;8(2):94-104.

36.

Muñiz J, Elosua P, Hambleton RK, International Test Commission. [International Test Commission Guidelines for test translation and adaptation: second edition]. *Psicothema*. 2013;25(2):151-7.

37.

Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976)*. 2000 Dec;25(24):3186-91.

38.

Bushnell C, Bettger JP, Cockcroft KM, Cramer SC, Edelen MO, Hanley D, et al. Chronic Stroke Outcome Measures for Motor Function Intervention Trials: Expert Panel Recommendations. *Circ Cardiovasc Qual Outcomes*. 2015 Oct;8(6 Suppl 3):S163-169.

39.

Santisteban L, Térémetz M, Bleton JP, Baron JC, Maier MA, Lindberg PG. Upper Limb Outcome Measures Used in Stroke Rehabilitation Studies: A Systematic Literature Review. *PLoS One*. 2016;11(5):e0154792.

40.

Yu B, Kaku A, Liu K, Parnandi A, Fokas E, Venkatesan A, et al. Quantifying impairment and disease severity using AI models trained on healthy subjects. *NPJ Digit Med*. 2024 Jul;7(1):180.

41.

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*. mayo de 2018;27(5):1147-57.

42.

Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res*. mayo de 2010;19(4):539-49.

43.

Elsman EBM, Mokkink LB, Terwee CB, Beaton D, Gagnier JJ, Tricco AC, et al. Guideline for reporting systematic reviews of outcome measurement instruments (OMIs): PRISMA-COSMIN for OMIs 2024. *J Clin Epidemiol*. 3 de julio de 2024;111422.

44.

Terwee CB, Mokkink LB, Knol DL, Ostelo RWJG, Bouter LM, De Vet HCW. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res*. mayo de 2012;21(4):651-7.

45.

Cronbach L. Coefficient alpha and the internal structure of tests. *Psyko Psykometrika*. 1951 Sep;16(3):2997-334.

46.

Walker M, Drummond A, Lincoln N. Evaluation of dressing practice for stroke patients after discharge from hospital: a crossover design study. *Clin Rehabil*. 1996 Feb;10(1):23-31.

47.

Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med*. 2016 Jun;15(2):155-63.

48.

De Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine: A Practical Guide* [Internet]. 1.^a ed. Cambridge University Press; 2011 [citado 26 de octubre de 2024]. Available from: <https://www.cambridge.org/core/product/identifier/9780511996214/type/book>

49.

Moral De La Rubia J. Revisión de los criterios para validez convergente estimada a través de la Varianza Media Extraída. *Psychol*. 2019 Nov;13(2):25-41.

50.

Méndez Martínez C, Rodón epúlveda MA. Introducción al análisis factorial exploratorio. *Revista Colombiana de Psiquiatría*. 2012;41(1):197-207.

51.

Martínez Pérez JA, Pérez Martin PS. [Intraclass correlation coefficient]. *Semergen*. 2023 Apr;49(3):101907.

52.

Liljequist D, Elfving B, Skavberg Roaldsen K. Intraclass correlation - A discussion and demonstration of basic features. *PLoS One*. 2019;14(7):e0219854.

53.

Akoglu H. User's guide to correlation coefficients. *Turk J Emerg Med*. 2018 Sept;18(3):91-3.

54.

Schober P, Boer C, Schwarte LA. Correlation Coefficients: Appropriate Use and Interpretation. *Anesth Analg*. 2018 May;126(5):1763-8.

55.

Luján Tangarife J, Crdona Arias J. Construcción y validación de escalas de medición en salud: revisión de propiedades psicométricas. *Archivos de Medicina*. 2015;11(3.1).

6.2.2 VALIDACIÓN DE LA VERSIÓN ESPAÑOLA DE FUGL-MEYER ASSESSMENT EN POBLACIÓN ESPAÑOLA TRAS UN ICTUS

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

VALIDACIÓN DE LA VERSIÓN ESPAÑOLA DE FUGL-MEYER
ASSESSMENT EN POBLACIÓN ESPAÑOLA TRAS UN ICTUS

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

El Artículo IV se encuentra bajo revisión y pendiente para su publicación en la revista “Disability and Health Journal”.

ABSTRACT

Introduction: The Fugl-Meyer Assessment (FMA) is one of the most recommended scales for assessing upper and lower limb sensorimotor function in stroke survivors. This study aims to evaluate the psychometric properties of the Spanish version of the FMA and validate it in the Spanish stroke population.

Methods: To validate the Spanish version of the FMA, 110 stroke survivors were assessed twice. Psychometric properties were evaluated using exploratory factor analysis, internal consistency, reliability, measurement error, criterion validity, hypothesis testing, and responsiveness.

Results: The exploratory analysis generated 7 factors. Internal consistency was excellent ($\alpha = 0.982$). Test-retest reliability was excellent (ICC = 0.98). The criterion validity found between FMA-Total and FMA-UE was 0.97, and between FMA-Total and FMA-LE, 0.89. The correlation between FMA and ENVI-R was demonstrated and significant ($r = 0.34-0.5$).

Conclusion: The Spanish version of the FMA can be used reliably and validly in research and clinical practice after demonstrating adequate and significant psychometric properties. The use of this scale will improve the quality of assessment of upper and lower limb sensorimotor function in stroke survivors and allow for reliable comparisons.

Keywords: Fugl-Meyer Assessment, sensorimotor function, assessment, psychometric properties, validation.

RESUMEN

Introducción: La Fugl-Meyer Assessment (FMA) es una de las escalas más recomendadas para evaluar la función sensoriomotora de miembros superiores e inferiores en personas que han sufrido un ictus. Este estudio pretende evaluar las propiedades psicométricas de la versión española de FMA, para validarla en población española que ha sufrido un ictus.

Métodos: Para validar la versión española de FMA, se evaluó en dos ocasiones a ciento diez personas que habían sufrido un ictus. Las propiedades psicométricas fueron evaluadas con análisis factorial exploratorio, consistencia interna, fiabilidad, error de medición, validez de criterio, prueba de hipótesis y sensibilidad.

Resultados: el análisis exploratorio generó 7 factores. La consistencia interna fue excelente ($\alpha = 0,982$). La fiabilidad test-retest fue excelente ($ICC=0.98$). La validez de criterio encontrada entre FMA-Total y FMA-UE fue de 0,97, y entre FMA-Total y FMA-LE de 0,89. La evaluación entre FMA y ENVI-R fue demostrada y significativa ($r=0.34-0.5$).

Conclusión: la versión española de FMA se puede utilizar de forma fiable y válida en investigaciones y en la práctica clínica tras mostrar propiedades psicométricas adecuadas y significativas. El uso de esta escala mejorará la calidad de la evaluación de la función sensoriomotora de miembros superiores e inferiores en personas que han sufrido un ictus y permitirá comparaciones fiables.

Palabras clave: Fugl-Meyer Assessment, función sensoriomotora, evaluación, propiedades psicométricas, validación

1.INTRODUCCIÓN

El ictus es la segunda causa de muerte¹ y la primera causa de discapacidad a nivel mundial². Quince millones de personas sufren un ictus cada año en todo el mundo; de ellas, 5 millones fallecen y otros 5 millones quedan con discapacidad permanente, lo que supone una carga para la familia y la comunidad³.

La carga de discapacidad tras un ictus es significativa y aumenta rápidamente, y su incidencia está aumentando en personas jóvenes y de mediana edad menores de 55 años a nivel mundial¹. Se estima que la cantidad de ictus aumentará un 50 % en los próximos años, pasando de 6,6 millones de casos en 2020 a 9,7 millones en 2050¹.

En España, los ACV son la primera causa de discapacidad en el adulto y la segunda causa de demencia⁴. En 2023 se registraron 23.173 defunciones por ictus, de las cuales 12.817 fueron mujeres⁵. Los estudios estiman que la incidencia de ictus es de 187 casos por cada 100.000 habitantes/año, mayor que en Inglaterra, Alemania y Francia, pero menor que en otros países de nuestro entorno como Italia, Grecia, o Portugal⁵.

Las limitaciones físicas son secuelas comunes tras un ictus, que se manifiestan en diversos dominios fisiológicos y aspectos motores, como la espasticidad, la sinergia muscular, la coordinación, la fuerza, la amplitud de movimiento, la destreza, la sensibilidad cutánea, la propiocepción y la estabilidad del brazo⁶⁻⁸. Todas ellas provocan la incapacidad de usar la extremidad superior para alcanzar, agarrar y manipular objetos, lo que afecta la realización de tareas de la vida diaria como vestirse, alimentarse o bañarse^{6,9} y tiene un impacto negativo en la calidad de vida⁸.

Con estos datos, es necesario disponer de herramientas para evaluar los diferentes déficits físicos causados por el ictus. Existen numerosas herramientas para evaluar el constructo sensoriomotor y funcional de las personas con ictus. Algunas de ellas son Fugl-Meyer Assessment (FMA), Action Research Arm Test (ARAT), Wolf Motor Function Test (WMFT), Box and Block Test (BBT), etc.¹⁰⁻¹².

Elegir el instrumento más adecuado para evaluar la función sensoriomotora de las extremidades superiores e inferiores en personas con ictus es un desafío, y la elección debe depender de las propiedades de medición y la viabilidad del instrumento¹³. Algunas revisiones sistemáticas demuestran que la FMA es la más utilizada en la práctica clínica y la investigación, y su uso ha aumentado en los últimos años^{11, 14, 15}.

La FMA es conocida internacionalmente, con diferentes versiones y traducciones. Ha sido traducida a diferentes idiomas y culturas^{10-12, 15-17}.

La FMA es el “gold standard” para evaluar la función sensoriomotora de la extremidad superior, por su capacidad de evaluar aspectos como el movimiento dentro de sinergias, mezcla de sinergias, reflejos, movimientos de muñeca, mano, agarre, coordinación y velocidad; aportando así gran cantidad de información de gran utilidad para comprender la capacidad sensoriomotora de la extremidad superior afectada tras un ictus^{11, 12, 14}.

Una revisión reciente ha evaluado las propiedades psicométricas de 22 versiones y traducciones de FMA, comparándolas según la metodología

COSMIN, hallando que la versión completa de FMA de 50 ítems, es la más válida para evaluar la función sensoriomotora del miembro superior en población post-ictus¹⁸.

Tras revisar y analizar la revisión sistemática de las diferentes versiones de FMA se ha comprobado que no se han evaluado propiedades psicométricas de la versión española de FMA¹⁹, a excepción de content validity y cross-cultural validity.

La importancia de validar los instrumentos de evaluación a los entornos de investigación y práctica clínica actuales es universalmente reconocida y necesaria para garantizar que los conceptos explorados por una herramienta sean iguales entre el contexto original y de destino; de tal manera que la validación de estos en el contexto de destino permite el uso de esa herramienta por los terapeutas de dicho país, y poder ofrecer una mejor intervención²⁰. Estos aspectos se hacen más relevantes en el caso de la neurorrehabilitación de miembros superiores e inferiores, los cuales cuentan con menor número de escalas de valoración en comparación con otras especialidades o constructos, y en su mayor parte en idiomas distintos al castellano, por lo que, o no se pueden utilizar en España, o su uso se encuentra sesgado, por falta de validación.

El objetivo de este estudio fue validar la versión española de FMA, disponible en la web de la Universidad de Gothenburg (<https://www.gu.se/en/neuroscience-physiology/fugl-meyer-assessment#spanish>), en una muestra de personas españolas post-ictus.

2.MATERIAL Y MÉTODOS

Diseño del estudio

Estudio de validación de la versión española de la escala FMA en población española que ha sufrido un ictus.

Fugl-Meyer Assessment

Fugl-Meyer Assessment es una evaluación de déficit que consta de 155 ítems, calificados en una escala ordinal de 3 puntos (0-2). 33 ítems para extremidades superiores, 17 ítems para extremidades inferiores, 7 ítems para equilibrio, 12 ítems para sensación y 22 ítems para movimiento articular pasivo y dolor articular.

La escala FMA incluye tres categorías: (1) función motora, (2) equilibrio y (3) sensibilidad. La puntuación máxima para cada una de estas categorías y subcategorías fue la siguiente: función motora de las extremidades superiores (66 puntos) e inferiores (34 puntos); equilibrio en sedestación (6 puntos) y de pie (8 puntos); tacto (24 puntos) y sensibilidad a la posición (16 puntos).

La parte motora de la escala incluye 50 ítems de función motora de las extremidades superiores e inferiores, con un rango de puntuación posible de 0 a 100 puntos (66 de la subescala de extremidades superiores y 34 de la subescala de extremidades inferiores^{18, 21}).

Proceso de validación

Para validar la versión española de FMA se administró dicha escala a 110 personas que habían sufrido un ictus.

Los participantes fueron reclutados de manera continua entre los meses de febrero y septiembre de 2024, a lo largo de 9 centros de neurorrehabilitación privados y concertados repartidos por el territorio español y contando con participantes que se encontraban tanto en fase aguda como en fase crónica que cumplieron los siguientes criterios de inclusión y de exclusión.

Criterios de inclusión: a) Personas mayores de edad (18-99 años); b) Haber sufrido, al menos, un ictus; c) Ser capaz de mover activamente (Daniels 4/5) el hemicuerpo “sano”.

Criterios de exclusión: a) Ausencia/Disminución de respuesta al medio (Glasgow inferior a 15); b) Incapacidad para comprender las instrucciones del terapeuta; c) Estar conectado a un dispositivo de soporte vital básico; d) Agresividad o suponer un riesgo para sí mismo o los demás; e) Estar encamado.

A todas las personas potencialmente elegibles a incluir en el estudio se les informó y se les pidió que leyesen y firmasen un consentimiento informado en el que aparecía toda la información referente a la investigación y su participación. Una vez firmado, se les citaba un día concreto en su centro de rehabilitación de referencia para participar en este estudio mediante la valoración con dos escalas de evaluación.

Para evaluar la validez de criterio de FMA, se evaluó a los mismos participantes, y el mismo día, con la Escala Nottingham de Vestido en Ictus-Revisada (ENVI-R), una escala de evaluación clínica del desempeño e independencia en la

actividad de la vida diaria del vestido. Se trata de una escala empleada en la práctica clínica y en la investigación con personas post-ictus, el cual es un instrumento válido y fiable en dicha población²¹⁻²⁹.

El investigador principal, terapeuta ocupacional clínico, que conoce y domina las escalas FMA y ENVI-R, administró la escala y evaluó la función sensoriomotora del miembro superior e inferior.

Se llevó a cabo una reevaluación por parte del mismo terapeuta de la evaluación inicial en el 59,1% de la muestra. Esta reevaluación fue realizada entre 1 y 4 meses después de la primera evaluación.

El examinador no comparó las puntuaciones durante la sesión de evaluación ni después, y las pruebas completadas permanecieron ciegas para el resto de los investigadores hasta que se completaron todas las evaluaciones.

Análisis estadístico

Las variables sociodemográficas y clínicas fueron expresadas en medias y porcentajes. La edad, fase (aguda/crónica), ubicación geográfica y sexo fueron especificados.

La prueba de Kaiser-Meyer-Olkin (KMO) midió la adecuación muestral y se utilizó la prueba de esfericidad de Bartlett para comprobar la redundancia entre variables; un KMO > 0,5 se consideró bueno, junto con valores significativos en la prueba de esfericidad de Bartlett («IBM Kaiser-Meyer-Olkin Measure for Identity Correlation Matrix» Estados Unidos, 2011). Se estimó un análisis factorial exploratorio (EFA) con extracción de máxima verosimilitud y rotación varimax para la estructura interna del nuevo cuestionario.

La consistencia interna del cuestionario se midió con el análisis estadístico α de Cronbach³⁰. Este dato fue arrojado en términos absolutos y por factores. El coeficiente de correlación intraclase (ICC) se empleó para evaluar la fiabilidad intra-evaluador. El error estándar de la medida fue usado para conocer el error de medición.

La validación del criterio convergente se refirió a la evaluación de instrumentos que evaluaban variables similares o que debían mostrar una relación similar. La relación de estas puntuaciones con las variables se utilizó para validar la escala. El coeficiente de correlación de Spearman se utilizó para evaluar la validez convergente entre la versión FMA total, FMA de Extremidades Superiores (FMA-UE), FMA de Extremidades Inferiores (FMA-LE) y la escala ENVI-R. Para la estimación de las mencionadas correlaciones se utilizó en primer lugar el test de Kolmogorov-Smirnov para conocer si la muestra era paramétrica o no paramétrica. Al ser una muestra significativa no paramétrica se empleó el test de Spearman para obtener coeficientes de correlación.

En todos los contrastes de hipótesis se consideró un nivel de significación estadística del 0,05.

El análisis estadístico fue realizado empleando el software SPSS versión 29 (IBM, USA).

Consideraciones éticas

Esta investigación contó con la aprobación del Comité de Ética de la Universidad de Málaga, con número de registro 7-2024-H, y respeta los preceptos éticos fundamentales según la declaración de Helsinki y el Informe Belmont, además de la normativa legal vigente en nuestro país en materia de investigación clínica, especialmente la Ley 14/2007, de 3 de julio, de Investigación Biomédica. Se garantizó en todo momento la protección y confidencialidad de la información, ya que el tratamiento, la comunicación y la cesión de los datos de carácter personal de todos los sujetos participantes se ajustará a lo dispuesto en la Ley Orgánica 3/2018, del 5 de diciembre, de protección de datos personales y garantía de los derechos digitales, y en su reglamento de desarrollo de datos. Los datos serán tratados de forma confidencial, asignándose un código correlativo y en ningún caso figurará el nombre y apellidos del paciente u otros datos que comprometan su intimidad. Los datos fueron tratados informáticamente y se incorporaron a un fichero automatizado de datos de carácter personal.

3.RESULTADOS

Participantes

Ciento diez participantes que cumplieron con los criterios de inclusión formaron parte de la muestra de este estudio. No hubo pérdidas de participantes en la recogida de datos. De los participantes del estudio, el 73.64% se encontraba en fase crónica (n=81), y el 26.36% en fase aguda (n=29). La edad media de los participantes fue de 62 ± 13.62 años. Los hombres representaron un 60% de la muestra total (n=66) y las mujeres el 40% (n=44).

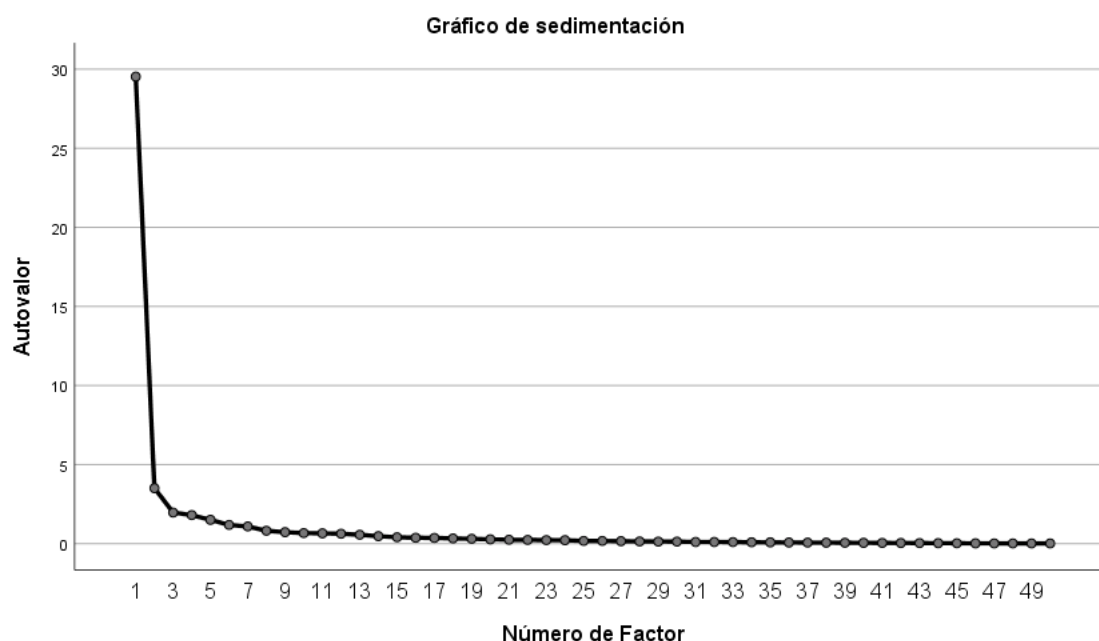
Validación clínica

Análisis factorial exploratorio

La escala en su versión traducida fue sometida a un análisis factorial exploratorio (EFA). La puntuación en el test de esfericidad de Bartlett's fue significativa ($X^2 = \infty$; $df = 7528.38$; $p < 0.001$); y para Chi-cuadrado también resultó significativa (Valor = 1567.54; $p < 0.001$). Por tanto, los datos fueron sometidos a un análisis factorial exploratorio con extracción de máxima verosimilitud y rotación Varimax. El valor propio > 1 y el guión (Figura 1) mostraron una solución de siete factores que explicó un 81% de la varianza total.

La carga de cada uno de los ítems en cada factor fue determinada por una carga superior a 0,3. (Tablas 1 y 2).

Figura 1: Guión



Fuente: elaboración propia.

Tabla 1: Características de los factores

Factor	Valor propio	Cargas SumSq	Solución Rotada	
			Proporción var.	% Acumulativo
1	29,529	18,255	36,511	36,511
2	3,507	8,340	16,681	53,191
3	1,966	4,543	9,085	62,277
4	1,803	3,153	6,306	68,583
5	1,506	1,434	2,868	71,450
6	1,183	1,029	2,058	73,509
7	1,086	,374	,748	74,257

Método de extracción: máxima verosimilitud.

Fuente: elaboración propia

Tabla 2: Cargas factoriales

	Factor						
	r 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
FM1	,066	,127	,034	,077	,140	-,095	,044
FM2	,002	,087	-,006	-,003	,373	,026	,017
FM3	,654	,346	,216	,517	,106	,007	,094
FM4	,593	,359	,222	,513	,151	,029	,062
FM5	,567	,261	,241	,542	,310	,093	,018
FM6	,625	,303	,341	,425	,197	,188	-,104
FM7	,631	,512	,139	,441	-,002	-,065	,023
FM8	,779	,210	,304	,303	,127	,015	-,104
FM9	,576	,508	,148	,486	-,031	-,111	-,104
FM10	,617	,488	,165	,493	,016	,006	-,144
FM11	,663	,437	,220	,425	,035	-,133	-,124
FM12	,747	,320	,263	,316	-,168	,225	,143
FM13	,704	,331	,270	,422	-,083	,184	,045
FM14	,850	,235	,241	,257	-,033	,094	,064
FM15	,725	,267	,342	,270	-,071	,264	,102
FM16	,647	,196	,328	,174	,118	,418	,092
FM17	,820	,217	,261	,107	,099	,169	,034
FM18	,465	,131	,263	,015	,205	,375	-,013
FM19	,893	,264	,244	,114	-,044	-,014	-,067
FM20	,910	,279	,168	,070	-,010	,031	-,100
FM21	,903	,225	,172	,115	-,065	,072	-,112
FM22	,898	,198	,177	,090	,030	,133	-,206

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

FM23	,715	,145	,193	,117	,250	,298	-,067
FM24	,725	,397	,133	,276	-,098	-,130	,043
FM25	,853	,300	,187	,191	-,097	-,130	,017
FM26	,732	,231	,322	,054	,090	,114	-,044
FM27	,779	,402	,208	,073	-,083	-,105	,068
FM28	,845	,231	,192	,130	,016	-,030	,148
FM29	,839	,345	,131	,075	,093	-,218	,154
FM30	,859	,352	,194	,063	,013	-,198	,166
FM31	,519	,275	,117	,267	,247	,099	,026
FM32	,458	,218	,203	,348	,249	,072	,041
FM33	,597	,092	,214	,224	,170	,241	,026
FM34	-,112	,072	,082	,032	,419	,031	-,053
FM35	,008	,368	,023	,007	,300	-,083	,021
FM36	,255	,783	,115	,125	,228	,048	,049
FM37	,350	,774	,181	,170	,173	,033	-,060
FM38	,478	,599	,457	,034	,069	-,100	-,102
FM39	,216	,622	,182	,172	,425	,162	,107
FM40	,267	,753	,263	,093	,217	,078	,093
FM41	,221	,719	,090	,267	,212	,054	,041
FM42	,287	,691	,329	,183	,001	-,039	-,158
FM43	,429	,633	,333	,086	-,057	,002	-,040
FM44	,517	,506	,509	,063	-,070	-,053	-,086
FM45	,423	,371	,687	,139	,227	,035	,017
FM46	,415	,320	,787	,118	,060	,013	-,031
FM47	,299	,209	,752	,162	,050	,079	,074
FM48	,295	,621	,241	,041	,028	,146	,008
FM49	,398	,480	,404	,157	,079	,093	,034
FM50	,186	,335	,440	,086	,137	,059	-,034

Fuente: elaboración propia

Consistencia interna

El cuestionario en su versión española presentó buenos valores de consistencia interna totales ($\alpha = 0.982$). Para analizar en mayor profundidad esta propiedad psicométrica, se calculó la consistencia interna si se elimina un ítem, hallando buenos valores de consistencia interna, siendo el máximo para los ítems 1, 2, 34 y 35 ($\alpha = 0.983$) y todos los demás ítems presentaron un valor muy cercano ($\alpha = 0.982$) (Table 3).

Tabla 3: Alfa de Cronbach si un ítem es eliminado

	Alfa de Cronbach si se elimina un ítem
FMA 1	0.983
FMA 2	0.983
FMA 3	0.982
FMA 4	0.982
FMA 5	0.982
FMA 6	0.982
FMA 7	0.982
FMA 8	0.982
FMA 9	0.982
FMA 10	0.982
FMA 11	0.982
FMA 12	0.982
FMA 13	0.982
FMA 14	0.982
FMA 15	0.982
FMA 16	0.982
FMA 17	0.982
FMA 18	0.982
FMA 19	0.982
FMA 20	0.982
FMA 21	0.982
FMA 22	0.982
FMA 23	0.982
FMA 24	0.982
FMA 25	0.982
FMA 26	0.982
FMA 27	0.982
FMA 28	0.982
FMA 29	0.982
FMA 30	0.982
FMA 31	0.982
FMA 32	0.982
FMA 33	0.982
FMA 34	0.983

FMA 35	0.983
FMA 36	0.982
FMA 37	0.982
FMA 38	0.982
FMA 39	0.982
FMA 40	0.982
FMA 41	0.982
FMA 42	0.982
FMA 43	0.982
FMA 44	0.982
FMA 45	0.982
FMA 46	0.982
FMA 47	0.982
FMA 48	0.982
FMA 49	0.982
FMA 50	0.982

Fuente: elaboración propia

Fiabilidad test-retest

La versión española de FMA mostró una fiabilidad satisfactoria al aplicarse por segunda vez a 65 participantes (59.1% de la muestra) entre 1 mes y 4 meses después de la evaluación inicial. La fiabilidad test-retest fue medida a través de un índice de correlación intraclass (ICC), y los datos analizados mostraron excelentes propiedades test-retest (ICC 0.98, 95%).

Error de medición

El error estándar de medición mostró una puntuación conveniente, con una puntuación de error estándar de la media de 3.17.

Validez de criterio

La versión española de FMA mostró muy buen ajuste de correlación convergente al ser comparada su puntuación total con la puntuación de la subescala de Miembro Superior (FMA-UE); y con la subescala de Miembro Inferior (FMA-LE); mostrándose correlaciones significativas, siendo la correlación entre FMA-Total y FMA-UE fue de $r=0.97$; y entre FMA-Total y FMA-LE de $r=0.89$.

Hypothesis Testing

La versión española de FMA presentó buena correlación con la escala ENVI-R. Se confirmó la correlación significativa entre FMA y ENVI-R, con una puntuación de $r=0.50$.

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

También se obtuvo una correlación significativa entre FMA-UE y ENVI-R ($r=0.34$), y entre FMA-LE y ENVI-R ($r=0.39$). Por lo que se pudo confirmar la hipótesis planteada.

4.DISCUSIÓN

La versión española de la escala Fugl-Meyer Assessment está disponible desde 2016, cuando se tradujo al español de España³¹, sin embargo, sus propiedades psicométricas no han sido evaluadas. La versión española colombiana está disponible desde 2019¹⁹. Tras 9 y 6 años de aquello, aún no se ha validado en población española que haya sufrido un ictus, a pesar de ser una de las escalas más recomendadas para su uso en la evaluación de la función sensoriomotora del miembro superior en esta población¹⁸.

En España se dispone de otras escalas de evaluación de este constructo, como Box and Block Test³² y ARAT³³. Sin embargo, son necesarias más escalas de evaluación de este constructo en español, para poder contrastar los resultados, porque no todos los clínicos en España pueden emplear estas dos herramientas. Además, disponer de una mayor variedad de escalas de evaluación de este constructo, enriquece la valoración de estas personas supervivientes del ictus, y permite mayor rigor en el tratamiento o en la investigación.

La validación clínica de la versión española de FMA fue sometida a un análisis estadístico de las propiedades psicométricas (análisis factorial exploratorio, consistencia interna, fiabilidad, error de medición, validez de criterio, prueba de hipótesis y sensibilidad), mostrando buena concordancia para la mayoría de los ítems. Dichas propiedades estadísticas fueron adecuadas y significativas³⁴⁻³⁶. Ofreciendo una versión altamente válida y recomendada para la evaluación de la función sensoriomotora en personas que han sufrido un ictus.

Anteriormente se han evaluado todas las propiedades psicométricas de FMA, sin embargo, de la versión española solo se ha analizado la content validity durante la realización de la traducción y adaptación transcultural¹⁹. Además, los estudios previos fueron realizados con menor muestra ($n=10$ ¹⁹ y $n=103$ ³¹). Por lo que este estudio supone, a conocimiento de los autores, el mayor estudio de la versión española de FMA, así como el primer análisis factorial de dicha versión. El EFA arrojó un modelo de 7 factores, rotado con Varimax, que explicó el 81% de la varianza total³⁷. Ha sido la primera vez que se realiza un análisis exploratorio de la versión completa de 50 ítems de FMA, ya que existen análisis factoriales exploratorios de otras versiones de FMA, las cuales han mostrado diferentes factores en la validez estructural de sus versiones, como la FMA-UE-Rom³⁸, o la FMA-UE-w/h³⁹, presentando un constructo unidimensional; por otro lado, para FMA-UE se han encontrado soluciones de 1¹⁷ y 4⁴⁰ factores; aunque la versión en la que más factores se han encontrado ha sido FMA-UE-w/h, con 12 factores⁴¹. La versión más similar a la realizada en este estudio es la traducción de la versión original de FMA al urdu, la cual presentó 2 factores⁴². No se ha encontrado ninguna otra versión de FMA que tenga una estructura como la versión española de FMA.

En cuanto a la consistencia interna de la versión española de FMA se encontraron excelentes niveles de consistencia interna con una puntuación alta de Alfa de Cronbach ($\alpha = 0.982$)³⁴ para la escala en general. Estos valores resultan superiores a los de otras versiones en otros idiomas, como la versión japonesa ($\alpha = 0.973$)^{43,44}, o la versión urdu ($\alpha = 0.80-0.87$)⁴⁰, mostrando

concordancia con versiones transculturalmente más similares como la versión polaca⁴⁵. Al compararlo con otras escalas de evaluación del mismo constructo en población española post-ictus, como es la escala ARAT, se ha hallado una mayor puntuación de consistencia interna para FMA que para ARAT, la cual presenta una puntuación de alfa de Cronbach de 0.96³².

La fiabilidad test-retest del cuestionario en español fue evaluada a través del ICC, el cual mostró propiedades de fiabilidad excelentes^{46,47}, con puntuación de 0.98 (ICC 95%). Esta puntuación resulta muy similar a la obtenida en otros estudios⁴⁸, siendo ligeramente menor a la obtenida en la versión japonesa de FMA (ICC=0.99-1)⁴³, teniendo en cuenta, que esa puntuación fue para inter-evaluador simultáneo, ya que cuando se evaluó a dicha población con inter-evaluador no simultáneo, la puntuación de ICC fue de 0.81 a 0.98⁴³. Otros estudios de la versión original de FMA también presentó un ICC superior tanto para inter-rater como para intra-rater ((ICC=0.99)⁴⁹. La puntuación obtenida en esta versión española de FMA sigue resultando elevada y significativa, siendo ligeramente superior a la obtenida en estudios previos de la versión original de FMA (ICC=0.96⁵⁰; ICC=0.95-0.98⁵¹), y superior a otras traducciones como la versión coreana de FMA (ICC>0.72⁵²). No obstante, la versión española de FMA ha presentado propiedades de fiabilidad excelentes. Esto concuerda con lo mostrado previamente, sugiriendo que las diferentes traducciones y validaciones de FMA en distintos idiomas son versiones válidas y fiables para evaluar la función sensoriomotora de miembros superiores e inferiores en personas que han sufrido un ictus.

Los autores utilizaron la puntuación total de la versión española de FMA y de ENVI-R para medir la validez de criterio, mostrando correlaciones positivas (Spearman rho: FMA vs ENVI-R = 0.50; FMA-UE vs ENVI-R = 0.34; FMA-LE vs ENVI-R = 0.39), de manera que las variables varían, sin implicar causalidad, y que la distribución de la variable es anormal^{53,54}. Este es el primer estudio en analizar la correlación entre la versión española de FMA con cualquier otra escala, en este caso la escala ENVI-R. Mediante esta correlación se pudo confirmar la hipótesis planteada en este estudio y lograr una adecuada sensibilidad para la versión española de FMA.

Observando los diferentes hallazgos y la comparación con otras escalas, la versión española de FMA presenta valores similares e incluso superiores a otras versiones en otros idiomas o a otras escalas de evaluación del mismo constructo para la misma población, para algunas propiedades psicométricas como consistencia interna^{42,43} o fiabilidad⁴⁸⁻⁵². Con estos hallazgos, se puede afirmar que la versión española de FMA es una escala muy adecuada para evaluar la función sensoriomotora en población española que ha sufrido un ictus, presentando buenas propiedades psicométricas tras el proceso de validación, por lo que se recomienda su uso en población española que haya sufrido un ictus y que presente dificultades sensoriomotoras en miembros superiores y/o inferiores.

Limitaciones y fortalezas del estudio

Una limitación de este estudio fue que el tamaño de la muestra de la validación no fue muy amplio y que dicha muestra fue obtenida a través de un muestreo por conveniencia en la que el paciente aceptaba participar en su centro de rehabilitación tras la invitación por parte del investigador para implicarse en el estudio. Además, no se realizaron las evaluaciones por 2 evaluadores independientes, impidiendo evaluar la fiabilidad inter-evaluador⁵⁵. También sería recomendable poder generalizar y validar la versión española de FMA a personas con otras patologías que causen limitaciones en la función sensoriomotora de miembros superiores e inferiores. Otra de las limitaciones, fue que se podría haber evaluado a los participantes con más escalas, que permitiesen analizar la correlación de la escala FMA con otros cuestionarios. Como principal fortaleza se destaca que este artículo es la primera validación de la versión española de FMA en población española. Además, ningún estudio previo analizó las propiedades psicométricas estudiadas en esta investigación para la versión española de FMA.

5.CONCLUSIÓN

El proceso de validación de la versión española de FMA ha demostrado que esta escala es una herramienta que podría ser empleada en población española para la evaluación de la función sensoriomotora de miembros superiores e inferiores afectados en personas que han sufrido un ictus. También se verificó que los valores obtenidos para las propiedades psicométricas fueron adecuados y significativos en el uso de la versión española de FMA en una muestra de personas españolas con ictus. Basándose en estos resultados, se puede defender y recomendar el uso de la versión española de FMA en la investigación y en la práctica clínica. El empleo de esta versión española de FMA permitirá identificar limitaciones en la función sensoriomotora en los miembros superiores e inferiores de personas españolas que han sufrido un ictus y, por lo tanto, permitirá comparaciones fiables de resultados del ictus entre regiones y países donde existen otras versiones transculturalmente adaptadas y validadas, tanto en la práctica clínica, como en la investigación.

6. AGRADECIMIENTOS

Los autores agradecen a todas las personas que participaron en este estudio y a los centros colaboradores que permitieron la realización de las evaluaciones en sus instalaciones, como Integra Daño Cerebral, Dacemur, CMVCaridad Puerto Lumbreras, Asociación DISMO, CIAN Grupo 5, Neuroavanzo, Caser Salud, Neuron y Fundación AISSE. También agradecen a la Universidad de Nottingham, así como a los autores originales y a todas las personas involucradas en esta investigación.

7.DECLARACIÓN DE CONFLICTO DE INTERESES

Los autores no informaron de ningún conflicto de intereses potencial.

8.CONTRIBUCIÓN DE LOS AUTORES

Todos los autores han acordado la versión final y cumplen al menos uno de los siguientes criterios (recomendados por el ICMJE): contribuciones sustanciales a la concepción y el diseño, adquisición de datos o análisis e interpretación de datos; redacción del artículo; o revisión crítica del mismo para determinar su contenido intelectual importante.

9.FINANCIACIÓN

Esta investigación no recibió ninguna subvención específica de agencias de financiación de los sectores público, comercial o sin fines de lucro.

10.REFERENCIAS

1.
Feigin VL, Owolabi MO, World Stroke Organization–Lancet Neurology Commission Stroke Collaboration Group. Pragmatic solutions to reduce the global burden of stroke: a World Stroke Organization–Lancet Neurology Commission. *Lancet Neurol.* 2023 Dec;22(12):1160-206.
2.
Hua X, Liu M, Wu S. Definition, prediction, prevention and management of patients with severe ischemic stroke and large infarction. *Chin Med J (Engl).* 2023 Dec;136(24):2912-22.
3.
Katan M, Luft A. Global Burden of Stroke. *Semin Neurol.* 2018 Apr;38(2):208-11.
4.
Masjuan J, Alvarez-Sabín J, Arenillas J, Calleja S, Castillo J, Dávalos A, et al. [Stroke health care plan (ICTUS II. 2010)]. *Neurologia.* 2011 Sep;26(7):383-96.
5.
Díaz-Guzmán J, Egido JA, Gabriel-Sánchez R, Barberá-Comes G, Fuentes-Gimeno B, Fernández-Pérez C, et al. Stroke and transient ischemic attack incidence rate in Spain: the IBERICTUS study. *Cerebrovasc Dis.* 2012;34(4):272-81.
6.
Ingram LA, Butler AA, Lord SR, Gandevia SC. Use a physiological profile to document upper limb motor impairment in ageing and neurological conditions. *The Journal of Physiology* 2023;601:2251–62. <https://doi.org/10.1113/JP283703>.
7.
Li S, Ghuman J, Gonzalez-Buonomo J, Huang X, Malik A, Yozbatiran N, et al. Does Spasticity Correlate With Motor Impairment in the Upper and Lower Limbs in Ambulatory Chronic Stroke Survivors? *Am J Phys Med Rehabil* 2023;102:907–12. <https://doi.org/10.1097/PHM.0000000000002247>.
8.
Zhao K, He C, Xiang W, Zhou Y, Zhang Z, Li J, et al. Evidence of synergy coordination patterns of upper-limb motor control in stroke patients with mild and

moderate impairment. Front Physiol 2023;14:1214995.
<https://doi.org/10.3389/fphys.2023.1214995>.

9.

Dawson J, Engineer ND, Cramer SC, Wolf SL, Ali R, O'Dell MW, et al. Vagus Nerve Stimulation Paired With Rehabilitation for Upper Limb Motor Impairment and Function After Chronic Ischemic Stroke: Subgroup Analysis of the Randomized, Blinded, Pivotal, VNS-REHAB Device Trial. *Neurorehabil Neural Repair* 2023;37:367–73. <https://doi.org/10.1177/15459683221129274>.

10.

Kwakkel G, Stinear C, Essers B, Munoz-Novoa M, Branscheidt M, Cabanas-Valdés R, et al. Motor rehabilitation after stroke: European Stroke Organisation (ESO) consensus-based definition and guiding framework. *European Stroke Journal* 2023;8:880–94. <https://doi.org/10.1177/23969873231191304>.

11.

Bushnell C, Bettger JP, Cockcroft KM, Cramer SC, Edelen MO, Hanley D, et al. Chronic Stroke Outcome Measures for Motor Function Intervention Trials: Expert Panel Recommendations. *Circ: Cardiovascular Quality and Outcomes* 2015;8. <https://doi.org/10.1161/CIRCOUTCOMES.115.002098>.

12.

Gunduz ME, Bucak B, Keser Z. Advances in Stroke Neurorehabilitation. *JCM* 2023;12:6734. <https://doi.org/10.3390/jcm12216734>.

13.

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;27:1147–57. <https://doi.org/10.1007/s11136-018-1798-3>.

14.

Santisteban L, Térémetz M, Bleton J-P, Baron J-C, Maier MA, Lindberg PG. Upper Limb Outcome Measures Used in Stroke Rehabilitation Studies: A Systematic Literature Review. *PLoS ONE* 2016;11:e0154792. <https://doi.org/10.1371/journal.pone.0154792>.

15.

Yu B, Kaku A, Liu K, Parnandi A, Fokas E, Venkatesan A, et al. Quantifying Impairment and Disease Severity Using AI Models Trained on Healthy Subjects. *ArXiv* 2023:arXiv:2311.12781v1.

16.

Lin G-H, Huang C-Y, Lee S-C, Chen K-L, Lien J-JJ, Chen M-H, et al. A 10-item Fugl-Meyer Motor Scale Based on Machine Learning. *Physical Therapy* 2021;101:pzab036. <https://doi.org/10.1093/ptj/pzab036>.

17.

Tauchi Y, Kyougoku M, Takahashi K, Okita Y, Takebayashi T. Dimensionality and item-difficulty hierarchy of the Fugl-Meyer assessment of the upper extremity among Japanese patients who have experienced stroke. *Topics in Stroke Rehabilitation* 2022;29:579–87. <https://doi.org/10.1080/10749357.2021.1965797>.

18.

de Blas Zamorano P, Montagut-Martínez P, Pérez-Cruzado D, Merchan-Baeza JA. Fugl-Meyer Assessment for upper extremity in stroke: A psychometric systematic review. *Journal of Hand Therapy* [Internet]. 2025; Disponible en: <https://www.sciencedirect.com/science/article/pii/S0894113025000572>

19.

Barbosa NE, Forero SM, Galeano CP, Hernández ED, Landinez NS, Sunnerhagen KS, et al. Translation and cultural validation of clinical observational scales – the Fugl-Meyer assessment for post-stroke sensorimotor function in Colombian Spanish. *Disability and Rehabilitation* 2019;41:2317–23. <https://doi.org/10.1080/09638288.2018.1464604>.

20.

Gjersing L, Caplehorn JRM, Clausen T. Cross-cultural adaptation of research instruments: language, setting, time and statistical considerations. *BMC Med Res Methodol*. 2010 Feb;10:13.

21.

Sanford J, Moreland J, Swanson LR, Stratford PW, Gowland C. Reliability of the Fugl-Meyer Assessment for Testing Motor Performance in Patients Following Stroke. *Physical Therapy* 1993;73:447–54. <https://doi.org/10.1093/ptj/73.7.447>.

22.

Walker MF, Sunderland A, Fletcher-Smith J, Drummond A, Logan P, Edmans JA, et al. The DRESS trial: a feasibility randomized controlled trial of a neuropsychological approach to dressing therapy for stroke inpatients. *Clin Rehabil*. 2012 Aug;26(8):675-85.

23.

Walker CM, Sunderland A, Sharma J, Walker MF. The impact of cognitive impairment on upper body dressing difficulties after stroke: a video analysis of patterns of recovery. *J Neurol Neurosurg Psychiatry*. 2004 Jan;75(1):43-8.

24.

Walker M. Dressing ability after stroke. *Nurs Times*. 1992 Jan;88(1):51.

25.

Walker MF, Lincoln NB. Factors influencing dressing performance after stroke. *J Neurol Neurosurg Psychiatry*. 1991 Aug;54(8):699-701.

26.

Walker MF, Lincoln NB. Reacquisition of dressing skills after stroke. *Int Disabil Stud*. 1990;12(1):41-3.

27.

Walker CM, Walker MF, Sunderland A. Dressing after a Stroke: A Survey of Current Occupational Therapy Practice. *British Journal of Occupational Therapy*. 2003 Jun;66(6):263-8.

28.

Fletcher-Smith JC. Recovery of dressing ability after stroke [Internet]. [Nottingham, UK]: University of Nottingham; 2011. Available from: <https://eprints.nottingham.ac.uk/11913/>

29.

Fletcher-Smith J, Walker M, Sunderland A, Garvey K, Wan A, Turner H. An Interrater Reliability Study of the Nottingham Stroke Dressing Assessment. *British Journal of Occupational Therapy*. 2010 Dec;73(12):570-8.

30.

Cronbach L. Coefficient alpha and the internal structure of tests. *Psyko Psykometrika*. 1951 Sep;16(3):2997-334.

31.

Ferrer González BM. Adaptación y validación al español de la escala Fugl-Meyer en el manejo de la rehabilitación de pacientes con ictus [Internet]. [España]: Universidad de Sevilla; 2016. Available from: <https://idus.us.es/items/d3473ef7-451d-4f90-b708-2d49838ab673>

32.

Noce MP, Valera-Gran D, Hurtado-Pomares M, Serrano-Reina E, Soler-Pons C, Navarrete-Muñoz EM. Spanish translation and cross-cultural adaptation of the Box and Block Test: a pilot study in adults with chronic acquired brain injury. *Brain Impair.* 2024 Sep;25:1B24014.

33.

Fernández-Solana J, Pardo-Hernández R, González-Bernal JJ, Sánchez-González E, González-Santos J, Soto-Cámara R, et al. Psychometric Properties of the Action Research Arm Test (ARAT) Scale in Post-Stroke Patients-Spanish Population. *Int J Environ Res Public Health.* 2022 Nov 13;19(22):14918.

34.

Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med.* 2016 Jun;15(2):155-63.

35.

De Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine: A Practical Guide* [Internet]. 1.^a ed. Cambridge University Press; 2011 [citado 26 de octubre de 2024]. Available from: <https://www.cambridge.org/core/product/identifier/9780511996214/type/book>

36.

Moral De La Rubia J. Revisión de los criterios para validez convergente estimada a través de la Varianza Media Extraída. *Psychol.* 2019 Nov;13(2):25-41.

37.

Méndez Martínez C, Rodón epúlveda MA. Introducción al análisis factorial exploratorio. *Revista Colombiana de Psiquiatría.* 2012;41(1):197-207.

38.

Roman N, Miclaus R, Repanovici A, Nicolau C. Equal Opportunities for Stroke Survivors' Rehabilitation: A Study on the Validity of the Upper Extremity Fugl-Meyer Assessment Scale Translated and Adapted into Romanian. *Medicina* 2020;56:409. <https://doi.org/10.3390/medicina56080409>.

39.

Chen K-L, Chen C-T, Chou Y-T, Shih C-L, Koh C-L, Hsieh C-L. Is the Long Form of the Fugl-Meyer Motor Scale More Responsive Than the Short Form in Patients With Stroke? *Archives of Physical Medicine and Rehabilitation* 2014;95:941–9. <https://doi.org/10.1016/j.apmr.2014.01.014>.

40.

Woodbury ML, Velozo CA, Richards LG, Duncan PW, Studenski S, Lai S-M. Dimensionality and Construct Validity of the Fugl-Meyer Assessment of the Upper Extremity. *Archives of Physical Medicine and Rehabilitation* 2007;88:715–23. <https://doi.org/10.1016/j.apmr.2007.02.036>.

41.

Persch AC, Gugiu PC, Velozo CA, Page SJ. Rasch Analysis of the Wrist and Hand Fugl-Meyer: Dimensionality and Item-Level Characteristics. *Journal of Neurologic Physical Therapy* 2015;39:185–92. <https://doi.org/10.1097/NPT.0000000000000096>.

42.

Ikram M, Rehman SSU, Sunnerhagen KS, Alt Murphy M. Urdu translation and cross-cultural validation of the Fugl-Meyer assessment in people with stroke. *Disability and Rehabilitation* 2022;44:8048–53. <https://doi.org/10.1080/09638288.2021.2003449>.

43.

Amano S, Umeji A, Uchita A, Hashimoto Y, Takebayashi T, Takahashi K, et al. Clinimetric properties of the Fugl-Meyer assessment with adapted guidelines for the assessment of arm function in hemiparetic patients after stroke. *Topics in Stroke Rehabilitation* 2018;25:500–8. <https://doi.org/10.1080/10749357.2018.1484987>.

44.

Tauchi Y, Kyougoku M, Okita Y, Takebayashi T. Structural validity and internal consistency of a hypothesized factor structure of the Fugl-Meyer Assessment of the upper extremity. *Top Stroke Rehabil.* 2023 Jul;30(5):501–11.

45.

Goliwās M, Małecka J, Adamczewska K, Flis-Masłowska M, Lewandowski J, Kocur P. Polish Cultural Adaptation and Reliability of the Fugl-Meyer Assessment of Motor Performance and Sensory Assessment Scale in Stroke Patients. *J Clin Med.* 2024 Jun 26;13(13):3710.

46.

Martínez Pérez JA, Pérez Martin PS. [Intraclass correlation coefficient]. *Semerger.* 2023 Apr;49(3):101907.

47.

Liljequist D, Elfving B, Skavberg Roaldsen K. Intraclass correlation - A discussion and demonstration of basic features. *PLoS One.* 2019;14(7):e0219854.

48.

Lin G-H, Huang C-Y, Lee S-C, Chen K-L, Lien J-JJ, Chen M-H, et al. A 10-item Fugl-Meyer Motor Scale Based on Machine Learning. *Physical Therapy* 2021;101:pzab036. <https://doi.org/10.1093/ptj/pzab036>.

49.

See J, Dodakian L, Chou C, Chan V, McKenzie A, Reinkensmeyer DJ, et al. A Standardized Approach to the Fugl-Meyer Assessment and Its Implications for Clinical Trials. *Neurorehabil Neural Repair* 2013;27:732–41. <https://doi.org/10.1177/1545968313491000>.

50.

Sanford J, Moreland J, Swanson LR, Stratford PW, Gowland C. Reliability of the Fugl-Meyer Assessment for Testing Motor Performance in Patients Following Stroke. *Physical Therapy* 1993;73:447–54. <https://doi.org/10.1093/ptj/73.7.447>.

51.

Hsueh I-P, Hsu M-J, Sheu C-F, Lee S, Hsieh C-L, Lin J-H. Psychometric Comparisons of 2 Versions of the Fugl-Meyer Motor Scale and 2 Versions of the Stroke Rehabilitation Assessment of Movement. *Neurorehabil Neural Repair* 2008;22:737–44. <https://doi.org/10.1177/1545968308315999>.

52.

Kim T, Hwang SH, Lee WJ, Hwang JW, Cho I, Kim E-H, et al. The Korean Version of the Fugl-Meyer Assessment: Reliability and Validity Evaluation. *Ann Rehabil Med* 2021;45:83–98. <https://doi.org/10.5535/arm.20225>.

53.

Akoglu H. User's guide to correlation coefficients. *Turk J Emerg Med*. 2018 Sept;18(3):91-3.

54.

Schober P, Boer C, Schwarte LA. Correlation Coefficients: Appropriate Use and Interpretation. *Anesth Analg*. 2018 May;126(5):1763-8.

55.

Luján Tangarife J, Crdona Arias J. Construcción y validación de escalas de medición en salud: revisión de propiedades psicométricas. *Archivos de Medicina*. 2015;11(3.1).

7. DISCUSIÓN

La presente Tesis Doctoral ha incorporado el protocolo de investigación COSMIN, junto con los criterios de Terwee y Prinsen, y el método GRADE para elaborar dos revisiones sistemáticas psicométricas. La primera de ellas dirigida a conocer la mejor escala para evaluar la independencia y la funcionalidad de supervivientes de ictus, entre todas las versiones de BI y de FIM (Artículo I). La segunda de ellas sobre la mejor versión de FMA para la evaluación de la función sensoriomotora del miembro superior en personas que habían sufrido un ictus (Artículo II) (de Blas-Zamorano et al., 2025).

La investigación recopiló de forma exhaustiva la información disponible hasta la fecha de las diferentes versiones de BI, FIM y FMA.

Posteriormente se revisó y actualizó la escala NSDA a NSDA-R. A continuación, se siguió la guía ISPOR para la traducción y adaptación transcultural de la escala del vestido NSDA-R, a la versión española ENVI-R; la cual fue seguidamente validada en población española que había sufrido un ictus, analizando sus propiedades psicométricas siguiendo la normativa COSMIN (Artículo III).

Por último, se validó la versión española de la escala FMA, disponible en la Universidad de Göteborgs; en población española que había sufrido un ictus; empleando nuevamente, la normativa COSMIN (Artículo IV).

Ambas validaciones presentaron buenas propiedades psicométricas, permitiendo recomendar su uso para la evaluación de personas españolas que hayan sufrido un ictus.

7.1 Análisis psicométrico y revisión de los cuestionarios de evaluación

A través del Artículo I y del Artículo II incluidos en esta tesis empleamos la metodología COSMIN para revisiones sistemáticas, enumerando y evaluando las propiedades psicométricas de las diferentes versiones del Índice de Barthel y de Functional Independence Measure para medir la independencia y funcionalidad en las AVDs en personas que han sufrido un ictus, en el caso del Artículo I; y de las diferentes versiones de Fugl-Meyer Assessment para la evaluación de la función sensoriomotora del miembro superior en supervivientes del ictus, en el caso del Artículo II.

Este estudio empleó la metodología COSMIN para llevar a cabo una revisión sistemática de cuestionarios existentes. Este enfoque se centra en instrumentos con información disponible sobre sus propiedades de medición, idoneidad evaluadora y aplicabilidad para propósitos específicos. La metodología se dividió en varias fases: búsqueda y selección sistemática de literatura, evaluación de la calidad de los estudios elegibles, valoración de las propiedades de medición de los instrumentos, análisis de interpretabilidad y factibilidad, formulación de conclusiones y recomendaciones, y difusión de los resultados (U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research et al., 2006).

Los primeros pasos se adhirieron a las directrices estándar para revisiones sistemáticas. Sin embargo, los pasos subsiguientes implicaron un proceso más complejo de evaluación de las propiedades de medición, culminando con la selección del cuestionario apropiado. Esta evaluación se realizó en tres etapas:

1. La calidad metodológica de cada estudio se evaluó utilizando la lista de verificación de riesgo de sesgo de COSMIN (COSMIN, 2011; US Department of Health, 2006).
2. La calidad de los resultados de cada estudio se calificó según los criterios actualizados para las propiedades de medición de Terwee y Prinsen (COSMIN, 2011; Mokkink et al., 2010b; Prinsen et al., 2018).
3. La evidencia se sintetizó y su calidad se calificó mediante el método GRADE modificado (COSMIN, 2011; Mokkink et al., 2010b; Prinsen et al., 2018).

7.1.1 Normativa COSMIN

La investigación en ciencias de la salud depende de la precisión de sus herramientas de evaluación. La proliferación de Instrumentos de Medición de Resultados (OMI) exige un análisis y cribado riguroso de su calidad metodológica, especialmente en el contexto de las revisiones sistemáticas. Diferentes autores han propuesto metodologías variadas para esta evaluación, considerando desde el tipo de pruebas realizadas (Haywood et al., 2006) hasta la idoneidad de los análisis y el tamaño muestral (Marinus et al., 2002). Sin embargo, la complejidad inherente a la selección de OMI para propósitos específicos se extiende a aspectos conceptuales (definición del constructo y la población de estudio), prácticos (carga para personas y costos), y, fundamentalmente, a la evaluación de sus propiedades de medición: fiabilidad, validez y sensibilidad al cambio (Ioannidis et al., 2014).

La toma de decisiones clínicas y diagnósticas se basa en las puntuaciones de los OMI, lo que subraya la necesidad crítica de que sean fiables y válidos. La ausencia de estas cualidades no solo compromete la precisión de los resultados, induciendo sesgos y el consiguiente desperdicio de recursos, sino que también plantea un dilema ético para las personas que participan en los estudios (European Medicines Agency, 2005). Marshall et al. (2000) demostraron que el uso de escalas no validadas en ensayos clínicos sobre esquizofrenia se correlacionaba con una mayor probabilidad de reportar resultados estadísticamente significativos, sugiriendo una posible sobrestimación de la efectividad de los tratamientos (Marshall et al., 2000).

Antes de su aplicación en investigación o práctica clínica, todo instrumento de salud debe ser evaluado en sus propiedades de medición. Estas incluyen la validez de contenido, validez estructural, consistencia interna, fiabilidad, error de medición, prueba de hipótesis, validez transcultural/invariancia de medición, validez de criterio y sensibilidad al cambio (Ioannidis et al., 2014). Al menos, las agrupadas en los dominios fundamentales de validez, fiabilidad y sensibilidad al cambio deben ser consideradas. Es imperativo que los estudios que evalúen estas propiedades posean una alta calidad metodológica para asegurar conclusiones apropiadas (Prinsen et al., 2018).

La iniciativa COSMIN representa un esfuerzo internacional concertado para mejorar la selección de OMI. A través de un estudio Delphi internacional (Terwee et al., 2018), COSMIN ha desarrollado guías metódicas para evaluar la calidad metodológica de los estudios sobre propiedades de medición, aplicables tanto en la investigación como en la práctica clínica.

El manual COSMIN se fundamenta en guías de revisión sistemática existentes, como el manual Cochrane (Higgins, 2024), la Declaración PRISMA (Moher et al., 2010), la lista de verificación STARD (Bossuyt et al., 2013), los criterios actualizados de Terwee y Prinsen, y la metodología GRADE (COSMIN, 2011;

Mokkink et al., 2010b; Prinsen et al., 2018). Esta sólida base garantiza que el manual cumple con los requisitos para revisiones sistemáticas de alta calidad metodológica y una selección adecuada de instrumentos.

COSMIN se estructura en un protocolo para revisiones sistemáticas de OMI, basado en una guía de 10 puntos. Este protocolo detalla la síntesis de la evidencia, permitiendo combinar el número de estudios, su calidad y la consistencia de los resultados para determinar la solidez de la evidencia de cada propiedad de medición. Por ejemplo, se considera evidencia concluyente de fiabilidad adecuada con resultados positivos consistentes (coeficiente de correlación intraclase (ICC) o Kappa > 0.70) en al menos dos estudios de buena calidad o uno de excelente calidad, similar a la metodología GRADE (Montagut, 2022).

A pesar de una fiabilidad adecuada, un alfa de Cronbach elevado no asegura la medición completa del constructo de interés ni la ausencia de conceptos importantes (Cortina, 1993; Sijtsma, 2009). Un instrumento puede ser muy fiable, pero incompleto o incorrecto, lo que podría sobrestimar o subestimar un cambio real debido a la inclusión de conceptos irrelevantes o la omisión de conceptos cruciales. La frustración del usuario ante preguntas irrelevantes puede llevar a sesgos en las respuestas o bajas tasas de respuesta (De Leeuw et al., 2012).

Factores como la estabilidad del constructo, el intervalo de tiempo apropiado, la selección del método estadístico adecuado (ICC, Pearson, Spearman, Kappa) (Cohen, 1968; de Vet et al., 2013; Shrout et al., 1979) y la independencia en la administración de las pruebas, pueden diferenciar una fiabilidad "muy buena" de una "dudosa".

No obstante, la validez de contenido emerge como una propiedad psicométrica de mayor relevancia que la fiabilidad. COSMIN recomienda establecer la validez de contenido antes de evaluar otras propiedades de medición (Elsman et al., 2024). La validez de contenido asegura que todos los elementos del instrumento sean pertinentes para el constructo de interés en una población y contexto específicos, que el instrumento sea exhaustivo en cuanto a las preocupaciones del usuario y que sea comprendido como se pretende (Brod et al., 2009).

La validez de contenido de los OMI existentes puede evaluarse mediante la consulta a usuarios y profesionales sobre la relevancia, exhaustividad y comprensibilidad de los ítems (Elsman et al., 2024). Sin embargo, los métodos varían ampliamente, y muchos estudios solo abordan la comprensibilidad (Chiarotto et al., 2018). Por lo tanto, se necesitan nuevas directrices para evaluar la calidad metodológica de los estudios de validez de contenido, reflejada en la necesidad de actualizar la sección de estándares de COSMIN en esta área, incorporando criterios para evaluar la calidad del desarrollo del instrumento y la comprensibilidad, y detallando el "cómo" de estas acciones (Terwee et al., 2018).

Además, se requieren criterios claros para definir una buena validez de contenido que permitan recomendaciones transparentes y basadas en la evidencia para la selección de OMI en revisiones sistemáticas. Esto es vital para determinar la robustez de un instrumento para medir resultados, su aprobación regulatoria o su inclusión en registros de herramientas de evaluación (Terwee et al., 2018).

Existen diversas guías y estándares para evaluar la calidad de los OMI, como los de la American Psychological Association (APA), la Agency for Healthcare Research and Quality (AHRQ), SAC-MOT, EMPRO, entre otros (Montagut, 2022). La principal distinción entre estas guías y COSMIN radica en que la mayoría son más concisas, carecen de una descripción exhaustiva de los pasos para determinar la calidad de los OMI, y no establecen que las conclusiones deban basarse en una consideración sistemática de toda la evidencia existente. A menudo, combinan estándares de calidad del estudio con criterios para calificar los resultados de las propiedades de medición en un único estándar, lo que limita la transparencia y sistematicidad de la evaluación del riesgo de sesgo (Montagut, 2022).

A nivel individual, otras diferencias con COSMIN incluyen que estas guías no siempre abordan los estudios de propiedades psicométricas de los resultados informados por los usuarios en ensayos clínicos, ni cubren el informe de estudios que prueban la propiedad de medida de un instrumento (Calvert et al., 2013). No todas evalúan la totalidad de las propiedades psicométricas, ni describen cómo combinar los resultados de estudios de diferente calidad para una conclusión general sobre un instrumento (Chou et al., 2018; Mokkink et al., 2010b). Algunas no se centran específicamente en OMI, o no incluyen la valoración de la unidimensionalidad dentro del estándar de consistencia interna (Chou et al., 2018). Además, ciertas propiedades de medición, como la fiabilidad y el error de medición, no se mencionan explícitamente en algunas de ellas (Chou et al., 2018; Montagut, 2022).

7.1.2 Criterios Terwee y Prinsen

COSMIN enfatiza la necesidad de evaluar la calidad de los estudios que examinan las propiedades de medición de los cuestionarios. Si bien las revisiones sistemáticas han contrastado estas propiedades, se requieren criterios explícitos para definir una "buena" propiedad de medición y discernir la superioridad de una herramienta sobre otra. No obstante, la calidad psicométrica no es el único factor determinante en la selección de un instrumento; también influyen consideraciones pragmáticas y conceptuales.

El Comité Asesor Científico sobre las Propiedades de Medición (SAC-MOT) ha delineado ocho atributos esenciales para evaluar la idoneidad de los instrumentos de medición (Lohr, 2002; Mokkink et al., 2010a):

- Modelo Conceptual y de Medición: Coherencia del marco teórico subyacente.
- Validez: Precisión en la medición de lo que se pretende medir.
- Fiabilidad: Consistencia y estabilidad de las mediciones.
- Sensibilidad al Cambio: Habilidad para detectar cambios significativos en el estado de salud.
- Interpretabilidad: Facilidad de comprensión y aplicación de los resultados.
- Demandas y Carga Administrativa: Practicidad en términos de tiempo y recursos.
- Formas Alternativas: Disponibilidad de versiones que se adapten a distintas necesidades.
- Cultura y Adaptaciones del Idioma: Aplicabilidad transcultural y lingüística.

Aunque otros autores han propuesto criterios similares (Andresen, 2000; Bombardier et al., 1987; McDowell et al., 1996), una limitación recurrente es la ausencia de razonamientos explícitos sobre lo que constituye una "buena" propiedad de medición. A menudo, no se especifican el número de hipótesis a confirmar para una validez adecuada, los umbrales de acuerdo aceptables (error de medición), los criterios de buena capacidad de respuesta, la interpretabilidad óptima o el tamaño muestral apropiado.

En respuesta a estas deficiencias, Terwee ha analizado y perfeccionado los criterios de calidad para estudios de desarrollo y evaluación de cuestionarios sobre el estado de salud. Su trabajo se centra en la formulación de criterios explícitos para propiedades de medición clave, que incluyen validez de contenido, validez estructural, consistencia interna, fiabilidad, error de medición, prueba de hipótesis, validez transcultural/invarianza de medición, validez de criterio y sensibilidad al cambio (Montagut, 2022).

Terwee identifica dos preocupaciones fundamentales:

1. Cuestionarios de calidad deficiente pueden obtener valoraciones positivas debido a la laxitud de los criterios.
2. Los cuestionarios con más valoraciones positivas no son inherentemente superiores, ya que la calificación depende de la disponibilidad y calidad de la información reportada.

En sus aplicaciones iniciales, Terwee observó que muchas propiedades de medición no se evaluaban, analizaban o comunicaban adecuadamente, y la información para la validez de contenido era frecuentemente limitada. Una posible explicación es que los criterios se basan en consensos de expertos más que en evidencia empírica robusta, lo que puede llevar a decisiones subjetivas.

A diferencia de las revisiones sistemáticas de ensayos controlados aleatorizados (ECAs) (Jadad et al., 1996), estos criterios no se consolidan en una puntuación de calidad global, evitando así la suposición de que todas las propiedades de

medición tienen igual importancia. Para Terwee, la validez de contenido es la propiedad psicométrica más significativa, una perspectiva compartida por COSMIN, ya que solo si esta es adecuada, el uso del cuestionario es pertinente y la evaluación de sus demás propiedades resulta relevante (Mokkink et al., 2010; Mokkink et al., 2010b; Prinsen et al., 2018; Terwee et al., 2018).

Para desarrollar criterios robustos, es imprescindible considerar varios aspectos (Strauss, 2005):

- Un estudio de validación deficientemente documentado puede llevar a bajas calificaciones para cuestionarios que no son intrínsecamente defectuosos. Esto subraya la necesidad de mejorar la calidad de los estudios de desarrollo y evaluación.
- Es vital determinar si el cuestionario fue diseñado con fines discriminatorios o evaluativos, información que a menudo no se especifica. Además, los conceptos a medir deben estar bien definidos y la selección de ítems debe describirse con precisión para juzgar la aplicabilidad en una población específica.
- Es fundamental verificar que el número de escalas se corresponda con el número de dimensiones identificadas en el análisis factorial. Se han observado discrepancias donde cuestionarios afirmaban cubrir múltiples dimensiones, pero consistían en una única escala, o viceversa.
- Se debe verificar si las hipótesis para la evaluación de la validez de constructo han sido detalladas, lo cual es frecuente su omisión o su insuficiencia en especificidad. Una hipótesis debe evaluar directamente las afirmaciones de una teoría e incluir una justificación de la dirección y magnitud esperada de la correlación para mitigar el riesgo de sesgo. Es común que los autores concluyan la validez de su instrumento incluso cuando el recuento objetivo de las hipótesis confirmadas sugiere lo contrario.
- Finalmente, la sensibilidad al cambio a menudo se encuentra mal definida y evaluada. Algunos autores emplean una multitud de índices con diferencias conceptuales y algebraicas, y aún así concluyen que su instrumento es sensible, incluso si los índices no lo respaldan consistentemente (Beaton et al., 1997; Stratford et al., 2005; Stuchi et al., 1995; Wright et al., 1997; Zou, 2005).

Estos puntos destacan la complejidad y la importancia de establecer criterios claros y empíricamente fundamentados para la evaluación de la calidad de los instrumentos de medición en salud, con el objetivo de optimizar la investigación y la toma de decisiones clínicas.

7.1.3 Método GRADE

En la investigación formal, la evaluación de la calidad de la evidencia es crucial y complementa la valoración metodológica y la robustez de los resultados, especialmente en el contexto de la metodología COSMIN. Este proceso se inicia con una búsqueda sistemática y automatizada de literatura científica para maximizar la certeza de los hallazgos. Es fundamental considerar la validez externa, analizando cómo factores culturales, poblacionales, económicos, ambientales y tecnológicos pueden influir en la aplicabilidad de los resultados (Prinsen et al., 2018).

La interpretación y ponderación de la evidencia requiere la selección de un sistema de clasificación idóneo para discernir los niveles de certeza y formular recomendaciones contextualizadas. La comunidad científica ha desarrollado diversas taxonomías, desde las pioneras de Sackett (Sackett et al., 1996) hasta esquemas más recientes como CACG (Boulet et al., 1999; Ernst et al., 1996), AATM (Jovell et al., 1995), SAS (Servicio Andaluz de Salud, 2012), SORT (Ebell et al., 2004), NHMRC (Merlin et al., 2009), CEBM (Centre for Evidence Based Medicine, 2009; Manterola et al., 2014), NICE (NICE, 2024) y PCCRP (Harrison, s.f.). Existen también sistemas especializados como SIGN (Sign, 2025), ADA (American Diabetes Association, 2008) y ACCF/AHA (Granhholm et al., 2019; Greenland et al., 2010), diseñados para situaciones clínicas específicas.

El sistema GRADE (Prinsen et al., 2018) ha ganado importancia, siendo cada vez más empleado; aproximadamente 90 instituciones internacionales han adoptado esta metodología para la valoración de la evidencia y el desarrollo de guías de práctica clínica (Montagut, 2022). El consorcio COSMIN recomienda el uso exclusivo de GRADE en revisiones sistemáticas por varias razones: optimiza la reproducibilidad, proporciona un marco integral y sistemático para todo el proceso de revisión, identifica áreas con evidencia robusta y aquellas que requieren más investigación, y su difusión global junto con la sencillez de su terminología facilitan su aplicación (Mokkink et al., 2018).

COSMIN utiliza una adaptación modificada de GRADE para calificar la calidad de la evidencia en revisiones sistemáticas de instrumentos de evaluación (Mokkink et al., 2018). Esta adaptación incorpora cuatro de los cinco factores originales de GRADE para degradar la calidad de la evidencia, en función de la fiabilidad de los resultados:

- Riesgo de sesgo: Evalúa la calidad metodológica de los estudios primarios.
- Inconsistencia: Se refiere a la variabilidad inexplicada de los resultados entre estudios.
- Imprecisión: Relacionada con el tamaño total de la muestra de los estudios.

- Evidencia indirecta: Presencia de evidencia de poblaciones distintas a la de interés.

La calidad de la evidencia se determina individualmente para cada propiedad de medición y cada instrumento de evaluación. Se parte de una presunción inicial de alta calidad, la cual se degrada progresivamente (a moderada, baja o muy baja) en uno o dos niveles por factor, si se detecta riesgo de sesgo, inconsistencia, imprecisión o evidencia indirecta. Se recomienda una evaluación secuencial de cada factor (Mokkink et al., 2018).

7.1.4 Revisión psicométrica del Índice de Barthel y FIM

En esta revisión sistemática psicométrica se identificaron veinte versiones del Índice de Barthel y quince versiones de la escala FIM, diseñadas para evaluar la independencia y la funcionalidad en las AVD en personas con ictus mediante la inclusión de 35 estudios.

Los profesionales clínicos e investigadores deben considerar diversos factores, como el entorno y la población donde se utilizará el instrumento, sus dimensiones, el número de ítems y la evidencia demostrada en la evaluación de cada propiedad psicométrica, para decidir qué instrumento es el más adecuado para medir este constructo. Además, deben tenerse en cuenta el idioma y la cultura de las diferentes versiones, incluyendo la versión original de BI o FIM que se utilizará.

No existía ninguna versión de BI ni de FIM en la que se evaluaran todas las propiedades psicométricas. Las propiedades evaluadas en la mayoría de los estudios fueron la sensibilidad, la fiabilidad, la prueba de hipótesis y la consistencia interna.

7.1.4.1 Calidad metodológica

Hubo cinco estudios (Hachisuka et al., 1997a; Yang et al., 2022; Prasad et al., 2018; Leung et al., 2007; Aminimalroaya et al., 2021) en los que la calidad metodológica se calificó como "adecuada", para la validez estructural (Leung et al., 2007; Aminimalroaya et al., 2021), debido al análisis factorial desarrollado, un análisis de Rasch y un buen tamaño de muestra. Para el error de medición (Yang et al., 2022), debido a las características similares de los usuarios y un tamaño de muestra suficiente de personas incluidas. Para la fiabilidad (Hachisuka et al., 1997a; Yang et al., 2022), debido a la estabilidad de los usuarios, el intervalo de tiempo apropiado y las condiciones de prueba similares. Para la validez de criterio (Prasad et al., 2018), debido a una buena correlación entre versiones; y para la prueba de hipótesis (Prasad et al., 2018), debido a una comparación adecuada entre escalas.

Hubo diez estudios (Shah et al., 1989; van Hartingsveld et al., 2006; Zhou et al., 2021; Aminalroaya et al., 2021; Naghdi et al., 2016; Sandaria et al., 2001; Adunsky et al., 2002; Stillman et al., 2009; Bohannon et al., 2003; Nayar et al., 2016) donde la puntuación de calidad metodológica se calificó solo como "muy buena" para todas sus propiedades psicométricas analizadas.

A pesar de estos cinco estudios con una calidad metodológica media-alta, la calidad metodológica de la mayoría de los estudios no fue la esperada, resultando "inadecuada" o "dudosa" en muchas versiones y propiedades debido a la falta de evaluación de contenido, y en el caso de la validez de contenido, debido a la relevancia del ítem en la población de estudio.

En el caso de la validez estructural, la baja calidad metodológica se debe a que no se calculó para cada dimensión por separado o a la falta de descripción del porcentaje de gestión de los valores faltantes. Para la validez transcultural, la principal razón de la baja calidad es la pequeña muestra de personas y la falta de comparación entre los diferentes grupos después de realizar el análisis transcultural. En el caso de la fiabilidad y la validez de criterio, la baja calidad metodológica se debe a la falta de descripción de los porcentajes y la gestión de los valores faltantes y a la fase posterior al ictus, ya que la mayoría de los estudios se realizaron en personas en fase aguda que no estaban neurológicamente estables en el momento de la evaluación. Finalmente, la baja calidad metodológica en la prueba de hipótesis resultó de que las hipótesis debían formularse a priori, aunque se podía deducir lo que se esperaba encontrar en ese estudio.

Ira-BI (Aminalroaya et al., 2021), Ita-BI (Galeoto et al., 2019), Tel-BI (Prasad et al., 2018), Int-BI (Sandaria et al., 2001), BI-8 (van Hartingsveld et al., 2006), BI-16 (Jansa et al., 2004), G-BI (Hachisuka et al., 1997a), Tel-mBI (Shinar et al., 1987), PMA-mBI (Shinar et al., 1987), Chi-MBI (Leung et al., 2007), Bra-MBI (Cincura et al., 2009), Ira-MBI (Aminalroaya et al., 2021), Tur-MBI (Kuşukdeveci et al., 2000), PET-MBI (Ohura et al., 2017), Tur-FIM (Kuşukdeveci et al., 2001), Per-FIM (Naghdi et al., 2016), FIM-Mot 3-Irv (Hachisuka et al., 1997a), Int-3item-FIM (Sandaria et al., 2016), Alpha-FIM (Stillman et al., 2009), FIM-5 ítems (Bohannon et al., 2003), FIM-Acu (Roberts et al., 2016), SR-FIM-Mot (Vadassery et al., 2019), Spa-FIM+FAM (Colomer et al., 2023), Jap-FIM+FAM (Miki et al., 2016) y UK-FIM+FAM (Nayar et al., 2016) solo se evaluaron en un estudio. Por lo tanto, se requieren estudios adicionales para evaluar las propiedades psicométricas de estas versiones.

Las propiedades psicométricas calificadas con mayor frecuencia como "muy buenas" fueron la consistencia interna (31 veces) y la sensibilidad (31 veces), seguidas de la prueba de hipótesis (24 veces) y la validez de criterio (18 veces). Las propiedades calificadas con mayor frecuencia como "inadecuadas" fueron principalmente la fiabilidad (28 veces), seguida de la sensibilidad (6 veces), la prueba de hipótesis y la validez de criterio (5 veces cada una), como se muestra en la tabla 3.

7.1.4.2 Calidad de la evidencia

La calidad de la evidencia fue «baja» o «muy baja» para la mayoría de las propiedades de medición analizadas en todas las versiones incluidas en esta revisión sistemática.

Validez del contenido

La validez de contenido solo se presentó en el BI-SS (Lee et al., 2014), y se calificó como "dudosa" debido a la probabilidad de que la propiedad de medición real se aleje de la estimada, aun cuando existe la posibilidad de que sea sustancialmente diferente. Además, la calidad de este resultado se calificó como "-", lo que indica que no se han evaluado la relevancia, la comprensibilidad ni la exhaustividad de los ítems en las diferentes versiones.

Validez estructural

La validez estructural se evaluó en trece estudios para catorce versiones: BI (2006), BI-8, Ita-BI (Galeoto et al., 2019), BI-16 (Jansa et al., 2004), BI-SS (Lee et al., 2014), MBI (Yang et al., 2021), Chi-MBI (Leung et al., 2007), Ira-MBI (Aminalroaya et al., 2021), Tur-MBI (Kuşukdeveci et al., 2000), FIM (Heinemann et al., 1994), Tur-FIM (Kuşukdeveci et al., 2001), Per-FIM (Naghdi et al., 2016), Alpha-FIM (Stillman et al., 2009) y UK-FIM+FAM (Nayar et al., 2016). Sólo seis versiones (BI, BI-8, MBI, FIM, Alpha-FIM y UK-FIM+FAM) obtuvieron una calidad de evidencia calificada como "muy buena". En dos estudios (Chi-MBI, Ira-MBI) se calificó como "adecuado". Para dos estudios (Ita-BI, Tur-FIM) se calificó como "dudoso"; y cuatro estudios (BI-16, BI-SS, Tur-MBI, Per-FIM) se calificaron como "inadecuados". Para todos ellos, excepto UK-FIM+FAM, la calidad de los resultados se calificó como "?" porque se proporciona poca información sobre el procedimiento de evaluación de esta propiedad.

Consistencia interna

Finalmente, para lograr coherencia interna, treinta y un estudios evaluaron esta propiedad psicométrica para veinticinco versiones: BI (Shah et al., 1989; Hachisuka et al., 1997a; al., 2019), Per-BI (Oveisgharan et al., 2006), Tel-BI (Prasad et al., 2018), Int-BI (Sandaria et al., 2001), SR-BI (Hachisuka et al., 1997a), BI-5 (Park, 2018; Hsueh et al., 2002), BI-8 (van Hartingsveld et al., 2006), BI-16 (Jansa et al., 2004), G-BI (Hachisuka et al., 1997a), MBI (Shah et al., 1989), Chi-MBI (Leung et al., 2007), Ira-MBI (Aminalroaya et al., 2021), Tur-MBI (Kuşukdeveci et al., 2021), Tur-FIM (Kuşukdeveci et al., 2001), Per-FIM (Naghdi et al., 2016), FIM-Mot-13 (Hachisuka et al., 1997a), Int-3item-FIM (Sandaria et al., 2001), Alpha-FIM (Stillman et al., 2009), Artículos FIM-5 (Bohannon et al., 2003), FIM-Acu (Roberts et al., 2016), Spa-FIM+FAM (Colomer et al., 2023), Jap-FIM+FAM (Miki et al., 2016) y UK-FIM+FAM (Nayar et al., 2016). Todos ellos obtuvieron una calidad de evidencia calificada como "muy buena" y mostraron una calificación de criterio "+", indicando que el alfa de Cronbach es $\geq 0,70$.

7.1.4.3 Principales conclusiones

El índice de Barthel original (IB) (Loewen et al., 1988; Shah et al., 1989; Wolfe et al., 1991; Hachisuka et al., 1997a; van der Putten et al., 1999; Gosman-Hedstrom et al., 2000; Green et al., 2001; Hsueh et al., 2001). 2001; Hsueh Hsueh et al., 2002; van Hartingsveld et al., 2006; Hsieh et al., 2007; Sarker et al., 2012; Zhou et al., 2021; Yang et al., 2022) ha sido la versión del índice de Barthel incluida en más estudios (14 en total, y tres traducciones). La segunda versión fue FIM Motor de 13 ítems (FIM-Mot-13) (Hachisuka et al., 1997a; van der Putten et al., 1999; Gosman-Hedstrom et al., 2000; Hsueh et al., 2002), apareciendo en 4 estudios. La versión con más traducciones es el índice de Barthel modificado por Shah (MBI) con cuatro traducciones (Chi-MBI (Leung et al., 2007), Bra-MBI (Cincura et al., 2009), Ira-MBI (Aminalroaya et al., 2021), Tur-MBI (Kuçukdeveci et al., 2000)), seguido del índice de Barthel (BI) con tres traducciones en 4 estudios; (Ira-BI (Aminalroaya et al., 2021), Ita-BI (Galeoto et al., 2019), Per-BI (Oveisgharan et al., 2006; Naghdi et al., 2016)). La versión del índice de Barthel o FIM con más propiedades psicométricas evaluadas ha sido la versión original del índice de Barthel (Loewen et al., 1988; Shsh et al., 1989; Wolfe et al., 1991; Hachisuka et al., 1997a; Van der Putten et al., 1999; Gosman-Hedstrom et al., 2000; Green et al., 2001; Hsueh et al., 2001; Hsueh et al., 2002; Van Hartingsveld et al., 2006; Hsieh et al., 2007; Sarker et al., 2012; Zhou et al., 2021; Yang et al., 2022) en los cuales se han evaluado siete propiedades psicométricas, casi todas ellas, excepto validez de contenido y validez transcultural, que se evalúa en las traducciones de esta versión. Se realizaron seis versiones de esta versión con seis propiedades psicométricas evaluadas, de las cuales casi todas fueron evaluadas, excepto la validez de contenido y la validez transcultural. BI-5 (Park, 2018; Hsueh et al., 2002), BI-SS (Lee et al., 2014; Lee et al., 2017), MBI (Shah et al., 1989; Yang et al., 2021; Yang et al., 2022), Ira-MBI (Aminalroaya et al., 2021), FIM (Heinemann et al., 1994; van der Putten et al., 1999) y Per-FIM (Naghdi et al., 2016).

Esto convierte a esta versión del Índice de Barthel (Índice de Barthel Original) en la más validada y recomendada para evaluar la independencia y la funcionalidad en las actividades de la vida diaria en personas con ictus, tanto en la práctica clínica como en la investigación. La versión original del Índice de Barthel Original se destaca como la más válida y recomendada porque, según los resultados de la presente investigación, es la más utilizada, con más propiedades psicométricas analizadas y ha demostrado mejores puntuaciones. Esto convierte a esta versión del Índice de Barthel en la más validada y recomendada, así como en el estándar de oro para evaluar la independencia y la funcionalidad en las actividades de la vida diaria en personas con ictus.

7.1.5 Revisión psicométrica de Fugl-Meyer Assessment

Se identificaron veintidós versiones de FMA diseñadas para evaluar la función sensoriomotora de las extremidades superiores en personas con ictus mediante la inclusión de 36 estudios.

Se incluyeron 11 traducciones en esta revisión sistemática psicométrica, mostrando heterogeneidad entre ellas, tanto en la población incluida como en el número de propiedades psicométricas evaluadas. Sin embargo, mostraron resultados similares en cuanto a las diferentes propiedades psicométricas. Además, los resultados fueron muy positivos, lo que indica que las diferentes traducciones se realizaron con métodos adecuados y que su versión final puede utilizarse en el país de destino, sin afectar la fiabilidad ni la sensibilidad de la prueba. La traducción más completa es la de FMA-Urd (Ikram et al., 2022), que evalúa 7 propiedades psicométricas y presenta puntuaciones psicométricas positivas. No existía una versión de FMA para medir la función sensoriomotora de las extremidades superiores, para la cual se evaluaron todas las propiedades psicométricas.

La versión completa de 50 ítems de FMA (Chen et al., 2014; Hsueh et al., 2008; Lin et al., 2021; Lin et al., 2004; Malouin et al., 1994; Sanford et al., 1993; See et al., 2013), y FMA-UE (Amano et al., 2020; De Weerdt, 1985; Hernández et al., 2019; Hsieh et al., 2009; Kim et al., 2012; Kim et al., 2016; Lin et al., 2009; Michaelsen et al., 2011; Rabadi et al., 2006; Sullivan et al., 2011; Tauchi et al., 2022; Van der Lee et al., 2011; Wei et al., 2011; Woodbury et al., 2007) han sido las dos versiones con mayor volumen de estudios, con más propiedades psicométricas analizadas y las más traducidas; además de ser las dos versiones más antiguas de FMA. Esto las convierte en las versiones de FMA más extendidas y utilizadas por clínicos e investigadores para evaluar la función sensoriomotora de la extremidad superior en personas post-ictus. Esto hace que estas dos (Fugl-Meyer Assessment (Versión completa, 50-items, FMA) y Subescala de miembro superior de Fugl-Meyer Assessment (FMA-UE)) sean las más ampliamente validadas y las más sugeridas para su uso en la evaluación de la función sensoriomotora de la extremidad superior en personas con ictus tanto en la práctica clínica como en la investigación. Destacando la versión completa inicial de Fugl-Meyer Assessment (Versión completa, 50-items, FMA) como la más válida y recomendada porque, con base en los resultados de la presente investigación, es la versión más utilizada con más propiedades psicométricas analizadas y ha demostrado mejores puntuaciones. Esto hace que esta versión de FMA sea la más ampliamente validada y recomendada como la mejor versión y el "estándar de oro" para evaluar la función sensoriomotora de las extremidades superiores en personas con ictus.

Esto no significa que no se deban utilizar las otras 20 versiones, ya que son muy útiles en ciertos contextos y entornos específicos. Por ejemplo, en Italia se pueden utilizar tanto FMA-Ita (Cecchi et al., 2021) como FMA-EU-Ita (Hochleitner et al., 2023); el evaluador debe elegir cuál seleccionar en función del constructo que desee evaluar o del tiempo que desee dedicar. Al elegir entre FMA-Ita y

FMA-UE-Ita, será útil consultar esta revisión sistemática psicométrica, ya que se recopilan las puntuaciones de las propiedades psicométricas de ambas versiones.

Otra posibilidad que ofrecen las versiones encontradas en esta revisión es poder utilizar una versión reducida de FMA, que permite una evaluación rápida y fiable, pudiendo elegir entre FMA-37 (Chen et al., 2014; Hsieh et al., 2007; Lin et al., 2021) y FMA-12 (Chen et al., 2014; Fu et al., 2012; Hsieh et al., 2007; Hsueh et al., 2008; Lin et al., 2021) si se desea una evaluación general de FMA; o por categorías de la escala, como las partes proximales del miembro superior (FMA-UE-SE) (See et al., 2013; Wei et al., 2011), las partes distales del miembro superior (FMA-UE-w/h) (Page et al., 2012; Page et al., 2015; Perch et al., 2015; See et al., 2013; Wei et al., 2011), la subescala sensitiva (FMA-S) (Lin et al., 2004; Sullivan et al., 2011), o la subescala de miembro superior de 33 ítems (FMA-UE) (Amano et al., 2020; De Weerd, 1985; Hernández et al., 2019; Hsieh et al., 2009; Kim et al., 2012; Kim et al., 2016; Lin et al., 2009; Michaelson et al., 2011; Rabadi et al., 2006; Sullivan et al., 2011; Tauchi et al., 2022; Van der Lee et al., 2011; Wei et al., 2011; Woodbury et al., 2007), que es un método rápido de evaluación exclusiva de la función motora del miembro superior, y que cuenta con un gran respaldo de múltiples estudios y propiedades psicométricas evaluadas. Si un terapeuta dispone de muy poco tiempo, se recomienda usar la versión FMA-12, ya que realiza una evaluación completa con FMA, no solo una subescala o una porción corporal. Además, se ha evaluado en cinco estudios con un total de 948 participantes, donde se evaluaron las propiedades psicométricas de esta versión y las compararon con otras versiones y escalas de evaluación, con las que presenta correlaciones positivas. A su vez, esta versión muestra buenos niveles de fiabilidad y sensibilidad.

A lo largo de los diferentes estudios incluidos en esta revisión, se ha evaluado un tamaño muestral amplio, con aspectos muy diversos, como fase aguda/crónica, secuelas leves y graves, evaluación del progreso con evaluación inicial y seguimientos con evaluaciones periódicas; y la FMA (Chen et al., 2014; Hsueh et al., 2008; Lin et al., 2021; Lin et al., 2004; Malouin et al., 1994; Sanford et al., 1993; See et al., 2013) se ha mostrado útil para la evaluación del constructo en las diferentes situaciones que se pueden presentar con la variabilidad de estos aspectos; por tanto, es una escala válida y fiable para aplicar en las diferentes situaciones que se pueden presentar en la población post-ictus; aunque con mayor preferencia por el uso de la versión completa de 50 ítems de FMA; y FMA-EU.

La principal implicación de esta revisión es ofrecer una lista de las diferentes versiones de la FMA, con información sobre sus características, como el número de ítems, las partes que evalúa, el material necesario, el idioma, las propiedades psicométricas evaluadas para cada una de estas propiedades y sus puntuaciones correspondientes. Esto permite a los profesionales clínicos seleccionar la escala que mejor se adapte a sus recursos y necesidades de atención médica, y consultar rápidamente el artículo original para resolver cualquier duda. Asimismo, permite a un terapeuta clínico que dude entre usar

dos versiones, poder compararlas rápidamente consultando el presente estudio. Al mismo tiempo, permite a los investigadores disponer de la versión de referencia de la FMA y basarse en esta investigación para justificar su uso en futuras investigaciones; así como consultar las propiedades psicométricas de cualquier versión de FMA, lo que ahorrará mucho tiempo y permitirá un uso mucho más válido y justificado de dicha herramienta.

7.2 Validación de Escalas de Evaluación

Además de la lista de verificación de riesgo de sesgo para revisiones sistemáticas, la guía COSMIN establece un conjunto de estándares de validación para escalas de evaluación en ciencias de la salud. El propósito de estos estándares es uniformar el proceso de validación de dichos instrumentos, promoviendo una metodología consistente para la evaluación de las propiedades psicométricas inherentes (Terwee et al., 2018).

En este contexto, la presente tesis doctoral aplicó los estándares COSMIN para la validación de la versión española del cuestionario NSDA-R, es decir, de la Escala Nottingham de Vestido en Ictus-Revisada (ENVI-R) (Artículo III), así como de la versión en español de la escala FMA (Artículo IV). El proceso abarcó la evaluación de diversas propiedades psicométricas como validez estructural, consistencia interna, fiabilidad, error de medición, validez de criterio, etc.

En el desarrollo de estos estudios, se documentó detalladamente la metodología empleada y se presentaron los resultados obtenidos para cada una de las propiedades psicométricas investigadas.

Es relevante señalar que, de acuerdo con las directrices de COSMIN, la validez de contenido se considera la propiedad psicométrica de mayor relevancia en el proceso de validación de una escala. Sin embargo, en esta investigación no fue posible evaluar dicha propiedad.

7.2.1 Traducción y adaptación transcultural de cuestionarios

La traducción y adaptación de instrumentos de medida, trasciende la mera traducción lingüística del idioma de origen al de destino. Constituye un proceso metodológico riguroso cuyo objetivo fundamental es garantizar la equivalencia psicométrica entre la versión original y la adaptada (Ferrer, 2016; Montagut, 2022; Wild et al., 2005). Esto implica que las puntuaciones obtenidas con el instrumento traducido deben ser equiparables y comparables a las alcanzadas con la versión nativa, asegurando así la validez y fiabilidad de las conclusiones realizadas a partir de dichas mediciones en el nuevo contexto cultural y lingüístico (Cecchi et al., 2021).

Para conseguir esta equivalencia instrumental, es imprescindible considerar críticamente al menos cuatro dimensiones interrelacionadas (Martín-Arribas, 2004):

- Marco cultural de adaptación: Se refiere al análisis exhaustivo de las características sociodemográficas, normativas y contextuales del grupo poblacional al que se dirige el instrumento. Este análisis es vital para identificar posibles sesgos culturales o conceptuales que puedan afectar

la comprensión de los ítems y, consecuentemente, la validez del constructo medido.

- Aspectos técnicos del desarrollo y adaptación del cuestionario: Engloba las decisiones metodológicas inherentes al proceso de traducción y retrotraducción, la depuración de ítems, el análisis de la pertinencia de las escalas de respuesta y la aplicación de métodos estadísticos avanzados para evaluar la equivalencia estructural y de medida.
- Administración del instrumento: Conlleva la estandarización de las condiciones bajo las cuales se aplica el cuestionario en el nuevo entorno, incluyendo las instrucciones proporcionadas a los participantes, el formato de presentación y cualquier variable que pueda influir en la respuesta y la obtención de datos consistentes.
- Representación de las puntuaciones: Implica la evaluación de la comparación de las propiedades psicométricas de las puntuaciones obtenidas con la versión adaptada frente a las del instrumento original. Esto incluye el análisis de la distribución de las puntuaciones, la fiabilidad y la validez, garantizando que el instrumento mantenga su capacidad para medir el constructo latente de manera consistente y precisa en la nueva cultura.

La implementación rigurosa de estas consideraciones es indispensable para que los instrumentos de medida adaptados puedan ser empleados con confianza en la investigación científica y la práctica profesional, permitiendo comparaciones válidas entre poblaciones y culturas.

7.2.2 Propiedades psicométricas de los cuestionarios

7.2.2.1 Validez estructural

La inclusión de análisis estadísticos para evaluar las características de escalas e ítems dentro de la validez de contenido es un punto de debate entre investigadores. Mientras que el grupo de trabajo del *Patient-Reported Outcomes Measurement Information System* (PROMIS) consideró la asignación de ítems como parte de esta evaluación (Magasi et al., 2012), la *International Society for Pharmacoeconomics and Outcomes Research* (ISPOR) optó por no incluirlos (Leidy et al., 2008; Messick, 1980; Patrick et al., 2011a; Patrick et al., 2011b; Rothman et al., 2009).

Nos alineamos con la perspectiva de Magasi et al. (2012) en que la prueba de la estructura interna de una herramienta de evaluación es fundamental, ya que permite identificar ítems que podrían no medir el constructo deseado. Por ello, COSMIN recomienda una evaluación secuencial: primero la validez de contenido de forma independiente, seguida por la validez estructural y, finalmente, la consistencia interna.

La validez estructural se evalúa comúnmente mediante análisis factorial o análisis de la teoría de respuesta al ítem (IRT/Rasch). Para determinar la validez estructural se realiza un análisis factorial sobre todos los ítems de un instrumento con el objetivo de identificar el número hipotético de subescalas y la agrupación de ítems dentro de estas (estudios de validez estructural). Para evaluar la unidimensionalidad por subescala, se llevan a cabo análisis de factores múltiples en los ítems de cada subescala de forma individual, con el fin de determinar si cada subescala mide un único constructo (estudios de unidimensionalidad). Si bien estos estudios son suficientes para interpretar los análisis de consistencia interna y los de IRT/Rasch, no proporcionan por sí mismos evidencia de validez estructural como parte de la validez de constructo (Montagut, 2022).

El panel Delphi determinó que la evaluación de la validez estructural puede realizarse tanto con el EFA como con el CFA. Sin embargo, se prefiere el CFA al EFA. Esto se debe a que el CFA permite verificar si los datos se ajustan a una estructura factorial hipotética preestablecida, mientras que el EFA se emplea cuando no existen hipótesis claras sobre las dimensiones subyacentes, lo que lo hace menos robusto para la prueba de hipótesis (Floyd et al., 1995).

Es importante señalar que la lista de verificación de COSMIN no aborda los requisitos específicos para los análisis factoriales, como la elección del tipo de EFA (análisis de componentes principales o análisis factorial común), la selección y justificación del método de rotación (ortogonal u oblicua) o la decisión sobre el número de factores relevantes.

7.2.2.2 Consistencia interna

La consistencia interna es una propiedad métrica crucial en cuestionarios diseñados para medir un único constructo subyacente a través de múltiples ítems. Sin embargo, esta propiedad no es relevante para cuestionarios cuyos ítems representan distintos aspectos de un fenómeno clínico complejo y no necesariamente correlacionados (Bollen, 1989; Streiner, 1994).

Cuando la consistencia interna es una propiedad relevante, es necesario aplicar un análisis de componentes principales o un análisis factorial para determinar si los ítems conforman una única escala (dimensión) o múltiples dimensiones. Si no existe una hipótesis previa sobre la dimensionalidad de un cuestionario, puede emplearse el análisis exploratorio de componentes principales o el EFA. No obstante, si se dispone de una hipótesis clara sobre la estructura factorial, se recomienda el CFA (Bollen, 1989; Streiner, 1994).

Una vez que se ha determinado el número de escalas o subescalas, se debe calcular el alfa de Cronbach para cada una de ellas por separado, ya que es una medida adecuada de consistencia interna. Un valor bajo de alfa de Cronbach sugiere una falta de correlación entre los ítems de una escala, mientras que un valor muy alto puede indicar correlaciones excesivas, lo que podría implicar

redundancia entre uno o más ítems. Nunnally y Bernstein (1994) propusieron un criterio de 0.70-0.90 como indicador de buena consistencia interna.

7.2.2.3 Fiabilidad

El panel Delphi alcanzó un consenso sobre los métodos estadísticos preferidos para evaluar la fiabilidad de medidas de resultado con diferentes tipos de escalas de puntuación: continuas, ordinales, dicotómicas o nominales. Los métodos estadísticos seleccionados para esta evaluación fueron el ICC y el Kappa de Cohen ponderado (Montagut, 2022).

El ICC es el parámetro de fiabilidad más idóneo y comúnmente utilizado para medidas continuas. Es frecuente que muchos autores no especifiquen el tipo de ICC empleado, si un ICC para consistencia (ICC consistency) o un ICC para el grado de concordancia (ICC agreement) (McGraw y Wong, 1996). Según McGraw y Wong (1996), el método preferido es el ICC agreement, ya que las diferencias sistemáticas se consideran parte del error de medición.

Por otro lado, el coeficiente Kappa de Cohen ponderado se utiliza para medidas ordinales. Cabe destacar que, cuando se aplican ponderaciones cuadráticas, el coeficiente Kappa de Cohen ponderado es idéntico al ICC agreement (Cohen, 2013). Es importante señalar que el coeficiente de correlación de Pearson se considera inadecuado para la evaluación de la fiabilidad, dado que no tiene en cuenta las diferencias sistemáticas. De manera similar, el porcentaje absoluto de concordancia tampoco es apropiado, ya que no se ajusta por la correspondencia atribuible al azar (Cohen, 2013).

7.2.2.4 Error de medición

Inicialmente, el panel Delphi designó el Error Estándar de Medición (SEM) como el método principal para evaluar el error de medición. En cuanto a otros métodos complementarios, se observó un porcentaje reducido de acuerdo con el Cambio Mínimo Detectable (SDC) (20%) y los Límites de Acuerdo (LoA) (28%). A pesar de estos bajos niveles de consenso, el comité directivo decidió incorporar tanto el SDC como los LoA, como procedimientos válidos para la medición del error, incluyéndolos en la lista de verificación. El panel Delphi estableció que, para puntuaciones continuas, el SEM, SDC o LoA eran los métodos estadísticos preferidos. En contraste, para puntuaciones ordinales, dicotómicas o nominales, se optó por el porcentaje de concordancia específico (positivo y negativo). Adicionalmente, el comité consideró el coeficiente de variación (CV) como el método estadístico óptimo para evaluar el error de medición en escalas con puntuaciones de proporción o porcentaje, proponiendo su aceptación como método estándar para este fin (Montagut, 2022).

COSMIN categorizó como "muy bueno" el error de medición calculado mediante el SEM, SDC o LoA para puntuaciones continuas, y el porcentaje de concordancia (positivo y negativo) para puntuaciones dicotómicas, nominales u

ordinales. Por otro lado, COSMIN consideró "inadecuado" el cálculo del SEM basado en el alfa de Cronbach o derivado de la desviación estándar de otra población, ya que el alfa de Cronbach se utiliza a menudo incorrectamente en lugar del ICC para estimar el error de medición. Asimismo, el alfa de Cronbach se aplica erróneamente cuando la escala no es unidimensional o se fundamenta en un modelo formativo, lo que impide su interpretación válida (Montagut, 2022).

7.2.2.5 Validez de criterio

El panel Delphi concluyó que no existe un "estándar de oro" universalmente aceptado para los instrumentos de evaluación en el ámbito de la salud. Se debatió si la propiedad de validez de criterio debía ser incluida en la lista de verificación de COSMIN.

Finalmente, se estableció una única excepción para la existencia de un "estándar de oro": cuando un instrumento abreviado se compara con su versión original completa (Elsman et al., 2024; Montagut, 2022).

7.2.2.6 Prueba de hipótesis

La prueba de hipótesis es un componente fundamental para establecer la dirección y magnitud de correlaciones o diferencias, permitiendo inferir su alineación con los constructos bajo estudio. Un mayor número de hipótesis que confirman la correspondencia entre los datos y las predicciones a priori fortalece la evidencia de la validez de constructo.

El panel Delphi acordó que las hipótesis específicas deben formularse a priori. Estas hipótesis deben predecir diferencias medias esperadas entre grupos conocidos o correlaciones esperadas entre las puntuaciones del instrumento y otras variables relevantes. Es crucial especificar la dirección esperada (positiva o negativa) y la magnitud (absoluta o relativa) de estas correlaciones o diferencias (Altman, 1990; Messick, 1975).

Las hipótesis también pueden referirse a la magnitud relativa de las correlaciones o a las diferencias en las puntuaciones entre grupos. Al evaluar estas diferencias, su significación estadística es menos crítica, ya que depende del tamaño de la muestra.

El panel Delphi recomendó evitar el uso de valores p en las hipótesis, argumentando que no es relevante examinar si las correlaciones o diferencias difieren estadísticamente de cero (Norman, 1989). La magnitud de la diferencia es más importante que la significancia estadística entre los grupos, dado que esta última está influenciada por el número de participantes en cada grupo. Por lo tanto, se prefiere que las pruebas de hipótesis formales se basen en la magnitud esperada de las correlaciones y diferencias, en lugar de los valores p .

Para la lista de verificación de COSMIN, la propiedad psicométrica de prueba de hipótesis se clasifica según si el instrumento de medida de resultados se comparó con otros instrumentos de evaluación similares (validez convergente) o si se comparó un mismo instrumento entre subgrupos (validez discriminativa) (Montagut, 2022).

7.2.2.7 Sensibilidad al cambio

La sensibilidad al cambio constituye una medida crucial de validez longitudinal, que evalúa la capacidad de un instrumento para detectar cambios reales a lo largo del tiempo. De manera análoga a la validez de constructo, la validez longitudinal debe ser determinada mediante la prueba de hipótesis predefinidas, que predigan correlaciones entre cambios en las medidas o diferencias esperadas en los cambios entre grupos conocidos. Esto no solo demuestra la idoneidad de un cuestionario para registrar cambios cuando estos ocurren, sino también su habilidad para discernir variaciones clínicamente importantes en contraste con el error de medición (Montagut, 2022).

Inicialmente, el panel Delphi propuso considerar la sensibilidad al cambio como una propiedad de medición independiente. Sin embargo, tras un debate exhaustivo, se acordó que la principal distinción entre la validez transversal (de constructo y de criterio) y la sensibilidad al cambio reside en que la validez se refiere a la adecuación de una puntuación única, mientras que la sensibilidad al cambio concierne a la validez de una puntuación de cambio. En consecuencia, se decidió que los estándares para esta propiedad de medición debían ser análogos a los establecidos para la validez de constructo y de criterio (Elsman et al., 2024; Terwee et al., 2018).

Las metodologías apropiadas para evaluar la sensibilidad al cambio son idénticas a las empleadas en la prueba de hipótesis y la validez de criterio, con la salvedad de que las hipótesis deben centrarse en la puntuación de cambio del instrumento.

El panel Delphi consideró inapropiados varios parámetros propuestos en la literatura para evaluar la capacidad de respuesta. Se alcanzó un consenso que desaconseja el uso del tamaño del efecto y otros métodos relacionados, como la respuesta media estandarizada, el coeficiente de sensibilidad al cambio S de Norman, y la estadística de eficiencia relativa, como medidas adecuadas para la sensibilidad al cambio. Del mismo modo, se consideraron inadecuadas la prueba t pareada (la cual mide un cambio significativo, no válido, y es dependiente del tamaño de la muestra) y la estadística de sensibilidad al cambio de Guyatt (que incorpora el Cambio Mínimo Clínicamente Importante (MIC), pero el panel Delphi dictaminó que este no se refiere a la validez de la puntuación de cambio) (Montagut, 2022; Terwee et al., 2018).

7.2.3 Validación de ENVI-R

El uso de una escala o instrumento de medida en una lengua diferente y usada en un país diferente del origen de creación de la escala requiere un proceso de traducción y adaptación cultural para disponer de una herramienta equivalente a la escala original.

Este proceso garantiza la valoración del mismo concepto en diferentes culturas, permitiendo la comparación de resultados (Beaton et al., 200; Ferraz et al., 1990; Guillemin, 1995). Durante el proceso de adaptación de la escala es importante resaltar que esta ha mantenido la equivalencia lingüística, semántica y conceptual con la versión original durante la traducción y adaptación transcultural en orden a mantener la apariencia y contenido de la nueva versión de la escala. Para ello, en esta investigación se siguió el protocolo multipasos recomendado por la guía ISPOR, en el que participaron ocho terapeutas ocupacionales internacionales que conformaron el comité de expertos, así como personas nativas anglosajonas que realizaron traducciones. Tras este proceso con traducciones de ida, traducciones de vuelta, unificaciones y armonizaciones de diferentes versiones, tanto en inglés como en español, se obtuvo la versión española de la escala NSDA-R, la que se denominó Escala Nottingham de Vestido en Ictus-Revisada (ENVI-R).

Siendo el vestido una AVD comúnmente afectada en personas que han padecido un ictus (Walker et al., 2012; Fletcher-Smith, 2011), existen pocas escalas de evaluación que profundicen en el vestido, concretamente UBDS (Endo et al., 2015; Ng et al., 2023; Suzuki et al., 2008), Jacket Test (Chen et al., 2017) y NSDA (Fletcher-Smith et al., 2010; Fletcher-Smith, 2011; Walker, 1992; Walker et al., 1990; Walker et al., 1991; Walker et al., 2003; Walker et al., 2004; Walker et al., 2012); siendo NSDA la más antigua, lo que la hace la escala del vestido más estudiada y empleada de todas en la población de personas supervivientes de ictus; además de ser la que se ha creado y utilizado en una cultura más similar a la española, la anglosajona; ya que UBDS y Jacket Test se han desarrollado en la cultura oriental.

Además, a juicio de los investigadores involucrados en esta investigación, NSDA es la más completa de todas, ya que evalúa una gran variedad de prendas de vestir, tiene en cuenta cuáles de estas son significativas para la persona que ha sufrido el ictus, descompone la colocación de cada prenda en diversos pasos, y analiza los errores cometidos durante la colocación de las mismas (Fletcher-Smith et al., 2010; Fletcher-Smith, 2011; Walker, 1992; Walker et al., 1990; Walker et al., 1991; Walker et al., 2003; Walker et al., 2004; Walker et al., 2012); mientras que UBDS contempla solo el vestido de una sola prenda de la parte superior del cuerpo (Endo et al., 2015; Ng et al., 2023; Suzuki et al., 2018), y la escala Jacket Test es una subescala de la escala Physical Performance Test (PPT), y únicamente evalúa vestirse con una chaqueta (Chen et al., 2017).

Una vez se realizó la traducción y adaptación transcultural de la escala ENVI-R se procedió a validar dicho instrumento, evaluando con esta herramienta a 110 personas que habían sufrido un ictus, que cumplían los criterios de inclusión y de

exclusión, y que aceptaron participar en el estudio. Tras un mes y medio de media, se reevaluó a 65 de ellos.

Una vez se evaluó y reevaluó a todos ellos, se analizaron los datos estadísticamente, obteniendo que el EFA presentó una solución de doce factores que explicaron el 84% de la varianza. Fueron eliminados dos ítems por una carga factorial insuficiente (<0.30). Siendo el primer análisis factorial exploratorio de la escala NSDA-R. Esta varianza es menor a la hallada en la escala UBDS, lo cual se puede deber a que la escala UBDS es una escala más reducida, con menos posibilidades de que surjan más factores (Suzuki et al., 2008).

Presentó una excelente consistencia interna con un valor de alfa de Cronbach de 0.921; mientras que, si elimina un ítem, la puntuación de alfa de Cronbach oscila de 0.918 a 0.923; siendo mayor a la de la escala UBDS (Suzuki et al., 2008); lo cual se debe a que ENVI-R ha seguido un proceso riguroso de traducción y adaptación transcultural que le permite seguir midiendo el constructo para el que fue creada la escala original.

La fiabilidad test-retest logró una puntuación de ICC de 0.92, la cual es considerada como excelente; sin embargo, resulta ligeramente menor a la escala original NSDA (ICC=0.99), siendo igualmente elevada y significativa. Los valores de fiabilidad obtenidos para ENVI-R son similares a los de otras escalas del vestido como UBDS (Ng et al., 2023; Suzuki et al., 2008) y Jacket Test (Chen et al., 2017); lo que sugiere que las tres escalas son fiables para emplearse en la evaluación del vestido de personas con ictus.

Mediante rho de Spearman se halló la correlación entre ENVI-R y la subescala de miembro superior de FMA, con un valor de 0.34; entre ENVI-R y la subescala de miembro inferior de FMA, con una puntuación de 0.39; y entre ENVI-R y FMA completa de 50 ítems, con una r de Spearman de 0.50. Estando en concordancia con las otras escalas de vestido (Chen et al., 2017; Ng et al., 2023), con puntuaciones moderadas y significativas de estas con la escala FMA, debido a que, aunque FMA es comúnmente empleada en la evaluación de personas que han sufrido un ictus, no evalúa el vestido. No se ha podido hallar correlación entre ENVI-R y UBDS o Jacket Test porque no hay más versiones españolas de estas escalas de evaluación del vestido en personas post-ictus.

Observando los diferentes hallazgos y la comparación con otras escalas, ENVI-R presenta valores similares para algunas propiedades psicométricas como fiabilidad, e incluso superiores, como es el caso de la validez de criterio y del análisis factorial. Con estos hallazgos, se puede afirmar que ENVI-R podría ser actualmente la única escala en español que evalúe de manera específica y exhaustiva el vestido en personas que han sufrido un ictus, presentando buenas propiedades psicométricas tras el proceso de traducción, adaptación transcultural y validación, por lo que se recomienda su uso en población española que haya sufrido un ictus y que presente dificultades para el vestido.

7.2.4 Validación de Fugl-Meyer Assessment

Dado que el miembro superior se ve afectado en un 60-80% de los supervivientes de ictus (Kwakkel et al., 2023), la evaluación de la función sensoriomotora del miembro superior tras un ictus es indispensable para poder plantear correctamente los objetivos de tratamiento y orientar la intervención hacia una recuperación de la funcionalidad de ese miembro superior, así como para poder medir la evolución de dicha función con el tiempo y el tratamiento.

Tras la realización de una revisión sistemática psicométrica analizando las diferentes versiones de FMA (Artículo I), y hallar que la versión original completa de 50 ítems es la más completa (de Blas-Zamorano et al., 2025); se procedió a buscar si dicha versión estaba disponible en población española. Para ello se acudió al origen de dicha escala, la Universidad de Goteborgs, donde se creó dicha escala, y donde está disponible la versión española de dicha escala completa de 50 ítems, la cual no había sido validada previamente en población española.

Por lo que para evaluar la función del miembro superior en personas que habían sufrido un ictus en España, se empleaban otras escalas que si habían sido validadas al español, como ARAT (Fernández-Solana et al., 2022) y Box and Block Test (Noce et al., 2024).

Para la validación de la versión española de FMA, se evaluó a 110 personas españolas que habían sufrido un ictus y que cumplían con los criterios de inclusión y de exclusión del estudio. Tras un periodo de entre 1 y 4 meses se reevaluó a 65 de ellos. Esto convierte a este estudio en el mayor estudio de FMA en población española, ya que los anteriores, habían empleado una muestra más reducida ($n=10$ (Barbosa et al., 2019) y $n=103$ (Ferrer, 2016)). Para la validación se analizaron las propiedades psicométricas de los datos obtenidos en dichas evaluaciones.

Se realizó un análisis factorial exploratorio, para conocer la validez estructural de la escala, obteniendo una solución de siete factores que explicó el 81% de la varianza total. Siendo este el primer análisis factorial exploratorio de la versión española de FMA, y también de la versión completa de 50 ítems de FMA.

La consistencia interna se halló con el alfa de Cronbach, el cual logró una puntuación de 0.982, lo que indica una excelente consistencia interna. Se calculó también dicho estadístico si se eliminaba un ítem, y la puntuación oscilaba de 0.982 a 0.983. Estos valores de consistencia interna fueron superiores a los obtenidos en otras versiones de FMA, como la versión japonesa (Amano et al., 2018; Tauchi et al., 2023) o la versión urdu (Ikram et al., 2022).

La versión española de FMA mostró fiabilidad test-retest excelente, mostrando una puntuación de ICC de 0.98; siendo similar e incluso superior a otras versiones de FMA como la versión japonesa o la versión coreana (Amano et al., 2018; Hsueh et al., 2008; Lin et al., 2021).

El error estándar de la medida fue de 3.17, mostrando una puntuación conveniente.

Se calculó la correlación entre FMA-Total y FMA-MS empleando Spearman rho, la cual fue de 0.97; y entre FMA-Total y FMA-MI, la cual fue de 0.89; mostrando una excelente validez de criterio.

También se halló correlación positiva significativa entre FMA y ENVI-R con una puntuación de 0.50. Siendo este el primer estudio en mostrar la correlación de la versión española de FMA con cualquier otra escala; y también el primer estudio en analizar la correlación de FMA con la escala ENVI-R.

Observando los diferentes hallazgos y la comparación con otras escalas, la versión española de FMA presenta valores similares e incluso superiores a otras versiones de FMA en otros idiomas o a otras escalas de evaluación del mismo constructo para la misma población (ARAT y Box and Block Test), para algunas propiedades psicométricas como consistencia interna (Amano et al., 2018; Ikram et al., 2022) o fiabilidad (Hsueh et al., 2008; Kim et al., 2021; Lin et al., 2021; Sandford et al., 1993; See et al., 2013). Con estos hallazgos, se puede afirmar que la versión española de FMA es una escala muy adecuada para evaluar la función sensoriomotora en población española que ha sufrido un ictus, presentando buenas propiedades psicométricas tras el proceso de validación, por lo que se recomienda su uso en población española que haya sufrido un ictus y que presente dificultades sensoriomotoras en miembros superiores y/o inferiores.

7.3 Limitaciones

La evaluación rigurosa de instrumentos de medición y herramientas de evaluación en el ámbito de la investigación científica es un pilar fundamental para garantizar la validez y fiabilidad de los hallazgos. Sin embargo, la aplicación de metodologías estandarizadas, como las propuestas por la lista de verificación de riesgo de sesgo de COSMIN o los criterios de buenas propiedades de Terwee y Prinsen, así como el marco GRADE, presenta una serie de desafíos inherentes que requieren una consideración crítica.

La lista de verificación de riesgo de sesgo de COSMIN, aunque diseñada para estandarizar la evaluación de la calidad metodológica de los estudios que informan sobre las propiedades de medición de los instrumentos de resultado informados por el paciente, adolece de varias limitaciones que comprometen su utilidad práctica y su aplicabilidad universal:

- Complejidad y barrera de entrada: Su extensión (116 elementos distribuidos en 10 categorías) impone una complejidad que podría obstaculizar su adopción por parte de revisores sistemáticos, investigadores y profesionales de la salud sin una formación especializada en propiedades de medición. Esta complejidad podría, en última instancia, limitar su utilidad en la práctica.
- Alcance restringido de aplicación: La lista de verificación se centra predominantemente en instrumentos de evaluación diseñados para cuantificar cambios a lo largo del tiempo. Consecuentemente, su aplicabilidad es limitada para instrumentos discriminatorios, herramientas predictivas o medidas relacionadas con la atención de la salud (p. ej., satisfacción de la persona usuaria, adherencia terapéutica).
- Efecto suelo y puntuación: Se ha observado un efecto suelo recurrente, donde estudios inicialmente considerados de alta calidad reciben puntuaciones bajas o incluso son clasificados como "inadecuados" debido a la penalización por un único ítem, una limitación del algoritmo de puntuación.
- Redundancia en la evaluación de fiabilidad: Se emplea un conjunto idéntico de ítems para evaluar la fiabilidad intra-observador, inter-observador y test-retest, lo que podría no capturar adecuadamente las especificidades de cada tipo de fiabilidad.
- Sesgos derivados de la cualificación del evaluador: La aplicación de la lista de verificación por parte de revisores sin experiencia en el desarrollo o evaluación de medidas ha dado lugar a la identificación de diversos sesgos. Esto subraya la necesidad de una especificación más precisa por parte de COSMIN sobre los criterios de calidad y los objetivos de las revisiones, evitando la realización de investigaciones cuyo fin es justificar la necesidad de nuevas medidas cuando las existentes son de alta calidad.

- Ausencia de imparcialidad en la evaluación: La confianza en las directrices sobre la calidad de instrumentos previamente revisados puede ser problemática, ya que la imparcialidad de los terceros evaluadores no siempre está garantizada. Esta cuestión se agrava cuando los revisores no diferencian adecuadamente los tipos de instrumentos, lo que puede llevar a calificaciones contradictorias sobre la misma herramienta de evaluación, minando la confianza en los resultados de las revisiones basadas en COSMIN.

Un estudio crítico de Szekeres (2018) señaló que, si bien existen aspectos positivos, la validez de contenido de la lista de verificación de COSMIN ha sido establecida únicamente por sus desarrolladores, requiriendo una evaluación independiente por expertos imparciales. Adicionalmente, se ha reportado una baja fiabilidad entre evaluadores para esta lista de verificación, con un porcentaje de acuerdo inferior al 80% en un tercio de las recomendaciones (Mokkink et al., 2010c).

Los criterios para buenas propiedades psicométricas propuestos por Terwee y Prinsen, aunque valiosos, también presentan limitaciones significativas:

- Integración de resultados heterogéneos: No proporcionan directrices claras sobre cómo sintetizar resultados de estudios con calidades metodológicas dispares para extraer una conclusión global sobre un instrumento.
- Evaluación inadecuada de propiedades específicas: Ciertas propiedades de medición, como la consistencia interna, a menudo no se evalúan o interpretan correctamente; por ejemplo, valores de Alfa de Cronbach superiores a 0.95, indicativos de redundancia entre ítems, no se califican explícitamente como negativos.
- Limitaciones en la validez de contenido: La información esencial para determinar la validez de contenido suele ser escasa.
- Dependencia de la opinión de expertos: Estos criterios se basan primordialmente en la opinión de expertos, dado que la evidencia empírica que respalde criterios de calidad explícitos en este campo es limitada.

El sistema GRADE constituye un marco robusto para evaluar la calidad de la evidencia y la fuerza de las recomendaciones. No obstante, su aplicación presenta ciertas limitaciones:

- Especificidad en el ámbito de aplicación: Se ha desarrollado para abordar preguntas sobre estrategias, intervenciones o políticas de gestión, no así para cuestiones relacionadas con el riesgo o el pronóstico.
- Dificultad en la aplicación a "Recomendaciones de buenas prácticas": La calificación de la calidad de la evidencia o la fuerza de las

recomendaciones para este tipo de directrices puede no ser ventajosa o, en algunos casos, resultar inútil.

- Brecha de conocimiento en la implementación: Los revisores y desarrolladores de guías a menudo carecen de una comprensión clara sobre la integración de GRADE en el proceso global de desarrollo de guías (Higgins, 2024; Schünemann et al., 2006a).
- Falta de experiencia en áreas específicas de la salud pública: La aplicación de GRADE a preguntas en salud pública, pruebas diagnósticas y sistemas de salud enfrenta desafíos particulares debido a la limitada experiencia en estas áreas (Schünemann et al., 2006b; Schünemann et al., 2008).
- Expectativas erróneas: GRADE no elimina la necesidad de emitir juicios sobre la calidad de la evidencia y la fuerza de las recomendaciones, sino que los expone de manera más sistemática y transparente. No debe esperarse que elimine los desacuerdos en la interpretación de la evidencia o en la toma de decisiones.

En las dos revisiones sistemáticas psicométricas (Artículo I y Artículo II) realizadas en esta tesis, es posible que se haya quedado algún artículo sin incluir, debido a la retricción idiomática, ya que solo se seleccionaron estudios en inglés y español.

Una de las limitaciones halladas en los estudios de validación realizados tanto para ENVI-R (Artículo III) como para FMA (Artículo IV), fue el tamaño de la muestra, el cual podría haber sido más amplio; además de la manera de obtención de la misma, ya que se realizó por muestreo de conveniencia, donde la persona superviviente del ictus aceptaba participar en el estudio de investigación desde su centro de rehabilitación.

No realizar las evaluaciones por parte de dos evaluadores independientes impidió evaluar la fiabilidad inter-evaluador.

Cabe destacar que todas las investigaciones realizadas en esta tesis han sido realizadas sin ninguna financiación, contando con la participación voluntaria de los colaboradores, lo cual influyó en los niveles de dedicación y tiempo dedicado a la colaboración en las diferentes investigaciones, ya que la conciliación familiar y laboral con proyectos altruistas como este, en ocasiones resulta altamente desafiante.

7.4 Implicaciones para la práctica clínica, para la investigación y recomendaciones para futuras líneas de investigación

Ante la ausencia de evidencia empírica consolidada para la selección óptima de instrumentos de evaluación y en respuesta a la complejidad inherente a esta tarea, el comité directivo de COSMIN desarrolló una metodología estandarizada para la realización de revisiones sistemáticas. El propósito primordial de esta metodología es identificar el instrumento de evaluación más apropiado para medir un constructo de interés específico en una población de estudio determinada. Estas revisiones sistemáticas proporcionan una descripción exhaustiva de las propiedades de medición de los cuestionarios, permiten la evaluación del riesgo de sesgo en los estudios individuales incluidos y fundamentan recomendaciones basadas en la evidencia para la selección del instrumento más adecuado para un fin específico.

Estos instrumentos de evaluación son cruciales para la identificación de aspectos relevantes de los problemas de salud en el individuo, abarcando dimensiones como el funcionamiento físico, emocional, social, de rol, cognitivo, la percepción general de la salud y el bienestar, así como síntomas específicos de diversas patologías.

El dominio exhaustivo de la metodología COSMIN y su aplicación en la realización de revisiones sistemáticas es esencial para la selección del instrumento idóneo y la facilitación de la validación de estas herramientas. La disponibilidad de un instrumento validado adecuadamente permite abordar problemas de salud con una base sólida, proporcionando a los profesionales sanitarios un recurso ineludible de gran aplicación clínica. Este recurso optimiza los procesos de valoración, diagnóstico, planificación estratégica, tipos de intervención y su evaluación, lo que se traduce en beneficios sustanciales en términos de salud de las personas, así como en la eficiencia de los recursos humanos y materiales. En el contexto de la investigación de esta tesis, la revisión sistemática del Artículo I permite la identificación de la versión óptima entre los cuestionarios Índice de Barthel y FIM, para la evaluación de la independencia y la funcionalidad en las AVD en personas que han sufrido un ictus; mientras que en el Artículo II se resalta la identificación de todas las versiones disponibles de la escala FMA y el análisis de sus propiedades psicométricas para hallar la versión más recomendada en la evaluación de la función sensoriomotora del miembro superior en personas que han sufrido un ictus; resaltando la trascendencia de los hallazgos de ambas revisiones sistemáticas psicométricas para la práctica clínica.

COSMIN también establece patrones metodológicos para la validación de cuestionarios en ciencias de la salud, buscando estandarizar este proceso al seguir una misma metodología y evaluar cada una de las propiedades psicométricas descritas en su guía. Estas directrices han sido aplicadas en las investigaciones llevadas a cabo en esta tesis doctoral, tanto para la validación de

la escala ENVI-R en población española que había sufrido un ictus, previamente actualizada, traducida y adaptada transculturalmente al castellano; como para la validación de la escala FMA en población española superviviente de ictus, la cual se considera la escala gold standard para la evaluación sensoriomotora en esta población.

Para investigaciones futuras, se sugiere la especificación y el consenso de directrices metodológicas estandarizadas para la validación de instrumentos de evaluación. Esta tarea podría ser encomendada a un comité de expertos o una comisión científica, basándose en la evidencia existente sobre propiedades de medición. La subsecuente difusión de estas pautas a la comunidad científica contribuiría a contrarrestar la heterogeneidad de información actual en este campo. Adicionalmente, se recomienda que las revisiones sistemáticas proporcionen información detallada sobre la metodología empleada, garantizando así la máxima transparencia.

Finalmente, la creación de programas de formación para profesionales interesados en la guía COSMIN, tanto para fines investigativos como clínicos, resultaría altamente beneficiosa. Complementariamente, el desarrollo de un manual conciso que resuma las pautas para su aplicación, tanto en revisiones sistemáticas como en procesos de validación de herramientas, facilitaría su implementación.

8. CONCLUSIONES

Conclusión general

El análisis de variables psicométricas genera una comparación exhaustiva de diferentes escalas o versiones de las mismas, recomendando la escala más rigurosa/completa para la evaluación de un constructo. La validación de las escalas de evaluación permite disponer de cuestionarios para la evaluación de variables específicas en población española que ha sufrido un ictus y, por tanto, permite una valoración más exhaustiva durante el proceso neurorrehabilitador.

Conclusiones específicas

Entre las diferentes versiones existente de las escalas BI y FIM, se detecta como la más recomendable para evaluar la funcionalidad e independencia en las AVDs de personas que han sufrido un ictus, la versión original del índice de Barthel, según los criterios de la metodología COSMIN.

La versión original completa de 50 ítems de FMA resulta la más recomendada para evaluar la función sensoriomotora del miembro superior en personas que han sufrido un ictus.

La actualización de la escala NSDA da lugar a una versión más ajustada a los tiempos actuales. La traducción, adaptación transcultural y validación al español de ENVI-R, y los excelentes valores psicométricos obtenidos, supone disponer de una escala validada y altamente adecuada para la evaluación del vestido, la cual es una AVD comúnmente afectada en la población española que ha sufrido un ictus.

La validación de la versión española de la escala Fugl-Meyer Assessment ha resultado en excelentes valores psicométricos que permiten disponer de una escala muy aconsejada para la evaluación de la función motora, equilibrio, rango articular y dolor de miembros superiores e inferiores afectos de personas supervivientes del ictus.

9. COMPETENCIAS ADQUIRIDAS

Competencias instrumentales

La dirección de esta tesis doctoral por parte de dos experimentados investigadores altamente cualificados, como el Dr. David Pérez Cruzado y el Dr. José Antonio Merchán Baeza, de los que aprender y nutrirse de forma intelectual, metodológica y práctica; así como la colaboración con diferentes investigadores de diversas partes de España, pertenecientes a diferentes universidades y grupos de investigación, así como colegas internacionales, ha tenido como desenlace la adquisición de varias competencias instrumentales que permiten al doctorando la realización de diferentes proyectos con una mayor autonomía.

La trayectoria investigadora durante la tesis doctoral me ha proporcionado una comprensión profunda del proceso de investigación en ciencias clínicas de la salud. He adquirido un conocimiento exhaustivo de la estructura de los artículos científicos, incluyendo la variabilidad en su extensión en función de las normativas editoriales de las revistas. Asimismo, domino el procedimiento de sometimiento de proyectos de investigación a un Comité de Ética, asegurando la aprobación necesaria para su ejecución.

Durante el desarrollo de la tesis doctoral, he cultivado una habilidad avanzada en la adaptación y presentación de manuscritos científicos a diversas publicaciones. Esta experiencia ha facilitado la familiarización con metodologías de adaptación específicas, la evaluación del factor de impacto de las revistas, y la selección estratégica de la revista de publicación más adecuada. En este contexto, he optimizado el uso de gestores de referencias bibliográficas como Zotero y Mendeley, comprendiendo la necesidad de ajustar los estilos de citación a las directrices de cada revista.

La faceta clínica de esta tesis implicó un riguroso análisis estadístico, lo que me permitió adquirir pericia en el manejo de programas estadísticos (principalmente SPSS) y en la creación de bases de datos para el procesamiento de resultados. Además, he desarrollado la capacidad de evaluar e interpretar propiedades psicométricas de escalas, tales como la validez estructural, consistencia interna, validez de criterio y fiabilidad test-retest entre otras.

La elaboración de múltiples artículos científicos me ha permitido observar una progresión significativa en la calidad de la redacción científica. Este progreso es atribuible a una comprensión integral de la estructura de investigación, la metodología, el análisis estadístico y el proceso de adaptación a las exigencias editoriales. Finalmente, la iteración de revisiones y adaptaciones solicitadas por los revisores, incluyendo la preparación de respuestas estructuradas ítem por ítem y modelos de respuesta al editor de la revista, ha consolidado mi dominio en el proceso de publicación científica y en la respuesta efectiva a la retroalimentación de pares.

Competencias interpersonales

La realización del proyecto clínico de esta tesis doctoral fue llevada a cabo en nueve centros de neurorrehabilitación, principalmente de la Región de Murcia, como Integra Daño Cerebral, Dacemur, Asociación DISMO, CMV Caridad Puerto Lumbreras, Caser Salud, Neuroavanzo, CIAN Grupo 5, Neuron y Fundación AISSE. Como terapeuta ocupacional, me he adaptado a los diferentes ambientes laborales y a los requisitos del puesto de terapeuta ocupacional de cada centro en el que se ha realizado la recogida de datos; reforzando las competencias personales y profesionales del doctorando. Sin embargo, previamente a la realización de esta tesis doctoral, nunca había experimentado realizar una recogida de datos en clínicas ajenas a mi centro de trabajo y en una empresa en el que no he prestado mis servicios como profesional sanitario, con instalaciones, organigramas, horarios y usuarios de perfiles variados. La ejecución de la fase clínica de esta tesis ha sido fundamental para el desarrollo de mis habilidades de comunicación interpersonal. Esta experiencia me permitió interactuar eficazmente con profesionales de la salud de diversas especialidades, directivos de distintas instituciones clínicas, y pacientes externos a mi ámbito de trabajo habitual, a quienes fue necesario entrevistar para la recolección de datos. Consecuentemente, valoro esta etapa como un catalizador significativo en el fortalecimiento de mis competencias interpersonales y comunicativas.

Competencias sistémicas

La participación en la recolección de datos en centros de neurorrehabilitación, específicamente con personas que han sufrido un ictus y presentan diversas secuelas y niveles de afectación, ha enriquecido significativamente el conocimiento previo en esta área. Esta experiencia ha facilitado una profundización de los fundamentos teóricos, así como la adquisición de conocimientos derivados de la interacción directa con los usuarios y sus familiares. Asimismo, la exposición a sus vivencias ha fomentado el desarrollo de habilidades empáticas, fundamentales en el ámbito clínico y de investigación.

La adscripción a un programa de doctorado durante estos años ha promovido el desarrollo de habilidades intrínsecas a la investigación científica, tales como la constancia, la resiliencia, el rigor metodológico y la organización sistemática. Este proceso ha subrayado la importancia de valorar los logros, desde los más pequeños hasta las grandes consecuciones, inherentes a la trayectoria académica y de investigación.

Un aspecto relevante de este proceso fue la dilatada espera de publicación de la investigación que fundamenta esta tesis. Si bien esta circunstancia supuso una demora, simultáneamente posibilitó la adquisición de competencias y conocimientos críticos para la carrera investigadora. Específicamente, el Artículo II, concebido como proyecto preclínico, fue la primera publicación derivada de esta tesis, siendo aceptado tras un riguroso proceso de revisión por pares que se extendió por más de seis meses y que incluyó tres rondas de revisiones. Posteriormente, se elaboró el Artículo I, el cual fue sometido a un extenso proceso de adaptación y se envió a más de diez revistas científicas. La considerable inversión de tiempo en la adecuación de ambos manuscritos a un total de más de veinte revistas generó un periodo de estrés emocional significativo, manifestado por sentimientos de frustración, inseguridad y desesperanza, lo que conllevó cuestionamientos sobre la propia capacidad académica. La subsiguiente aceptación del Artículo I mitigó estas emociones negativas, validando la calidad de la investigación realizada y proporcionando una lección fundamental en paciencia, tolerancia a la frustración y persistencia en la labor científica.

En retrospectiva, se concluye que la perseverancia, la paciencia, la confianza y la convicción constituyen requisitos indispensables para todo individuo que aspire a llevar a cabo una tesis doctoral en el ámbito de las ciencias de la salud.

10. PROSPECTIVA

La investigación doctoral se ha centrado en enriquecer el campo de las Ciencias de la Salud a través de la aplicación de metodologías rigurosas y la difusión de resultados en publicaciones científicas. Durante el periodo formativo del doctorando, la producción científica ha sido el resultado de la colaboración entre el doctorando, los directores de tesis y el tutor académico. Esta producción incluye:

- Dos revisiones sistemáticas psicométricas: Adheridas a los criterios COSMIN, Terwee y Prinsen, y GRADE, las cuales abordaron la evaluación de propiedades psicométricas de instrumentos de medida como la escala FMA, el Índice de Barthel y FIM.
- Actualización, traducción, adaptación transcultural y validación de la escala ENVI-R: Este proceso garantizó la pertinencia y aplicabilidad del instrumento en un nuevo contexto cultural y lingüístico, como es la población española que ha sufrido un ictus.
- Validación de la escala Fugl-Meyer Assessment: Se llevó a cabo la validación de esta escala para su aplicación en población española superviviente de ictus.

Los hallazgos derivados de esta tesis doctoral ofrecen directrices esenciales para futuras investigaciones en el ámbito de las Ciencias de la Salud, particularmente en la evaluación de personas que han experimentado un ictus. Se recomienda lo siguiente:

- Evaluación de la función sensoriomotora del miembro superior: Para estudios que evalúen esta variable en personas con ictus, se deberá emplear la versión de la escala FMA recomendada en la revisión sistemática psicométrica detallada en el Artículo II.
- Evaluación de la funcionalidad e independencia en las AVD: Investigadores interesados en esta área deberán considerar los resultados de la revisión sistemática psicométrica realizada mediante la metodología COSMIN, presentada en el Artículo I.
- Evaluación del desempeño en la AVD del vestido en personas que han sufrido un ictus: Para la población inglesa, se recomienda la versión actualizada de la escala del vestido NSDA-R. Para la población española, se recomienda la escala ENVI-R, la cual se encuentra disponible y validada, según lo expuesto en el Artículo III.
- Aplicación de la escala FMA en población española: Gracias a la validación realizada en el Artículo IV, la escala FMA puede ser utilizada tanto en la práctica clínica como en la investigación con población española post-ictus.

La adopción de estas recomendaciones contribuirá a mejorar la calidad de la evaluación de diversas variables en la población de personas que han sufrido un

ictus. Adicionalmente, estudios futuros deberían replicar la validación de las escalas ENVI-R y FMA para confirmar la consistencia de las variables que mostraron menor estabilidad en los análisis de esta tesis doctoral.

Actualmente, el doctorando no forma parte de ningún grupo de investigación, sin embargo, mantiene colaboraciones activas con diversos colegas y entidades como la Universidad de Murcia, donde ejerce como docente. Esta interacción fomenta el desarrollo de su futura labor investigadora en la línea de la clinimetría, aprovechando los conocimientos y habilidades adquiridos durante este proceso doctoral. Tras la obtención del título de Doctor, se explorarán activamente las oportunidades de becas ofrecidas por entidades públicas y privadas, con el objetivo de completar la formación postdoctoral en un centro de I+D nacional o, en su defecto, en una institución universitaria.

11. BIBLIOGRAFÍA

- Acquadro, C., Conway, K., Hareendran, A., Aaronson, N., & European Regulatory Issues and Quality of Life Assessment (ERIQA) Group. (2008). Literature review of methods to translate health-related quality of life questionnaires for use in multinational clinical trials. *Value in Health: The Journal of the International Society for Pharmacoeconomics and Outcomes Research*, 11(3), 509-521. <https://doi.org/10.1111/j.1524-4733.2007.00292.x>
- Afakir, S., & Martínez, D. (2017). *Intervención del terapeuta ocupacional en la rehabilitación de pacientes con daño cerebral adquirido*. 11(Extra 1), 107-112.
- Agudelo, L. S., Nieto, M. L., Montero, J. C., & Montoya Huertado, O. L. (2015). *Referentes teóricos de fisioterapia en neurorrehabilitación, una revisión exploratoria*. 9(1), 67-72.
- Aho, K., Harmsen, P., Hatano, S., Marquardsen, J., Smirnov, V. E., & Strasser, T. (1980). Cerebrovascular disease in the community: Results of a WHO collaborative study. *Bulletin of the World Health Organization*, 58(1), 113-130.
- Alarcón M, A. M., & Muñoz N, S. (2008). [Some methodological issues about measurements in health]. *Revista Medica De Chile*, 136(1), 125-130.
- Alberdi Odriozola, F., Iriarte Ibarrarán, M., Mendía Gorostidi, Á., Murgialdai, A., & Marco Garde, P. (2009). *Pronóstico de las secuelas tras la lesión cerebral*. 33(4), 171-181.
- Alcaraz, F. G., Espín, A. A., Martínez, A. H., & Alarcon, M. M. (2006). *Diseño de Cuestionarios para la recogida de información: Metodología y limitaciones*. 1(5), 232-236.
- Alconero-Camarero, A. R., -Romero, A. G., Sarabia-Cobo, C. M., & Arce, A. M.-. (2016). Clinical simulation as a learning tool in undergraduate nursing: Validation of a questionnaire. *Nurse Education Today*, 39, 128-134. <https://doi.org/10.1016/j.nedt.2016.01.027>
- Allegri, R. F. (2000). Atención y negligencia: Bases neurológicas, evaluación y trastornos. *Revista de Neurología*, 30(05), 491. <https://doi.org/10.33588/rn.3005.99645>
- Altman, D. G. (1990). *Practical Statistics for Medical Research* (0 ed.). Chapman and Hall/CRC. <https://doi.org/10.1201/9780429258589>
- Alonso-Coello, P., Oxman, A. D., Moberg, J., Brignardello-Petersen, R., Akl, E. A., Davoli, M., Treweek, S., Mustafa, R. A., Vandvik, P. O., Meerpohl, J., Guyatt, G. H., Schünemann, H. J., & GRADE Working Group. (2016a). GRADE Evidence to Decision (EtD) frameworks: A systematic and

transparent approach to making well informed healthcare choices. 2: Clinical practice guidelines. *BMJ (Clinical Research Ed.)*, 353, i2089. <https://doi.org/10.1136/bmj.i2089>

Alonso-Coello, P., Schünemann, H. J., Moher, J., Brignardello-Petersen, R., Akl, E. A., Davoli, M., Treweek, S., Mustafa, R. A., Rada, G., Rosenbaum, S., Morelli, A., Guyatt, G. H., Oxman, A. D., & GRADE Working Group. (2016b). GRADE Evidence to Decision (EtD) frameworks: A systematic and transparent approach to making well informed healthcare choices. 1: Introduction. *BMJ (Clinical Research Ed.)*, 353, i2016. <https://doi.org/10.1136/bmj.i2016>

Alotaibi, N. M., Reed, K., & Shaban Nadar, M. (2009). Assessments used in occupational therapy practice: An exploratory study. *Occupational Therapy in Health Care*, 23(4), 302-318. <https://doi.org/10.3109/07380570903222583>

Al-Senani, F., Al-Johani, M., Salawati, M., Alhazzani, A., Morgenstern, L. B., Seguel Ravest, V., Cuche, M., & Eggington, S. (2020). An Epidemiological Model for First Stroke in Saudi Arabia. *Journal of Stroke and Cerebrovascular Diseases: The Official Journal of National Stroke Association*, 29(1), 104465. <https://doi.org/10.1016/j.jstrokecerebrovasdis.2019.104465>

Amador G., A., & Tortosa M., C. (2018). *Participación de enfermería y otros profesionales en el proceso de rehabilitación en pacientes que han sufrido un ictus. Revisión de la literatura*. TecnoCampus. Escola Superior de Ciències de la Salut.

Amano, S., Umeji, A., Takebayashi, T., Takahashi, K., Uchiyama, Y., & Domen, K. (2020). Clinimetric properties of the shortened Fugl-Meyer Assessment for the assessment of arm motor function in hemiparetic patients after stroke. *Topics in Stroke Rehabilitation*, 27(4), 290-295. <https://doi.org/10.1080/10749357.2019.1701176>

Amano, S., Umeji, A., Uchita, A., Hashimoto, Y., Takebayashi, T., Takahashi, K., Uchiyama, Y., & Domen, K. (2018). Clinimetric properties of the Fugl-Meyer assessment with adapted guidelines for the assessment of arm function in hemiparetic patients after stroke. *Topics in Stroke Rehabilitation*, 25(7), 500-508. <https://doi.org/10.1080/10749357.2018.1484987>

American Diabetes Association. (2008). Standards of Medical Care in Diabetes—2008. *Diabetes Care*, 31(Supplement_1), S12-S54. <https://doi.org/10.2337/dc08-S012>

Andresen, E. M. (2000). Criteria for assessing the tools of disability outcomes research. *Archives of Physical Medicine and Rehabilitation*, 81(12 Suppl 2), S15-20. <https://doi.org/10.1053/apmr.2000.20619>

Andrews, A. W., & Bohannon, R. W. (2000). Distribution of muscle strength impairments following stroke. *Clinical Rehabilitation*, 14(1), 79-87. <https://doi.org/10.1191/026921500673950113>

AOTA. (2020). Occupational Therapy Practice Framework: Domain and Process-Fourth Edition. *The American Journal of Occupational Therapy: Official Publication of the American Occupational Therapy Association*, 74(Supplement 2), 7412410010p1-7412410010p87. <https://doi.org/10.5014/ajot.2020.74S2001>

Ara, M.J.N. & Saboya, P. R. (2006). *Métodos, diseño y técnicas de investigación psicológica*. UNED.

Argibay, J. C. (2006). Técnicas psicométricas. Cuestiones de validez y confiabilidad. *Subjetividad y Procesos Cognitivos*, 8, 15-33.

Argimon P., J. M. & Jiménez V., J. (2004). *Métodos de Investigación Clínica y Epidemiológica*. Masson.

Arias Cuadrado, A. (2009). *Rehabilitación del ACV: evaluación, pronóstico y tratamiento*. 70(3), 25-40.

Au-Yeung, S. S. Y., & Hui-Chan, C. W. Y. (2009). Predicting recovery of dextrous hand function in acute stroke. *Disability and Rehabilitation*, 31(5), 394-401. <https://doi.org/10.1080/09638280802061878>

Azzimondi, G., Bassein, L., Nonino, F., Fiorani, L., Vignatelli, L., Re, G., & D'Alessandro, R. (1995). Fever in acute stroke worsens prognosis. A prospective study. *Stroke*, 26(11), 2040-2043. <https://doi.org/10.1161/01.str.26.11.2040>

Badia, & Carné. (1998). [Evaluation of quality of life in clinical trials]. *Medicina Clínica*, 110(14), 550-556.

Bailey, R. R., Phad, A., McGrath, R., & Haire-Joshu, D. (2019). Prevalence of five lifestyle risk factors among U.S. adults with and without stroke. *Disability and Health Journal*, 12(2), 323-327. <https://doi.org/10.1016/j.dhjo.2018.11.003>

Baldelli, M. V., Boiardi, R., Ferrari, P., Bianchi, S., & Hunscoth Bianchi, M. (2004). OCCUPATIONAL THERAPY AND TOOLS FOR MONITORING THE EFFECTIVENESS OF THERAPY. *Archives of Gerontology and Geriatrics*, 38, 45-50. <https://doi.org/10.1016/j.archger.2004.04.008>

Balshem, H., Helfand, M., Schünemann, H. J., Oxman, A. D., Kunz, R., Brozek, J., Vist, G. E., Falck-Ytter, Y., Meerpohl, J., Norris, S., & Guyatt, G. H. (2011). GRADE guidelines: 3. Rating the quality of evidence. *Journal of*

Clinical Epidemiology, 64(4), 401-406.
<https://doi.org/10.1016/j.jclinepi.2010.07.015>

Barbosa, N. E., Forero, S. M., Galeano, C. P., Hernández, E. D., Landinez, N. S., Sunnerhagen, K. S., & Alt Murphy, M. (2019). Translation and cultural validation of clinical observational scales – the Fugl-Meyer assessment for post stroke sensorimotor function in Colombian Spanish. *Disability and Rehabilitation*, 41(19), 2317-2323.
<https://doi.org/10.1080/09638288.2018.1464604>

Barer, D. H., & Mitchell, J. R. (1989). Predicting the outcome of acute stroke: Do multivariate models help? *The Quarterly Journal of Medicine*, 70(261), 27-39.

Batista-Foguet, J. M., Coenders, G., & Alonso, J. (2004). [Confirmatory factor analysis. Its role on the validation of health related questionnaires]. *Medicina Clinica*, 122 Suppl 1, 21-27. <https://doi.org/10.1157/13057542>

Beaton, D. E., Hogg-Johnson, S., & Bombardier, C. (1997). Evaluating changes in health status: Reliability and responsiveness of five generic health status measures in workers with musculoskeletal disorders. *Journal of Clinical Epidemiology*, 50(1), 79-93. [https://doi.org/10.1016/s0895-4356\(96\)00296-x](https://doi.org/10.1016/s0895-4356(96)00296-x)

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, 25(24), 3186-3191. <https://doi.org/10.1097/00007632-200012150-00014>

Béjot, Y., Bailly, H., Durier, J., & Giroud, M. (2016). Epidemiology of stroke in Europe and trends for the 21st century. *Presse Medicale (Paris, France: 1983)*, 45(12 Pt 2), e391-e398. <https://doi.org/10.1016/j.lpm.2016.10.003>

Ben Shoham, A., Liberant-Taub, S., Sharon, M., & Zucker, I. (2019). The number of acute cerebrovascular events in Israel: A forecast until 2040. *Israel Journal of Health Policy Research*, 8(1), 67. <https://doi.org/10.1186/s13584-019-0337-1>

Berenguer Guarnaluses, L. J., & Pérez Ramos, A. (2016). *Factores de riesgo de los accidentes cerebrovasculares durante un bienio*. 20(5), 621-629.

Bergström, L., Irewall, A.-L., Söderström, L., Ögren, J., Laurell, K., & Moee, T. (2017). One-Year Incidence, Time Trends, and Predictors of Recurrent Ischemic Stroke in Sweden From 1998 to 2010: An Observational Study. *Stroke*, 48(8), 2046-2051.
<https://doi.org/10.1161/STROKEAHA.117.016815>

Bernal Jiménez, J. J. (2024). *Efectividad del uso combinado de terapia asistida por robot y estimulación eléctrica por corriente directa para la recuperación*

del miembro superior en el ictus [Universidad de Castilla la Mancha].
<https://ruidera.uclm.es/items/4816f896-a4c8-4c6c-a188-47404e419f93>

Bernhardt, J., Hayward, K. S., Kwakkel, G., Ward, N. S., Wolf, S. L., Borschmann, K., Krakauer, J. W., Boyd, L. A., Carmichael, S. T., Corbett, D., & Cramer, S. C. (2017). Agreed definitions and a shared vision for new standards in stroke recovery research: The Stroke Recovery and Rehabilitation Roundtable taskforce. *International Journal of Stroke: Official Journal of the International Stroke Society*, 12(5), 444-450.
<https://doi.org/10.1177/1747493017711816>

Bhimani, R., & Anderson, L. (2014). Clinical Understanding of Spasticity: Implications for Practice. *Rehabilitation Research and Practice*, 2014, 1-10.
<https://doi.org/10.1155/2014/279175>

Billinger, S. A., Arena, R., Bernhardt, J., Eng, J. J., Franklin, B. A., Johnson, C. M., MacKay-Lyons, M., Macko, R. F., Mead, G. E., Roth, E. J., Shaughnessy, M., Tang, A., American Heart Association Stroke Council, Council on Cardiovascular and Stroke Nursing, Council on Lifestyle and Cardiometabolic Health, Council on Epidemiology and Prevention, & Council on Clinical Cardiology. (2014). Physical activity and exercise recommendations for stroke survivors: A statement for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke*, 45(8), 2532-2553.
<https://doi.org/10.1161/STR.0000000000000022>

Bland, J. M., & Altman, D. G. (1997). Cronbach's alpha. *BMJ (Clinical Research Ed.)*, 314(7080), 572. <https://doi.org/10.1136/bmj.314.7080.572>

Bollen, K. A. (1989). *Structural Equations with Latent Variables* (1.^a ed.). Wiley.
<https://doi.org/10.1002/9781118619179>

Bombardier, C., & Tugwell, P. (1987). Methodological considerations in functional assessment. *The Journal of Rheumatology. Supplement*, 14 Suppl 15, 6-10.

Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., Gatsonis, C. A., Glasziou, P. P., Irwig, L. M., Moher, D., Rennie, D., de Vet, H. C. W., Lijmer, J. G., & Standards for Reporting of Diagnostic Accuracy. (2003). The STARD statement for reporting studies of diagnostic accuracy: Explanation and elaboration. *Annals of Internal Medicine*, 138(1), W1-12.
<https://doi.org/10.7326/0003-4819-138-1-200301070-00012-w1>

Bou, J. C., & Satorra, A. (2010). A Multigroup Structural Equation Approach: A Demonstration by Testing Variation of Firm Profitability Across EU Samples. *Organizational Research Methods*, 13(4), 738-766.
<https://doi.org/10.1177/1094428109340433>

- Boulet, L. P., Becker, A., Bérubé, D., Beveridge, R., & Ernst, P. (1999). Canadian Asthma Consensus Report, 1999. Canadian Asthma Consensus Group. *CMAJ: Canadian Medical Association Journal = Journal de l'Association Médicale Canadienne*, 161(11 Suppl), S1-61.
- Boyne, P., Welge, J., Kissela, B., & Dunning, K. (2017). Factors Influencing the Efficacy of Aerobic Exercise for Improving Fitness and Walking Capacity After Stroke: A Meta-Analysis With Meta-Regression. *Archives of Physical Medicine and Rehabilitation*, 98(3), 581-595.
<https://doi.org/10.1016/j.apmr.2016.08.484>
- Brea Rivero, M. (2008). Terminología Consensuada en Terapia Ocupacional: Marco para el Entendimiento. *TOG (A coruña)*, 5(1).
<https://www.revistatog.com/num7/pdfs/revision1.pdf>
- Brod, M., Tesler, L. E., & Christensen, T. L. (2009). Qualitative research and content validity: Developing best practices based on science and experience. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, 18(9), 1263-1278.
<https://doi.org/10.1007/s11136-009-9540-9>
- Broeks, J. G., Lankhorst, G. J., Rumping, K., & Prevo, A. J. (1999). The long-term outcome of arm function after stroke: Results of a follow-up study. *Disability and Rehabilitation*, 21(8), 357-364.
<https://doi.org/10.1080/096382899297459>
- Browne, M. W., & Cudeck, R. (1992). Alternative Ways of Assessing Model Fit. *Sociological Methods & Research*, 21(2), 230-258.
<https://doi.org/10.1177/0049124192021002005>
- Brunnstrom, S. (1966). Motor testing procedures in hemiplegia: Based on sequential recovery stages. *Physical Therapy*, 46(4), 357-375.
<https://doi.org/10.1093/ptj/46.4.357>
- Bushnell, C., Bettger, J. P., Cockcroft, K. M., Cramer, S. C., Edelen, M. O., Hanley, D., Katzan, I. L., Mattke, S., Nilsen, D. M., Piquado, T., Skidmore, E. R., Wing, K., & Yenokyan, G. (2015). Chronic Stroke Outcome Measures for Motor Function Intervention Trials: Expert Panel Recommendations. *Circulation: Cardiovascular Quality and Outcomes*, 8(6_suppl_3).
<https://doi.org/10.1161/CIRCOUTCOMES.115.002098>
- Busk, H., Alt Murphy, M., Korsman, R., Skou, S. T., & Wienecke, T. (2022). Cross-cultural translation and adaptation of the Danish version of the Fugl-Meyer assessment for post stroke sensorimotor function. *Disability and Rehabilitation*, 44(17), 4888-4895.
<https://doi.org/10.1080/09638288.2021.1919215>

- Buzzelli, C., Zerboni, C., & Dominguez, S. (2023). Intervención de Terapia Ocupacional luego de un accidente cerebro vascular: Reporte de caso clínico. *Revista de la Facultad de Ciencias Médicas de Córdoba*, 80(2), 153-155. <https://doi.org/10.31053/1853.0605.v80.n2.40202>
- Byrne, B. M. (2008). Testing for multigroup equivalence of a measuring instrument: A walk through the process. *Psicothema*, 20(4), 872-882.
- Cabrera Zamora, J. L. (2014). *Factores de riesgo y enfermedad cerebrovascular*. 15(2), 75-88.
- Cai, W., Mueller, C., Li, Y.-J., Shen, W.-D., & Stewart, R. (2019). Post stroke depression and risk of stroke recurrence and mortality: A systematic review and meta-analysis. *Ageing Research Reviews*, 50, 102-109. <https://doi.org/10.1016/j.arr.2019.01.013>
- Calvert, M., Blazeby, J., Altman, D. G., Revicki, D. A., Moher, D., Brundage, M. D., & CONSORT PRO Group. (2013). Reporting of patient-reported outcomes in randomized trials: The CONSORT PRO extension. *JAMA*, 309(8), 814-822. <https://doi.org/10.1001/jama.2013.879>
- Campo-Arias, A. (2006). Usos del coeficiente de alfa de Cronbach. *Biomédica*, 26(4), 585. <https://doi.org/10.7705/biomedica.v26i4.327>
- Campo-Arias, A., & Oviedo, H. C. (2008). [Psychometric properties of a scale: Internal consistency]. *Revista De Salud Publica (Bogota, Colombia)*, 10(5), 831-839. <https://doi.org/10.1590/s0124-00642008000500015>
- Camoes-Barbosa, A., Sequeira-Medeiros, L., Duarte, N., Morais, J., Mendes, M., & Meneses, C. (2012). Afasia y depresión post-ictus: Una relación predictiva. *Rehabilitación*, 46(1), 36-40. <https://doi.org/10.1016/j.rh.2011.11.001>
- Canning, C. G., Ada, L., Adams, R., & O'Dwyer, N. J. (2004). Loss of strength contributes more to physical disability after stroke than loss of dexterity. *Clinical Rehabilitation*, 18(3), 300-308. <https://doi.org/10.1191/0269215504cr715oa>
- Cano de la Cuerda, R., & Collado Vázquez, S. (2012). *Neurorrehabilitación. Métodos específicos de valoración y tratamiento*. Editorial Médica Panamericana.
- Cano-de-la-Cuerda, R., Molero-Sánchez, A., Carratalá-Tejada, M., Alguacil-Diego, I. M., Molina-Rueda, F., Miangolarra-Page, J. C., & Torricelli, D. (2015). Teorías y modelos de control y aprendizaje motor. Aplicaciones clínicas en neurorrehabilitación. *Neurología*, 30(1), 32-41. <https://doi.org/10.1016/j.nrl.2011.12.010>

Careaga, R. M. (2020). *Taller educativo para cuidadores principales de pacientes que han sufrido un ictus*. Universidad Pontificia Comillas.

Carey, L. M. (1995). Somatosensory Loss after Stroke. *Critical Reviews in Physical and Rehabilitation Medicine*, 7(1), 51-91. <https://doi.org/10.1615/CritRevPhysRehabilMed.v7.i1.40>

Carey, L. M. (2012). Touch and Body Sensations. En L. M. Carey (Ed.), *Stroke Rehabilitation* (pp. 157-172). Oxford University Press. <https://doi.org/10.1093/med/9780199797882.003.0012>

Carey, L. M. (Ed.). (2013). *Stroke rehabilitation: Insights from neuroscience and imaging*. <https://doi.org/10.1093/med/9780199797882.001.0001>

Carey, L. M., Matyas, T. A., & Baum, C. (2018). Effects of Somatosensory Impairment on Participation After Stroke. *The American Journal of Occupational Therapy: Official Publication of the American Occupational Therapy Association*, 72(3), 7203205100p1-7203205100p10. <https://doi.org/10.5014/ajot.2018.025114>

Carretero-Dios, H. & Pérez, C. (2005). Normas para el desarrollo y revision de estudios instrumentales. *International Journal of clinical and health psuchology*, 5(3), 521-551.

Carvajal, A., et al. (2011). ¿Cómo validar un innstrumento de medida de la salud? 34(1), 63-72.

Casado, V. (2011). Atención al paciente neurológico en los Servicios de Urgencias. Revisión de la situación actual en España. *Neurología*, 26(4), 233-238. <https://doi.org/10.1016/j.nrl.2010.07.033>

Cattell, R. B. (1966). The Scree Test For The Number Of Factors. *Multivariate Behavioral Research*, 1(2), 245-276. https://doi.org/10.1207/s15327906mbr0102_10

Cecchi, F., Carrabba, C., Bertolucci, F., Castagnoli, C., Falsini, C., Gnetti, B., Hochleitner, I., Lucidi, G., Martini, M., Mosca, I. E., Pancani, S., Paperini, A., Verdesca, S., Macchi, C., & Alt Murphy, M. (2021). Transcultural translation and validation of Fugl-Meyer assessment to Italian. *Disability and Rehabilitation*, 43(25), 3717-3722. <https://doi.org/10.1080/09638288.2020.1746844>

Centre for Evidence Based Medicine. (2009). *Oxford Centre for Evidence-Based Medicine: Levels of Evidence*. <https://www.cebm.ox.ac.uk/resources/levels-of-evidence/oxford-centre-for-evidence-based-medicine-levels-of-evidence-march-2009>

- Chen, K.-L., Chen, C.-T., Chou, Y.-T., Shih, C.-L., Koh, C.-L., & Hsieh, C.-L. (2014). Is the Long Form of the Fugl-Meyer Motor Scale More Responsive Than the Short Form in Patients With Stroke? *Archives of Physical Medicine and Rehabilitation*, 95(5), 941-949. <https://doi.org/10.1016/j.apmr.2014.01.014>
- Chen, P., Lai, C. K. Y., Chung, R. C. K., & Ng, S. S. M. (2017). The Jacket Test for assessing people with chronic stroke. *Disability and Rehabilitation*, 39(25), 2577-2583. <https://doi.org/10.1080/09638288.2016.1236413>
- Chiarotto, A., Ostelo, R. W., Boers, M., & Terwee, C. B. (2018). A systematic review highlights the need to investigate the content validity of patient-reported outcome measures for physical functioning in patients with low back pain. *Journal of Clinical Epidemiology*, 95, 73-93. <https://doi.org/10.1016/j.jclinepi.2017.11.005>
- Chiuve, S. E., Rexrode, K. M., Spiegelman, D., Logroscino, G., Manson, J. E., & Rimm, E. B. (2008). Primary prevention of stroke by healthy lifestyle. *Circulation*, 118(9), 947-954. <https://doi.org/10.1161/CIRCULATIONAHA.108.781062>
- Choi, B. C. K., & Pak, A. W. P. (2005). A catalog of biases in questionnaires. *Preventing Chronic Disease*, 2(1), A13.
- Chou, R., Baker, W. L., Bañez, L. L., Iyer, S., Myers, E. R., Newberry, S., Pincock, L., Robinson, K. A., Sardenga, L., Sathe, N., Springs, S., & Wilt, T. J. (2018). Agency for Healthcare Research and Quality Evidence-based Practice Center methods provide guidance on prioritization and selection of harms in systematic reviews. *Journal of Clinical Epidemiology*, 98, 98-104. <https://doi.org/10.1016/j.jclinepi.2018.01.007>
- Censori, B., Manara, O., Agostinis, C., Camerlingo, M., Casto, L., Galavotti, B., Partiguian, T., Servalli, M. C., Cesana, B., Belloni, G., & Mamoli, A. (1996). Dementia after first stroke. *Stroke*, 27(7), 1205-1210. <https://doi.org/10.1161/01.str.27.7.1205>
- Cieza, A., Causey, K., Kamenov, K., Hanson, S. W., Chatterji, S., & Vos, T. (2021). Global estimates of the need for rehabilitation based on the Global Burden of Disease study 2019: A systematic analysis for the Global Burden of Disease Study 2019. *Lancet (London, England)*, 396(10267), 2006-2017. [https://doi.org/10.1016/S0140-6736\(20\)32340-0](https://doi.org/10.1016/S0140-6736(20)32340-0)
- Clarke, P., Marshall, V., Black, S. E., & Colantonio, A. (2002). Well-being after stroke in Canadian seniors: Findings from the Canadian Study of Health and Aging. *Stroke*, 33(4), 1016-1021. <https://doi.org/10.1161/01.str.0000013066.24300.f9>

- Cohen, J. (1968). Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin*, 70(4), 213-220. <https://doi.org/10.1037/h0026256>
- Cohen, J. (2013). *Statistical Power Analysis for the Behavioral Sciences* (0 ed.). Routledge. <https://doi.org/10.4324/9780203771587>
- Colomer, C., Llorens, R., Sánchez, C., Ugart, P., Moliner, B., Navarro, M. D., Noé, E., & Ferri, J. (2023). Reliability and validity of the Spanish adaptation of the Functional Independence Measure + Functional Assessment Measure. *European Journal of Physical and Rehabilitation Medicine*, 59(4). <https://doi.org/10.23736/S1973-9087.23.07841-3>
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78(1), 98-104. <https://doi.org/10.1037/0021-9010.78.1.98>
- COSMIN-Improving the selection of outcome measurement instrument. (2011). [COSMIN]. <https://www.cosmin.nl/>
- Cramer, S. C. (2008). Repairing the human brain after stroke: I. Mechanisms of spontaneous recovery. *Annals of Neurology*, 63(3), 272-287. <https://doi.org/10.1002/ana.21393>
- Creek, J. (Ed.). (2009). *Occupational therapy and mental health* (4. ed., repr). Churchill Livingstone.
- Cronbach, L. J. (1951). Coefficient Alpha and the Internal Structure of Tests. *Psychometrika*, 16(3), 297-334. <https://doi.org/10.1007/BF02310555>
- Crosby, R. D., Kolotkin, R. L., & Williams, G. R. (2003). Defining clinically meaningful change in health-related quality of life. *Journal of Clinical Epidemiology*, 56(5), 395-407. [https://doi.org/10.1016/s0895-4356\(03\)00044-1](https://doi.org/10.1016/s0895-4356(03)00044-1)
- Cuesta García, C. (2016). *El procesamiento de la información somatosensorial y la funcionalidad de la mano en pacientes con Daño Cerebral Adquirido* [Universidad Rey Juan Carlos]. <http://hdl.handle.net/10115/14341>
- da Silva, F. C., Gonçalves, E., Arancibia, B. A. V., Bento, G. G., Castro, T. L. da S., Hernandez, S. S. S., & da Silva, R. (2015). [Estimators of internal consistency in health research: The use of the alpha coefficient]. *Revista Peruana De Medicina Experimental Y Salud Publica*, 32(1), 129-138.
- Dandekar, S. P., & Ganvir, S. (2019). Action research arm test as a tool for assessment of motor recovery in patients with stroke: A critical review. *International Journal of Research in Medical Sciences*, 7(8), 3224. <https://doi.org/10.18203/2320-6012.ijrms20193425>

- Dawson, J., Liu, C. Y., Francisco, G. E., Cramer, S. C., Wolf, S. L., Dixit, A., Alexander, J., Ali, R., Brown, B. L., Feng, W., DeMark, L., Hochberg, L. R., Kautz, S. A., Majid, A., O'Dell, M. W., Pierce, D., Prudente, C. N., Redgrave, J., Turner, D. L., ... Kimberley, T. J. (2023). Vagus nerve stimulation paired with rehabilitation for upper limb motor function after ischaemic stroke (VNS-REHAB): A randomised, blinded, pivotal, device trial. *Lancet (London, England)*, 397(10284), 1545-1553. [https://doi.org/10.1016/S0140-6736\(21\)00475-X](https://doi.org/10.1016/S0140-6736(21)00475-X)
- De Arróyabe Castillo, E. L., & Zumalde, E. C. (2013). Daño cerebral adquirido: Percepción del familiar de las secuelas y su malestar psicológico. *Clínica y Salud*, 24(1), 27-35. <https://doi.org/10.5093/cl2013a4>
- de Blas Zamorano, P., Montagut-Martínez, P., Pérez-Cruzado, D., & Merchán-Baeza, J.-A. (2025). Fugl-Meyer Assessment for upper extremity in stroke: A psychometric systematic review. *Journal of Hand Therapy*. <https://doi.org/10.1016/j.jht.2025.04.004>
- de Diego Alonso, C. (2024). *Estilo de vida de las personas supervivientes de un ictus: Actividad física y participación en la vida cotidiana* [Universidad San Jorge]. <https://www.educacion.gob.es/teseo/mostrarRef.do?ref=2580021>
- Deeks, J. J., Bossuyt, P. M., Leeflang, M. M., & Takwoingi, Y. (2023). *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy. Version 2.0 (updated July 2023)*. <https://training.cochrane.org/handbook-diagnostic-test-accuracy/current>.
- de Havenon, A., Heitsch, L., Sunmonu, A., Braun, R., Lohse, K. R., Cole, J. W., Mistry, E., Lindgren, A., Worrall, B. B., & Cramer, S. C. (2022). Accurate Prediction of Persistent Upper Extremity Impairment in Patients With Ischemic Stroke. *Archives of Physical Medicine and Rehabilitation*, 103(5), 964-969. <https://doi.org/10.1016/j.apmr.2021.10.023>
- De Las Heras, B., Rodrigues, L., Cristini, J., Moncion, K., Ploughman, M., Tang, A., Fung, J., & Roig, M. (2024). Measuring Neuroplasticity in Response to Cardiovascular Exercise in People With Stroke: A Critical Perspective. *Neurorehabilitation and Neural Repair*, 38(4), 303-321. <https://doi.org/10.1177/15459683231223513>
- De Leeuw, E. D., Hox, J., & Dillman, D. (Eds.). (2012). *International Handbook of Survey Methodology* (0 ed.). Routledge. <https://doi.org/10.4324/9780203843123>
- Delgado Losada, M. L. (2019). *Fundamentos de psicología: Para ciencias sociales y de la salud* (1ª ed., reimp). Editorial Médica Panamericana.
- Della Vecchia, C., Viprey, M., Haesebaert, J., Termoz, A., Giroudon, C., Dima, A., Rode, G., Préau, M., & Schott, A.-M. (2021). Contextual determinants of

participation after stroke: A systematic review of quantitative and qualitative studies. *Disability and Rehabilitation*, 43(13), 1786-1798. <https://doi.org/10.1080/09638288.2019.1679897>

de Noreña, D., Sánchez-Cubillo, I., García-Molina, A., Tirapu-Ustárriz, J., Bombín-González, I., & Ríos-Lago, M. (2010). [Effectiveness of neuropsychological rehabilitation in acquired brain injury (II): Executive functions, behavioural modification and psychotherapy, and the use of the new technologies]. *Revista De Neurologia*, 51(12), 733-744.

de Rooij, I. J. M., van de Port, I. G. L., & Meijer, J.-W. G. (2016). Effect of Virtual Reality Training on Balance and Gait Ability in Patients With Stroke: Systematic Review and Meta-Analysis. *Physical Therapy*, 96(12), 1905-1918. <https://doi.org/10.2522/ptj.20160054>

de Sousa, D. G., Harvey, L. A., Dorsch, S., & Glinsky, J. V. (2018). Interventions involving repetitive practice improve strength after stroke: A systematic review. *Journal of Physiotherapy*, 64(4), 210-221. <https://doi.org/10.1016/j.jphys.2018.08.004>

Desrosiers, J., Noreau, L., Rochette, A., Bravo, G., & Boutin, C. (2002). Predictors of handicap situations following post-stroke rehabilitation. *Disability and Rehabilitation*, 24(15), 774-785. <https://doi.org/10.1080/09638280210125814>

de Vet, H. C., Terwee, C. B., Ostelo, R. W., Beckerman, H., Knol, D. L., & Bouter, L. M. (2006). Minimal changes in health status questionnaires: Distinction between minimally detectable change and minimally important change. *Health and Quality of Life Outcomes*, 4, 54. <https://doi.org/10.1186/1477-7525-4-54>

de Vet, H. C. W., Adèr, H. J., Terwee, C. B., & Pouwer, F. (2005). Are factor analytical techniques used appropriately in the validation of health status questionnaires? A systematic review on the quality of factor analysis of the SF-36. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, 14(5), 1203-1218; discussion 1219-1221, 1223-1224. <https://doi.org/10.1007/s11136-004-5742-3>

De Vet, H. C. W., Terwee, C. B., Mokkink, L. B., & Knol, D. L. (2011). *Measurement in Medicine: A Practical Guide* (1.^a ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511996214>

de Vet, H. C. W., Mokkink, L. B., Terwee, C. B., Hoekstra, O. S., & Knol, D. L. (2013). Clinicians are right not to like Cohen's κ . *BMJ (Clinical Research Ed.)*, 346, f2125. <https://doi.org/10.1136/bmj.f2125>

DeVon, H. A., Block, M. E., Moyle-Wright, P., Ernst, D. M., Hayden, S. J., Lazzara, D. J., Savoy, S. M., & Kostas-Polston, E. (2007). A psychometric

toolbox for testing validity and reliability. *Journal of Nursing Scholarship: An Official Publication of Sigma Theta Tau International Honor Society of Nursing*, 39(2), 155-164. <https://doi.org/10.1111/j.1547-5069.2007.00161.x>

De Weerd, W. J. G. (1985). Measuring recovery of arm-hand function in stroke patients: A comparison of the Brunnstrom-Fugl-Meyer test and the Action Research Arm test. *Physiotherapy Canada*, 37(2), 65-70. <https://doi.org/10.3138/ptc.37.2.065>

Deyo, R. A., & Centor, R. M. (1986). Assessing the responsiveness of functional scales to clinical change: An analogy to diagnostic test performance. *Journal of Chronic Diseases*, 39(11), 897-906. [https://doi.org/10.1016/0021-9681\(86\)90038-x](https://doi.org/10.1016/0021-9681(86)90038-x)

Díaz-Guzmán, J., Egido, J.-A., Gabriel-Sánchez, R., Barberá-Comes, G., Fuentes-Gimeno, B., Fernández-Pérez, C., & IBERICTUS Study Investigators of the Stroke Project of the Spanish Cerebrovascular Diseases Study Group. (2012). Stroke and transient ischemic attack incidence rate in Spain: The IBERICTUS study. *Cerebrovascular Diseases (Basel, Switzerland)*, 34(4), 272-281. <https://doi.org/10.1159/000342652>

Dickstein, R., Shefi, S., Marcovitz, E., & Villa, Y. (2004). Electromyographic activity of voluntarily activated trunk flexor and extensor muscles in post-stroke hemiparetic subjects. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 115(4), 790-796. <https://doi.org/10.1016/j.clinph.2003.11.018>

Díez Tejedor, E. (2006). *Guía para el diagnóstico y tratamiento del ictus*. Prous Science Sociedad Española de Neurología.

Díez-Tejedor, E., del Brutto, O., Alvarez Sabín, J., Muñoz, M., Abiusi, G., & Sociedad Iberoamericana de Enfermedades Cerebrovasculares. (2001). [Classification of the cerebrovascular diseases. Iberoamerican Cerebrovascular diseases Society]. *Revista De Neurología*, 33(5), 455-464.

Dirette, D. P. & Gutman, S. A. (2020). *Occupational Therapy for Physical Disfunction* (8.ª ed.). Wolters Kluwer.

Dobkin, B. (1995). The economic impact of stroke. *Neurology*, 45(2 Suppl 1), S6-9.

Douiri, A., Grace, J., Sarker, S.-J., Tilling, K., McKevitt, C., Wolfe, C. D., & Rudd, A. G. (2017). Patient-specific prediction of functional recovery after stroke. *International Journal of Stroke: Official Journal of the International Stroke Society*, 12(5), 539-548. <https://doi.org/10.1177/1747493017706241>

- Doyle, S., Bennett, S., Fasoli, S. E., & McKenna, K. T. (2010). Interventions for sensory impairment in the upper limb after stroke. *Cochrane Database of Systematic Reviews*. <https://doi.org/10.1002/14651858.CD006331.pub2>
- Durà Mata, M. J., Molleda Marzo, M., García Almazán, C., Mallol Badellino, J., & Calderon Padilla, V. (2011). Factores pronósticos en el ictus. De la fase aguda a los tres años. *Rehabilitación*, 45(1), 18-23. <https://doi.org/10.1016/j.rh.2010.09.004>
- Ebell, M. H., Siwek, J., Weiss, B. D., Woolf, S. H., Susman, J., Ewigman, B., & Bowman, M. (2004). Strength of recommendation taxonomy (SORT): A patient-centered approach to grading evidence in the medical literature. *The Journal of the American Board of Family Practice*, 17(1), 59-67. <https://doi.org/10.3122/jabfm.17.1.59>
- Elsman, E. B. M., Mokkink, L. B., Terwee, C. B., Beaton, D., Butcher, N. J., Gagnier, J. J., Tricco, A. C., Lee Aiyegbusi, O., Berardi, A., Farmer, J., & Haywood, K. L. (2024). *Guideline for Reporting Systematic Reviews of Outcome Measurement Instruments*. PRISMA-COSMIN OUTCOME Measurement Instruments.
- Embretson, S. E., & Reise, S. P. (2013). *Item Response Theory* (0 ed.). Psychology Press. <https://doi.org/10.4324/9781410605269>
- Endo, A., Suzuki, M., Akagi, A., Chiba, N., Ishizaka, I., Matsunaga, A., & Fukuda, M. (2015). Reliability and validity of the upper-body dressing scale in Japanese patients with vascular dementia with hemiparesis. *Occupational Therapy International*, 22(1), 10-18. <https://doi.org/10.1002/oti.1378>
- Ergeletzis, D., Kevorkian, C. G., & Rintala, D. (2002). Rehabilitation of the older stroke patient: Functional outcome and comparison with younger patients. *American Journal of Physical Medicine & Rehabilitation*, 81(12), 881-889. <https://doi.org/10.1097/00002060-200212000-00001>
- Ernst, P., Fitzgerald, J. M., & Spier, S. (1996). Canadian Asthma Consensus Conference Summary of Recommendations. *Canadian Respiratory Journal*, 3(2), 89-101. <https://doi.org/10.1155/1996/314657>
- Espárrago Llorca, G., Castilla-Guerra, L., Fernández Moreno, M. C., Ruiz Doblado, S., & Jiménez Hernández, M. D. (2015). Post-stroke depression: An update. *Neurología (Barcelona, Spain)*, 30(1), 23-31. <https://doi.org/10.1016/j.nrl.2012.06.008>
- European Medicines Agency. (2005). *Committee for Medical Products for Human Use. Scientific Discussion: Taxotere (Docetaxel)*. Committee for Medical Products for Human Use. <https://www.ema.europa.eu/en/committees/committee-medicinal-products-human-use-chmp>

- Evers, S. M. A. A., Struijs, J. N., Ament, A. J. H. A., van Genugten, M. L. L., Jager, J. H. C., & van den Bos, G. A. M. (2004). International comparison of stroke cost studies. *Stroke*, 35(5), 1209-1215. <https://doi.org/10.1161/01.STR.0000125860.48180.48>
- Ezekiel, L., Collett, J., Mayo, N. E., Pang, L., Field, L., & Dawes, H. (2019). Factors Associated With Participation in Life Situations for Adults With Stroke: A Systematic Review. *Archives of Physical Medicine and Rehabilitation*, 100(5), 945-955. <https://doi.org/10.1016/j.apmr.2018.06.017>
- Fan, X., & Sivo, S. A. (2007). Sensitivity of Fit Indices to Model Misspecification and Model Types. *Multivariate Behavioral Research*, 42(3), 509-529. <https://doi.org/10.1080/00273170701382864>
- Fava, G. A., Tomba, E., & Sonino, N. (2012). Clinimetrics: The science of clinical measurements. *International Journal of Clinical Practice*, 66(1), 11-15. <https://doi.org/10.1111/j.1742-1241.2011.02825.x>
- Fayers, P. M., Hand, D. J., Bjordal, K., & Groenvold, M. (1997). Causal indicators in quality of life research. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, 6(5), 393-406. <https://doi.org/10.1023/a:1018491512095>
- Fearon, P., Langhorne, P., & Early Supported Discharge Trialists. (2012). Services for reducing duration of hospital care for acute stroke patients. *The Cochrane Database of Systematic Reviews*, 9, CD000443. <https://doi.org/10.1002/14651858.CD000443.pub3>
- Feigin, V. L., Brainin, M., Norrving, B., Martins, S., Sacco, R. L., Hacke, W., Fisher, M., Pandian, J., & Lindsay, P. (2022). World Stroke Organization (WSO): Global Stroke Fact Sheet 2022. *International Journal of Stroke: Official Journal of the International Stroke Society*, 17(1), 18-29. <https://doi.org/10.1177/17474930211065917>
- Feigin, V. L., Owolabi, M. O., & World Stroke Organization–Lancet Neurology Commission Stroke Collaboration Group. (2023). Pragmatic solutions to reduce the global burden of stroke: A World Stroke Organization–Lancet Neurology Commission. *The Lancet. Neurology*, 22(12), 1160-1206. [https://doi.org/10.1016/S1474-4422\(23\)00277-6](https://doi.org/10.1016/S1474-4422(23)00277-6)
- Feinstein, A. R. (1982). T. Duckett Jones Memorial Lecture. The Jones criteria and the challenges of clinimetrics. *Circulation*, 66(1), 1-5. <https://doi.org/10.1161/01.cir.66.1.1>
- Fernández-Solana, J., Pardo-Hernández, R., González-Bernal, J. J., Sánchez-González, E., González-Santos, J., Soto-Cámara, R., & Santamaría-Pelaez, M. (2022). Psychometric Properties of the Action Research Arm Test (ARAT) Scale in Post-Stroke Patients-Spanish Population.

International Journal of Environmental Research and Public Health, 19(22), 14918. <https://doi.org/10.3390/ijerph192214918>

Ferraz, M. B., Oliveira, L. M., Araujo, P. M., Atrá, E., & Tugwell, P. (1990). Crosscultural reliability of the physical ability dimension of the health assessment questionnaire. *The Journal of Rheumatology*, 17(6), 813-817.

Ferrer González, B. M. (2016). *Adaptación y validación al español de la escala Fugl-Meyer en el manejo de la rehabilitación de pacientes con ictus* [Universidad de Sevilla]. <https://idus.us.es/items/d3473ef7-451d-4f90-b708-2d49838ab673>

Figueroa, G. A., Montilla, M. A. C., & Melo, R. M. (2012). Método DELPHI: aplicaciones y posibilidades en la gestión prospectiva de la investigación y desarrollo. *Revista Venezolana de Análisis de Coyuntura*, 18(1), 41-52.

Fletcher-Smith, J. C. (2011). *Recovery of dressing ability after stroke* [University of Nottingham]. <https://eprints.nottingham.ac.uk/11913/>

Fletcher-Smith, J., Walker, M., Sunderland, A., Garvey, K., Wan, A., & Turner, H. (2010). An Interrater Reliability Study of the Nottingham Stroke Dressing Assessment. *British Journal of Occupational Therapy*, 73(12), 570-578. <https://doi.org/10.4276/030802210X12918167234127>

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, 7(3), 286-299. <https://doi.org/10.1037/1040-3590.7.3.286>

Foundas, A. L., & Duncan, E. S. (2019). Limb Apraxia: A Disorder of Learned Skilled Movement. *Current Neurology and Neuroscience Reports*, 19(10), 82. <https://doi.org/10.1007/s11910-019-0989-9>

French, B., Thomas, L., Leathley, M., Sutton, C., McAdam, J., Forster, A., Langhorne, P., Price, C., Walker, A., & Watkins, C. (2010). Does repetitive task training improve functional activity after stroke? A Cochrane systematic review and meta-analysis. *Journal of Rehabilitation Medicine*, 42(1), 9-14. <https://doi.org/10.2340/16501977-0473>

Fu, T. S.-T., Wu, C., Lin, K., Hsieh, C., Liu, J., Wang, T., & Ou-yang, P. (2012). Psychometric comparison of the shortened Fugl-Meyer Assessment and the streamlined Wolf Motor Function Test in stroke rehabilitation. *Clinical Rehabilitation*, 26(11), 1043-1047. <https://doi.org/10.1177/0269215511431474>

Fugl-Meyer, A. R., Jääskö, L., Leyman, I., Olsson, S., & Steglind, S. (1975). The post-stroke hemiplegic patient. 1. A method for evaluation of physical performance. *Scandinavian Journal of Rehabilitation Medicine*, 7(1), 13-31.

- Gagnier, J. J., Lai, J., Mokkink, L. B., & Terwee, C. B. (2021). COSMIN reporting guideline for studies on measurement properties of patient-reported outcome measures. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, 30(8), 2197-2218. <https://doi.org/10.1007/s11136-021-02822-4>
- García de Yébenes P., M. A. J., Rodríguez Salvanés, F., & Carmona Ortells, L. (2009). [Validation of questionnaires]. *Reumatología Clínica*, 5(4), 171-177. <https://doi.org/10.1016/j.reuma.2008.09.007>
- García Díez, E. (2004). Fisioterapia de la espasticidad: Técnicas y metodos. *Fisioterapia*, 26(1), 25-35. [https://doi.org/10.1016/S0211-5638\(04\)73080-1](https://doi.org/10.1016/S0211-5638(04)73080-1)
- García Peña, M., Sánchez Cabeza, A., & Miján De Castro, E. (2002). Evaluación funcional y terapia ocupacional en el daño cerebral adquirido. *Rehabilitación*, 36(3), 167-175. [https://doi.org/10.1016/S0048-7120\(02\)73264-7](https://doi.org/10.1016/S0048-7120(02)73264-7)
- García Ríos, N. J., Sánchez Moriones, D. M., & Montoya Hurtado, O. L. (2015). *Estrategias de Intervención de Fisioterapia en Neurorrehabilitación utilizadas en Colombia*. 9(1), 60-66.
- García Valdés, M. & Suarez Marín, M. (2013). El método Delphi para la consulta a expertos en la investigación científica. *Revista Cubana de Salud Pública*, 39(2), 253-267.
- GBD 2016 Lifetime Risk of Stroke Collaborators, Feigin, V. L., Nguyen, G., Cercy, K., Johnson, C. O., Alam, T., Parmar, P. G., Abajobir, A. A., Abate, K. H., Abd-Allah, F., Abejie, A. N., Abyu, G. Y., Ademi, Z., Agarwal, G., Ahmed, M. B., Akinyemi, R. O., Al-Raddadi, R., Aminde, L. N., Amlie-Lefond, C., ... Roth, G. A. (2018). Global, Regional, and Country-Specific Lifetime Risks of Stroke, 1990 and 2016. *The New England Journal of Medicine*, 379(25), 2429-2437. <https://doi.org/10.1056/NEJMoa1804492>
- GBD 2019 Stroke Collaborators. (2021). Global, regional, and national burden of stroke and its risk factors, 1990-2019: A systematic analysis for the Global Burden of Disease Study 2019. *The Lancet. Neurology*, 20(10), 795-820. [https://doi.org/10.1016/S1474-4422\(21\)00252-0](https://doi.org/10.1016/S1474-4422(21)00252-0)
- Gil, R. (2019). *Neuropsicología*. Elsevier España.
- Gil, J. A. P., Moscoso, S. C., & Rodríguez, R. M. (2000). Validez de constructo: El uso de análisis factorial exploratorio-confirmatorio para obtener evidencias de validez. *Psicothema*, Su2, 442-446.
- Glass, T. A., Matchar, D. B., Belyea, M., & Feussner, J. R. (1993). Impact of social support on outcome in first stroke. *Stroke*, 24(1), 64-70. <https://doi.org/10.1161/01.str.24.1.64>

- Gladstone, D. J., Danells, C. J., & Black, S. E. (2002). The fugl-meyer assessment of motor recovery after stroke: A critical review of its measurement properties. *Neurorehabilitation and Neural Repair*, 16(3), 232-240. <https://doi.org/10.1177/154596802401105171>
- Goldstein, L. B., Bushnell, C. D., Adams, R. J., Appel, L. J., Braun, L. T., Chaturvedi, S., Creager, M. A., Culebras, A., Eckel, R. H., Hart, R. G., Hinchey, J. A., Howard, V. J., Jauch, E. C., Levine, S. R., Meschia, J. F., Moore, W. S., Nixon, J. V. I., Pearson, T. A., American Heart Association Stroke Council, ... Council on Peripheral Vascular Disease, and Interdisciplinary Council on Quality of Care and Outcomes Research. (2011). Guidelines for the primary prevention of stroke: A guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke*, 42(2), 517-584. <https://doi.org/10.1161/STR.0b013e3181fcb238>
- Gomez-Benito, J. & Dolores Hidalgo, M. (2005). *La Validez de los Test, Escalas y Cuestionarios*. 12. <https://revistas.udea.edu.co/index.php/ceo/article/view/1750>
- González Bernal, J., González Santos, J., Cuesta Gómez, J. L., & Grau Rubio, C. (2014). *Implications of using dynamic orthoses in the improvement of the quality of life for patients with stroke*. 25(4), 196-199.
- Gracies, J.-M. (2005a). Pathophysiology of spastic paresis. I: Paresis and soft tissue changes. *Muscle & Nerve*, 31(5), 535-551. <https://doi.org/10.1002/mus.20284>
- Gracies, J.-M. (2005b). Pathophysiology of spastic paresis. II: Emergence of muscle overactivity. *Muscle & Nerve*, 31(5), 552-571. <https://doi.org/10.1002/mus.20285>
- GRADE Handbook. (2013). <https://grade.pro.org/handbook>
- Granhölm, A., Alhazzani, W., & Møller, M. H. (2019). Use of the GRADE approach in systematic reviews and guidelines. *British Journal of Anaesthesia*, 123(5), 554-559. <https://doi.org/10.1016/j.bja.2019.08.015>
- Grau G. (1995). *Metodología para la validación de cuestionarios*. 5, 351-359.
- Greenland, P., Alpert, J. S., Beller, G. A., Benjamin, E. J., Budoff, M. J., Fayad, Z. A., Foster, E., Hlatky, M. A., Hodgson, J. M., Kushner, F. G., Lauer, M. S., Shaw, L. J., Smith, S. C., Taylor, A. J., Weintraub, W. S., Wenger, N. K., Jacobs, A. K., Smith, S. C., Anderson, J. L., ... American Heart Association. (2010). 2010 ACCF/AHA guideline for assessment of cardiovascular risk in asymptomatic adults: A report of the American College of Cardiology Foundation/American Heart Association Task Force

on Practice Guidelines. *Journal of the American College of Cardiology*, 56(25), e50-103. <https://doi.org/10.1016/j.jacc.2010.09.001>

Griffiths, C., Armstrong-James, L., White, P., Rumsey, N., Pleat, J., & Harcourt, D. (2015). A systematic review of patient reported outcome measures (PROMs) used in child and adolescent burn research. *Burns: Journal of the International Society for Burn Injuries*, 41(2), 212-224. <https://doi.org/10.1016/j.burns.2014.07.018>

Guillemin, F. (1995). Cross-cultural adaptation and validation of health status measures. *Scandinavian Journal of Rheumatology*, 24(2), 61-63. <https://doi.org/10.3109/03009749509099285>

Gunduz, M. E., Bucak, B., & Keser, Z. (2023). Advances in Stroke Neurorehabilitation. *Journal of Clinical Medicine*, 12(21), 6734. <https://doi.org/10.3390/jcm12216734>

Gurari, N., Drogos, J. M., & Dewald, J. P. A. (2018). Ability of individuals with chronic hemiparetic stroke to locate their forearms during single-arm and between-arms tasks. *PloS One*, 13(10), e0206518. <https://doi.org/10.1371/journal.pone.0206518>

Guyatt, G., Walter, S., & Norman, G. (1987). Measuring change over time: Assessing the usefulness of evaluative instruments. *Journal of Chronic Diseases*, 40(2), 171-178. [https://doi.org/10.1016/0021-9681\(87\)90069-5](https://doi.org/10.1016/0021-9681(87)90069-5)

Guyatt, G. H., Oxman, A. D., Kunz, R., Atkins, D., Brozek, J., Vist, G., Alderson, P., Glasziou, P., Falck-Ytter, Y., & Schünemann, H. J. (2011a). GRADE guidelines: 2. Framing the question and deciding on important outcomes. *Journal of Clinical Epidemiology*, 64(4), 395-400. <https://doi.org/10.1016/j.jclinepi.2010.09.012>

Guyatt, G. H., Oxman, A. D., Kunz, R., Brozek, J., Alonso-Coello, P., Rind, D., Devereaux, P. J., Montori, V. M., Freyschuss, B., Vist, G., Jaeschke, R., Williams, J. W., Murad, M. H., Sinclair, D., Falck-Ytter, Y., Meerpohl, J., Whittington, C., Thorlund, K., Andrews, J., & Schünemann, H. J. (2011b). GRADE guidelines 6. Rating the quality of evidence—Imprecision. *Journal of Clinical Epidemiology*, 64(12), 1283-1293. <https://doi.org/10.1016/j.jclinepi.2011.01.012>

Guyatt, G. H., Oxman, A. D., Kunz, R., Falck-Ytter, Y., Vist, G. E., Liberati, A., Schünemann, H. J., & GRADE Working Group. (2008a). Going from evidence to recommendations. *BMJ (Clinical Research Ed.)*, 336(7652), 1049-1051. <https://doi.org/10.1136/bmj.39493.646875.AE>

Guyatt, G. H., Oxman, A. D., Kunz, R., Vist, G. E., Falck-Ytter, Y., Schünemann, H. J., & GRADE Working Group. (2008b). What is «quality of evidence»

and why is it important to clinicians? *BMJ (Clinical Research Ed.)*, 336(7651), 995-998. <https://doi.org/10.1136/bmj.39490.551019.BE>

Guyatt, G. H., Oxman, A. D., Kunz, R., Woodcock, J., Brozek, J., Helfand, M., Alonso-Coello, P., Falck-Ytter, Y., Jaeschke, R., Vist, G., Akl, E. A., Post, P. N., Norris, S., Meerpohl, J., Shukla, V. K., Nasser, M., Schünemann, H. J., & GRADE Working Group. (2011c). GRADE guidelines: 8. Rating the quality of evidence--indirectness. *Journal of Clinical Epidemiology*, 64(12), 1303-1310. <https://doi.org/10.1016/j.jclinepi.2011.04.014>

Guyatt, G. H., Oxman, A. D., Kunz, R., Woodcock, J., Brozek, J., Helfand, M., Alonso-Coello, P., Glasziou, P., Jaeschke, R., Akl, E. A., Norris, S., Vist, G., Dahm, P., Shukla, V. K., Higgins, J., Falck-Ytter, Y., Schünemann, H. J., & GRADE Working Group. (2011d). GRADE guidelines: 7. Rating the quality of evidence--inconsistency. *Journal of Clinical Epidemiology*, 64(12), 1294-1302. <https://doi.org/10.1016/j.jclinepi.2011.03.017>

Guyatt, G. H., Oxman, A. D., Montori, V., Vist, G., Kunz, R., Brozek, J., Alonso-Coello, P., Djulbegovic, B., Atkins, D., Falck-Ytter, Y., Williams, J. W., Meerpohl, J., Norris, S. L., Akl, E. A., & Schünemann, H. J. (2011e). GRADE guidelines: 5. Rating the quality of evidence--publication bias. *Journal of Clinical Epidemiology*, 64(12), 1277-1282. <https://doi.org/10.1016/j.jclinepi.2011.01.011>

Guyatt, G. H., Oxman, A. D., Schünemann, H. J., Tugwell, P., & Knottnerus, A. (2011f). GRADE guidelines: A new series of articles in the Journal of Clinical Epidemiology. *Journal of Clinical Epidemiology*, 64(4), 380-382. <https://doi.org/10.1016/j.jclinepi.2010.09.011>

Guyatt, G. H., Oxman, A. D., Sultan, S., Glasziou, P., Akl, E. A., Alonso-Coello, P., Atkins, D., Kunz, R., Brozek, J., Montori, V., Jaeschke, R., Rind, D., Dahm, P., Meerpohl, J., Vist, G., Berliner, E., Norris, S., Falck-Ytter, Y., Murad, M. H., ... GRADE Working Group. (2011g). GRADE guidelines: 9. Rating up the quality of evidence. *Journal of Clinical Epidemiology*, 64(12), 1311-1316. <https://doi.org/10.1016/j.jclinepi.2011.06.004>

Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., Schünemann, H. J., & GRADE Working Group. (2008c). GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. *BMJ (Clinical Research Ed.)*, 336(7650), 924-926. <https://doi.org/10.1136/bmj.39489.470347.AD>

Guyatt, G. H., Oxman, A. D., Vist, G., Kunz, R., Brozek, J., Alonso-Coello, P., Montori, V., Akl, E. A., Djulbegovic, B., Falck-Ytter, Y., Norris, S. L., Williams, J. W., Atkins, D., Meerpohl, J., & Schünemann, H. J. (2011h). GRADE guidelines: 4. Rating the quality of evidence--study limitations (risk of bias). *Journal of Clinical Epidemiology*, 64(4), 407-415. <https://doi.org/10.1016/j.jclinepi.2010.07.017>

Harrison, D. E. (s. f.). *Practicing Chiropractors' Committee on Radiology Protocols (PCCRP)*. <http://www.pccrp.org/>

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (Second edition). Guilford Press.

Haywood, K. L., Garratt, A. M., & Fitzpatrick, R. (2006). Quality of life in older people: A structured review of self-assessed health instruments. *Expert Review of Pharmacoeconomics & Outcomes Research*, 6(2), 181-194. <https://doi.org/10.1586/14737167.6.2.181>

He, Q., Wang, W., Zhang, Y., Xiong, Y., Tao, C., Ma, L., Ma, J., You, C., & Wang, C. (2024). Global, Regional, and National Burden of Stroke, 1990–2021: A Systematic Analysis for Global Burden of Disease 2021. *Stroke*, 55(12), 2815-2824. <https://doi.org/10.1161/strokeaha.124.048033>

Heinsius, T., Bogousslavsky, J., & Van Melle, G. (1998). Large infarcts in the middle cerebral artery territory. Etiology and outcome patterns. *Neurology*, 50(2), 341-350. <https://doi.org/10.1212/wnl.50.2.341>

Heller, A., Wade, D. T., Wood, V. A., Sunderland, A., Hower, R. L., & Ward, E. (1987). Arm function after stroke: Measurement and recovery over the first three months. *Journal of Neurology, Neurosurgery, and Psychiatry*, 50(6), 714-719. <https://doi.org/10.1136/jnnp.50.6.714>

Henson, R. K. (2001). Understanding Internal Consistency Reliability Estimates: A Conceptual Primer on Coefficient Alpha. *Measurement and Evaluation in Counseling and Development*, 34(3), 177-189. <https://doi.org/10.1080/07481756.2002.12069034>

Herrera, M., Lama, J., & De-la-Cruz, J. (2020). Factores de riesgo de depresión post ictus isquémico. *Revista Médica Herediana*, 31(3), 181-189. <https://doi.org/10.20453/rmh.v31i3.3808>

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (2024). *Cochrane Handbook for Systematic Reviews of Interventions version 6.5 (updated August 2024)*. www.training.cochrane.org/handbook. To cite the print edition of the Handbook, please use:

Hilari, K. (2011). The impact of stroke: Are people with aphasia different to those without? *Disability and Rehabilitation*, 33(3), 211-218. <https://doi.org/10.3109/09638288.2010.508829>

Hochleitner, I., Pellicciari, L., Castagnoli, C., Paperini, A., Politi, A. M., Campagnini, S., Pancani, S., Basagni, B., Gerli, F., Carrozza, M. C., Macchi, C., Alt Murphy, M., & Cecchi, F. (2023). Intra- and inter-rater

reliability of the Italian Fugl-Meyer assessment of upper and lower extremity. *Disability and Rehabilitation*, 45(18), 2989-2999. <https://doi.org/10.1080/09638288.2022.2114553>

Hofman, J. (2008). Clinimetría y la actividad médica. *Revista Argentina de Reumatología*, 2.

Howard, G. (2021). Rural-urban differences in stroke risk. *Preventive Medicine*, 152(Pt 2), 106661. <https://doi.org/10.1016/j.ypmed.2021.106661>

Howick, J. (2011). *The Philosophy of Evidence-Based Medicine* (1.^a ed.). Wiley. <https://doi.org/10.1002/9781444342673>

Hsieh, Y., Wu, C., Lin, K., Chang, Y., Chen, C., & Liu, J. (2009). Responsiveness and Validity of Three Outcome Measures of Motor Function After Stroke Rehabilitation. *Stroke*, 40(4), 1386-1391. <https://doi.org/10.1161/STROKEAHA.108.530584>

Hsieh, Y.-W., Hsueh, I.-P., Chou, Y.-T., Sheu, C.-F., Hsieh, C.-L., & Kwakkel, G. (2007). Development and Validation of a Short Form of the Fugl-Meyer Motor Scale in Patients With Stroke. *Stroke*, 38(11), 3052-3054. <https://doi.org/10.1161/STROKEAHA.107.490730>

Hsueh, I.-P., Hsu, M.-J., Sheu, C.-F., Lee, S., Hsieh, C.-L., & Lin, J.-H. (2008). Psychometric Comparisons of 2 Versions of the Fugl-Meyer Motor Scale and 2 Versions of the Stroke Rehabilitation Assessment of Movement. *Neurorehabilitation and Neural Repair*, 22(6), 737-744. <https://doi.org/10.1177/1545968308315999>

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55. <https://doi.org/10.1080/10705519909540118>

Hu, L. T., Bentler, P. M., & Kano, Y. (1992). Can test statistics in covariance structure analysis be trusted? *Psychological Bulletin*, 112(2), 351-362. <https://doi.org/10.1037/0033-2909.112.2.351>

Huertas Hoyas, E., Pedrero Pérez, E. J., Águila Maturana, A. M., García López-Alberca, S., & González Alted, C. (2015). Predictores de funcionalidad en el daño cerebral adquirido. *Neurología*, 30(6), 339-346. <https://doi.org/10.1016/j.nrl.2014.01.001>

Ikram, M., Rehman, S. S. U., Sunnerhagen, K. S., & Alt Murphy, M. (2022). Urdu translation and cross-cultural validation of the Fugl-Meyer assessment in people with stroke. *Disability and Rehabilitation*, 44(25), 8048-8053. <https://doi.org/10.1080/09638288.2021.2003449>

Ingram, L. A., Butler, A. A., Lord, S. R., & Gandevia, S. C. (2023). Use of a physiological profile to document upper limb motor impairment in ageing and in neurological conditions. *The Journal of Physiology*, 601(12), 2251-2262. <https://doi.org/10.1113/JP283703>

Institute of Medicine (US) Committee on Standards for Systematic Reviews of Comparative Effectiveness Research. (2011). *Finding What Works in Health Care: Standards for Systematic Reviews* (J. Eden, L. Levit, A. Berg, & S. Morton, Eds.). National Academies Press (US). <http://www.ncbi.nlm.nih.gov/books/NBK209518/>

Ioannidis, J. P. A., Greenland, S., Hlatky, M. A., Khoury, M. J., Macleod, M. R., Moher, D., Schulz, K. F., & Tibshirani, R. (2014). Increasing value and reducing waste in research design, conduct, and analysis. *Lancet (London, England)*, 383(9912), 166-175. [https://doi.org/10.1016/S0140-6736\(13\)62227-8](https://doi.org/10.1016/S0140-6736(13)62227-8)

Ioannidis, J. P. A. (2022). Correction: Why Most Published Research Findings Are False. *PLOS Medicine*, 19(8), e1004085. <https://doi.org/10.1371/journal.pmed.1004085>

Jackson, D. L., Gillaspy, J. A., & Purc-Stephenson, R. (2009). Reporting practices in confirmatory factor analysis: An overview and some recommendations. *Psychological Methods*, 14(1), 6-23. <https://doi.org/10.1037/a0014694>

Jadad, A. R., Moore, R. A., Carroll, D., Jenkinson, C., Reynolds, D. J., Gavaghan, D. J., & McQuay, H. J. (1996). Assessing the quality of reports of randomized clinical trials: Is blinding necessary? *Controlled Clinical Trials*, 17(1), 1-12. [https://doi.org/10.1016/0197-2456\(95\)00134-4](https://doi.org/10.1016/0197-2456(95)00134-4)

Jhong, M., Tang, N.-Y., Liu, C.-H., Huang, W.-H., Hsu, Y.-T., Liu, Y.-L., Li, T.-C., & Hsieh, C.-L. (2013). Relationship Between Chinese Medicine Pattern Types, Clinical Severity, and Prognosis in Patients With Acute Cerebral Infarct. *EXPLORE*, 9(4), 226-231. <https://doi.org/10.1016/j.explore.2013.04.002>

Jørgensen, H. S., Nakayama, H., Reith, J., Raaschou, H. O., & Olsen, T. S. (1996). Acute stroke with atrial fibrillation. The Copenhagen Stroke Study. *Stroke*, 27(10), 1765-1769. <https://doi.org/10.1161/01.str.27.10.1765>

Jovell, A. J., & Navarro-Rubio, M. D. (1995). [Evaluation of scientific evidence]. *Medicina Clínica*, 105(19), 740-743.

Kaiser, H. F. (1960). The Application of Electronic Computers to Factor Analysis. *Educational and Psychological Measurement*, 20(1), 141-151. <https://doi.org/10.1177/001316446002000116>

- Katrak, P., Bowring, G., Conroy, P., Chilvers, M., Poulos, R., & McNeil, D. (1998). Predicting upper limb recovery after stroke: The place of early shoulder and hand movement. *Archives of Physical Medicine and Rehabilitation*, 79(7), 758-761. [https://doi.org/10.1016/s0003-9993\(98\)90352-5](https://doi.org/10.1016/s0003-9993(98)90352-5)
- Kennedy, C., Bernhardt, J., Churilov, L., Collier, J. M., Ellery, F., Rethnam, V., Carvalho, L. B., Donnan, G. A., & Hayward, K. S. (2021). Factors associated with time to independent walking recovery post-stroke. *Journal of Neurology, Neurosurgery, and Psychiatry*, 92(7), 702-708. <https://doi.org/10.1136/jnnp-2020-325125>
- Khan, A. I., Abuzainah, B., Gutlapalli, S. D., Chaudhuri, D., Khan, K. I., Al Shouli, R., Allakky, A., Ferguson, A. A., & Hamid, P. (2023). Effect of Major Depressive Disorder on Stroke Risk and Mortality: A Systematic Review. *Cureus*. <https://doi.org/10.7759/cureus.40475>
- Kielhofner, G. (2006). *Fundamentos conceptuales de la terapia ocupacional* (3a. ed). Médica Panamericana.
- Kilbride, C., Kneafsey, R., & Kean, V. (2019). Management of Physical Impairments Post-Stroke. En J. Williams, L. Perry, & C. Watkins (Eds.), *Stroke Nursing* (1.^a ed., pp. 177-202). Wiley. <https://doi.org/10.1002/9781119581161.ch8>
- Kim, H., Her, J., Ko, J., Park, D., Woo, J.-H., You, Y., & Choi, Y. (2012). Reliability, Concurrent Validity, and Responsiveness of the Fugl-Meyer Assessment (FMA) for Hemiplegic Patients. *Journal of Physical Therapy Science*, 24(9), 893-899. <https://doi.org/10.1589/jpts.24.893>
- Kim, T., Hwang, S. H., Lee, W. J., Hwang, J. W., Cho, I., Kim, E.-H., Lee, J. A., Choi, Y., Park, J. H., & Shin, J.-H. (2021). The Korean Version of the Fugl-Meyer Assessment: Reliability and Validity Evaluation. *Annals of Rehabilitation Medicine*, 45(2), 83-98. <https://doi.org/10.5535/arm.20225>
- Kim, W.-S., Cho, S., Baek, D., Bang, H., & Paik, N.-J. (2016). Upper Extremity Functional Evaluation by Fugl-Meyer Assessment Scoring Using Depth-Sensing Camera in Hemiplegic Stroke Patients. *PLOS ONE*, 11(7), e0158640. <https://doi.org/10.1371/journal.pone.0158640>
- Kolmos, M., Christoffersen, L., & Kruuse, C. (2021). Recurrent Ischemic Stroke—A Systematic Review and Meta-Analysis. *Journal of Stroke and Cerebrovascular Diseases: The Official Journal of National Stroke Association*, 30(8), 105935. <https://doi.org/10.1016/j.jstrokecerebrovasdis.2021.105935>
- Krus, D. & Helmstadter, G. (1993). The Problem of Negative Reliabilities. *Educational and Psychological Measurement*, 53(3), 643-650.

- Kwakkel, G., Wagenaar, R. C., Kollen, B. J., & Lankhorst, G. J. (1996). Predicting disability in stroke—A critical review of the literature. *Age and Ageing*, 25(6), 479-489. <https://doi.org/10.1093/ageing/25.6.479>
- Kwakkel, G., Stinear, C., Essers, B., Munoz-Novoa, M., Branscheidt, M., Cabanas-Valdés, R., Lakičević, S., Lampropoulou, S., Luft, A. R., Marque, P., Moore, S. A., Solomon, J. M., Swinnen, E., Turolla, A., Alt Murphy, M., & Verheyden, G. (2023). Motor rehabilitation after stroke: European Stroke Organisation (ESO) consensus-based definition and guiding framework. *European Stroke Journal*, 8(4), 880-894. <https://doi.org/10.1177/23969873231191304>
- Lamprea, M. J. A. & Gómez, R. C. (2007). Validez en la evaluación de escalas. *Revista Colombiana de Psiquiatría*, 36(2), 340-348.
- Lance, J. W., & Gail, P. D. (1965). Spread of phasic muscle reflexes in normal and spastic subjects. *Journal of Neurology, Neurosurgery & Psychiatry*, 28(4), 328-334. <https://doi.org/10.1136/jnnp.28.4.328>
- Lang, C. E., Edwards, D. F., Birkenmeier, R. L., & Dromerick, A. W. (2008). Estimating minimal clinically important differences of upper-extremity measures early after stroke. *Archives of Physical Medicine and Rehabilitation*, 89(9), 1693-1700. <https://doi.org/10.1016/j.apmr.2008.02.022>
- Langer, G., Meerpohl, J. J., Perleth, M., Gartlehner, G., Kaminski-Hartenthaler, A., & Schünemann, H. (2012). [GRADE guidelines: 1. Introduction - GRADE evidence profiles and summary of findings tables]. *Zeitschrift Fur Evidenz, Fortbildung Und Qualitat Im Gesundheitswesen*, 106(5), 357-368. <https://doi.org/10.1016/j.zefq.2012.05.017>
- Langerak, A. J., McCambridge, A. B., Stubbs, P. W., Fabricius, J., Rogers, K., Quel de Oliveira, C., Nielsen, J. F., & Verhagen, A. P. (2021). Externally validated model predicting gait independence after stroke showed fair performance and improved after updating. *Journal of Clinical Epidemiology*, 137, 73-82. <https://doi.org/10.1016/j.jclinepi.2021.03.022>
- Langhammer, B., Sunnerhagen, K. S., Lundgren-Nilsson, Å., Sällström, S., Becker, F., & Stanghelle, J. K. (2017). Factors enhancing activities of daily living after stroke in specialized rehabilitation: An observational multicenter study within the Sunnaas International Network. *European Journal of Physical and Rehabilitation Medicine*, 53(5). <https://doi.org/10.23736/S1973-9087.17.04489-6>
- Langhorne, P., Coupar, F., & Pollock, A. (2009). Motor recovery after stroke: A systematic review. *The Lancet. Neurology*, 8(8), 741-754. [https://doi.org/10.1016/S1474-4422\(09\)70150-4](https://doi.org/10.1016/S1474-4422(09)70150-4)

- Langhorne, P., Taylor, G., Murray, G., Dennis, M., Anderson, C., Bautz-Holter, E., Dey, P., Indredavik, B., Mayo, N., Power, M., Rodgers, H., Ronning, O. M., Rudd, A., Suwanwela, N., Widen-Holmqvist, L., & Wolfe, C. (2005). Early supported discharge services for stroke patients: A meta-analysis of individual patients' data. *Lancet (London, England)*, 365(9458), 501-506. [https://doi.org/10.1016/S0140-6736\(05\)17868-4](https://doi.org/10.1016/S0140-6736(05)17868-4)
- Lee, C. D., Folsom, A. R., & Blair, S. N. (2003). Physical activity and stroke risk: A meta-analysis. *Stroke*, 34(10), 2475-2481. <https://doi.org/10.1161/01.STR.0000091843.02517.9D>
- Lee, Y., Nicholas, M. L., & Connor, L. T. (2023). Identifying emotional contributors to participation post-stroke. *Topics in Stroke Rehabilitation*, 30(2), 180-192. <https://doi.org/10.1080/10749357.2021.2008597>
- Leidy, N. K., & Vernon, M. (2008). Perspectives on patient-reported outcomes: Content validity and qualitative research in a changing clinical trial environment. *PharmacoEconomics*, 26(5), 363-370. <https://doi.org/10.2165/00019053-200826050-00002>
- Leno Díaz, C., Holguín Mohedas, M., Hidalgo Jiménez, N., Rodríguez-Ramos, M., & Lavado García, J. M. (2016). Calidad de vida relacionada con la salud en personas supervivientes a un ictus a largo plazo. *Revista Científica de la Sociedad Española de Enfermería Neurológica*, 44, 9-15. <https://doi.org/10.1016/j.sedene.2015.12.002>
- Li, S., Ghuman, J., Gonzalez-Buonomo, J., Huang, X., Malik, A., Yozbatiran, N., Francisco, G. E., Wu, H., & Frontera, W. R. (2023). Does Spasticity Correlate With Motor Impairment in the Upper and Lower Limbs in Ambulatory Chronic Stroke Survivors? *American Journal of Physical Medicine & Rehabilitation*, 102(10), 907-912. <https://doi.org/10.1097/PHM.0000000000002247>
- Li, Y., Zhang, X., Sang, H., Niu, X., Liu, T., Liu, W., & Li, J. (2019). Urban-rural differences in risk factors for ischemic stroke in northern China. *Medicine*, 98(21), e15782. <https://doi.org/10.1097/MD.00000000000015782>
- Lieberman, D., & Lieberman, D. (2005). Rehabilitation following stroke in patients aged 85 and above. *Journal of Rehabilitation Research and Development*, 42(1), 47-53.
- Lin, G.-H., Huang, C.-Y., Lee, S.-C., Chen, K.-L., Lien, J.-J. J., Chen, M.-H., Huang, Y.-H., & Hsieh, C.-L. (2021). A 10-item Fugl-Meyer Motor Scale Based on Machine Learning. *Physical Therapy*, 101(4), pzab036. <https://doi.org/10.1093/ptj/pzab036>
- Lin, J.-H., Hsu, M.-J., Sheu, C.-F., Wu, T.-S., Lin, R.-T., Chen, C.-H., & Hsieh, C.-L. (2009). Psychometric Comparisons of 4 Measures for Assessing Upper-

Extremity Function in People With Stroke. *Physical Therapy*, 89(8), 840-850. <https://doi.org/10.2522/ptj.20080285>

Lin, J.-H., Hsueh, I.-P., Sheu, C.-F., & Hsieh, C.-L. (2004). Psychometric properties of the sensory scale of the Fugl-Meyer Assessment in stroke patients. *Clinical Rehabilitation*, 18(4), 391-397. <https://doi.org/10.1191/0269215504cr737oa>

Lin, H. J., Wolf, P. A., Kelly-Hayes, M., Beiser, A. S., Kase, C. S., Benjamin, E. J., & D'Agostino, R. B. (1996). Stroke severity in atrial fibrillation. The Framingham Study. *Stroke*, 27(10), 1760-1764. <https://doi.org/10.1161/01.str.27.10.1760>

Lira C., E. U. M. T. (2015). IMPACTO DE LA HIPERTENSIÓN ARTERIAL COMO FACTOR DE RIESGO CARDIOVASCULAR. *Revista Médica Clínica Las Condes*, 26(2), 156-163. <https://doi.org/10.1016/j.rmcl.2015.04.004>

Lohr, K. N. (2002). Assessing health status and quality-of-life instruments: Attributes and review criteria. *Quality of Life Research*, 11(3), 193-205. <https://doi.org/10.1023/A:1015291021312>

Lugtmeijer, S., Lammers, N. A., de Haan, E. H. F., de Leeuw, F.-E., & Kessels, R. P. C. (2021). Post-Stroke Working Memory Dysfunction: A Meta-Analysis and Systematic Review. *Neuropsychology Review*, 31(1), 202-219. <https://doi.org/10.1007/s11065-020-09462-4>

Lui, F., Hui, C., Khan Suheb, M. Z., & Patti, L. (2025). Ischemic Stroke. En *StatPearls*. StatPearls Publishing. <http://www.ncbi.nlm.nih.gov/books/NBK499997/>

Lundquist, C. B., & Maribo, T. (2017). The Fugl-Meyer assessment of the upper extremity: Reliability, responsiveness and validity of the Danish version. *Disability and Rehabilitation*, 39(9), 934-939. <https://doi.org/10.3109/09638288.2016.1163422>

Macciocchi, S. N., Diamond, P. T., Alves, W. M., & Mertz, T. (1998). Ischemic stroke: Relation of age, lesion location, and initial neurologic deficit to functional outcome. *Archives of Physical Medicine and Rehabilitation*, 79(10), 1255-1257. [https://doi.org/10.1016/s0003-9993\(98\)90271-4](https://doi.org/10.1016/s0003-9993(98)90271-4)

Magasi, S., Ryan, G., Revicki, D., Lenderking, W., Hays, R. D., Brod, M., Snyder, C., Boers, M., & Cella, D. (2012). Content validity of patient-reported outcome measures: Perspectives from a PROMIS meeting. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, 21(5), 739-746. <https://doi.org/10.1007/s11136-011-9990-8>

- Malouin, F., Pichard, L., Bonneau, C., Durand, A., & Corriveau, D. (1994). Evaluating motor recovery early after stroke: Comparison of the Fugl-Meyer Assessment and the Motor Assessment Scale. *Archives of Physical Medicine and Rehabilitation*, 75(11), 1206-1212. [https://doi.org/10.1016/0003-9993\(94\)90006-X](https://doi.org/10.1016/0003-9993(94)90006-X)
- Maneesriwongul, W., & Dixon, J. K. (2004). Instrument translation process: A methods review. *Journal of Advanced Nursing*, 48(2), 175-186. <https://doi.org/10.1111/j.1365-2648.2004.03185.x>
- Manterola, C., Asenjo-Lobos, C., & Otzen, T. (2014). Jerarquización de la evidencia: Niveles de evidencia y grados de recomendación de uso actual. *Revista Chilena de Infectología*, 31(6), 705-718. <https://doi.org/10.4067/S0716-10182014000600011>
- Marinus, J., Ramaker, C., van Hilten, J. J., & Stiggelbout, A. M. (2002). Health related quality of life in Parkinson's disease: A systematic review of disease specific instruments. *Journal of Neurology, Neurosurgery, and Psychiatry*, 72(2), 241-248. <https://doi.org/10.1136/jnnp.72.2.241>
- Markus, H. S. (2022). Reducing disability after stroke. *International Journal of Stroke*, 17(3), 249-250. <https://doi.org/10.1177/17474930221080904>
- Marshall, M., Lockwood, A., Bradley, C., Adams, C., Joy, C., & Fenton, M. (2000). Unpublished rating scales: A major source of bias in randomised controlled trials of treatments for schizophrenia. *The British Journal of Psychiatry: The Journal of Mental Science*, 176, 249-252. <https://doi.org/10.1192/bjp.176.3.249>
- Martín Arribas, M. C. (2004). *Diseño y validación de cuestionarios*. 5(17), 23-29.
- Martin-Saez, M. M., & James, N. (2021). The experience of occupational identity disruption post stroke: A systematic review and meta-ethnography. *Disability and Rehabilitation*, 43(8), 1044-1055. <https://doi.org/10.1080/09638288.2019.1645889>
- Martinell, M. (2013). *Intervención Logopédica en la Afasia*. 83, 4-8.
- Mathiowetz, V. (1993). Role of physical performance component evaluations in occupational therapy functional assessment. *The American Journal of Occupational Therapy: Official Publication of the American Occupational Therapy Association*, 47(3), 225-230. <https://doi.org/10.5014/ajot.47.3.225>
- McDowell, F. H. (1994). Neurorehabilitation. *The Western Journal of Medicine*, 161(3), 323-327.

- McDowell, I., & Jenkinson, C. (1996). Development standards for health measures. *Journal of Health Services Research & Policy*, 1(4), 238-246. <https://doi.org/10.1177/135581969600100410>
- McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, 1(1), 30-46. <https://doi.org/10.1037/1082-989X.1.1.30>
- Mendis, S. (2013). Stroke disability and rehabilitation of stroke: World Health Organization perspective. *International Journal of Stroke: Official Journal of the International Stroke Society*, 8(1), 3-4. <https://doi.org/10.1111/j.1747-4949.2012.00969.x>
- Meneses, J., Barrios, M., Cosculluela, A., Lozano, L.M., Turbany, J., & et al. (2013). *Aproximación hitórica y conceptos básicos de la psicometría* (Psicometría, pp. 25-73). UOC.
- Merlin, T., Weston, A., & Tooher, R. (2009). Extending an evidence hierarchy to include topics other than treatment: Revising the Australian «levels of evidence». *BMC Medical Research Methodology*, 9, 34. <https://doi.org/10.1186/1471-2288-9-34>
- Messick, S. (1975). The standard problem: Meaning and values in measurement and evaluation. *American Psychologist*, 30(10), 955-966. <https://doi.org/10.1037/0003-066X.30.10.955>
- Messick, S. (1980). Test validity and the ethics of assessment. *American Psychologist*, 35(11), 1012-1027. <https://doi.org/10.1037/0003-066X.35.11.1012>
- Michaelson, S. M., Rocha, A. S., Knabben, R. J., Rodrigues, L. P., & Fernandes, C. G. C. (2011). Translation, adaptation and inter-rater reliability of the administration manual for the Fugl-Meyer assessment. *Revista Brasileira De Fisioterapia (Sao Carlos (Sao Paulo, Brazil))*, 15(1), 80-88.
- Miki, E., Yamane, S., Yamaoka, M., Fujii, H., Ueno, H., Kawahara, T., Tanaka, K., Tamashiro, H., Inoue, E., Okamoto, T., & Kuriyama, M. (2016). Validity and reliability of the Japanese version of the FIM + FAM in patients with cerebrovascular accident. *Scandinavian Journal of Occupational Therapy*, 23(5), 398-404. <https://doi.org/10.3109/11038128.2015.1095236>
- Ministerio de Sanidad y Política Social. (s. f.). *Estrategiae Ictus de Sistema Nacionalde Salud. Actualización 2024*. Ministerio de Sanidad y Política Social. <https://www.sanidad.gob.es/areas/calidadAsistencial/estrategias/ictus/docs/Estrategia en Ictus del SNS. Actualizacion 2024 accesible.pdf>

- Miyai, I., Blau, A. D., Reding, M. J., & Volpe, B. T. (1997). Patients with stroke confined to basal ganglia have diminished response to rehabilitation efforts. *Neurology*, 48(1), 95-101. <https://doi.org/10.1212/wnl.48.1.95>
- Mlinac, M. E., & Feng, M. C. (2016). Assessment of Activities of Daily Living, Self-Care, and Independence. *Archives of Clinical Neuropsychology*, 31(6), 506-516. <https://doi.org/10.1093/arclin/acw049>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2010). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *International Journal of Surgery (London, England)*, 8(5), 336-341. <https://doi.org/10.1016/j.ijsu.2010.02.007>
- Mokkink, L. B., de Vet, H. C. W., Prinsen, C. a. C., Patrick, D. L., Alonso, J., Bouter, L. M., & Terwee, C. B. (2018). COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, 27(5), 1171-1179. <https://doi.org/10.1007/s11136-017-1765-4>
- Mokkink, L. B., Prinsen, C. A. C., Bouter, L. M., Vet, H. C. W. de, & Terwee, C. B. (2016). The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) and how to select an outcome measurement instrument. *Brazilian Journal of Physical Therapy*, 20(2), 105-113. <https://doi.org/10.1590/bjpt-rbf.2014.0143>
- Mokkink, L. B., Prinsen, C. A., Patrick, D. L., Alonso, J., Bouter, L. M., De Vet, H. C., & et al. (2019). *COSMIN Study Design checklist for Patient-reported outcome measurement instruments*. https://www.cosmin.nl/wp-content/uploads/COSMIN-study-designing-checklist_final.pdf
- Mokkink, L. B., Prinsen, C., Patrick, D. L., Alonso, J., Bouter, L. M., De Vet, H. C., & et al. (2018). *COSMIN methodology for systematic reviews of patient-reported outcome measures (PROMs). User manual*. https://cosmin.nl/wp-content/uploads/COSMIN-syst-review-for-PROMs-manual_version-1_feb-2018.pdf
- Mokkink, L. B., Terwee, C. B., Knol, D. L., Stratford, P. W., Alonso, J., Patrick, D. L., Bouter, L. M., & de Vet, H. C. (2010a). The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content. *BMC Medical Research Methodology*, 10, 22. <https://doi.org/10.1186/1471-2288-10-22>
- Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. W. (2010b). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Quality of Life Research: An International Journal of Quality of Life Aspects*

of Treatment, Care and Rehabilitation, 19(4), 539-549.
<https://doi.org/10.1007/s11136-010-9606-8>

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. W. (2010b). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology*, 63(7), 737-745.
<https://doi.org/10.1016/j.jclinepi.2010.02.006>

Mokkink, L. B., Terwee, C. B., Gibbons, E., Stratford, P. W., Alonso, J., Patrick, D. L., Knol, D. L., Bouter, L. M., & de Vet, H. C. W. (2010c). Inter-rater agreement and reliability of the COSMIN (COnsensus-based Standards for the selection of health status Measurement Instruments) checklist. *BMC Medical Research Methodology*, 10, 82. <https://doi.org/10.1186/1471-2288-10-82>

Mokkink, L. B., Boers, M., van der Vleuten, C. P. M., Bouter, L. M., Alonso, J., Patrick, D. L., de Vet, H. C. W., & Terwee, C. B. (2020). COSMIN Risk of Bias tool to assess the quality of studies on reliability or measurement error of outcome measurement instruments: A Delphi study. *BMC Medical Research Methodology*, 20(1), 293. <https://doi.org/10.1186/s12874-020-01179-5>

Montagut Martínez, P. (2022). *Validación de herramientas de evaluación en terapia ocupacional y enfermería a través de la metodología COSMIN [UCAM]*. <http://hdl.handle.net/10952/5344>

Moreno Bacca, N. (2017). Práctica neurorrehabilitación paciente adulto, bajo el modelo de la Asociación Americana de Fisioterapia. *Boletín Informativo CEI*, 4(3).

Morone, G., Paolucci, S., & Iosa, M. (2015). In What Daily Activities Do Patients Achieve Independence after Stroke? *Journal of Stroke and Cerebrovascular Diseases: The Official Journal of National Stroke Association*, 24(8), 1931-1937.
<https://doi.org/10.1016/j.jstrokecerebrovasdis.2015.05.006>

Morris, D. M., Taub, E., & Mark, V. W. (2006). Constraint-induced movement therapy: Characterizing the intervention protocol. *Europa Medicophysica*, 42(3), 257-268.

Moruno Miralles, P., & Talavera-Valverde, M. Á. (2011). *Terapia Ocpacional en salud Mental*. Masson.

Moyano V., A. (2010). *El accidente cerebrovascular desde la mirada del rehabilitador*. 21(4), 348-355.

- Müller, R., & Büttner, P. (1994). A critical discussion of intraclass correlation coefficients. *Statistics in Medicine*, 13(23-24), 2465-2476. <https://doi.org/10.1002/sim.4780132310>
- Murciego Rubio, P., & García Atarés, N. (2019). Secuelas del daño cerebral adquirido, estudio sobre las necesidades terapéuticas. *Revista de Logopedia, Foniatría y Audiología*, 39(2), 52-58. <https://doi.org/10.1016/j.rlfa.2019.02.001>
- Murie-Fernández, M., Irimia, P., Martínez-Vila, E., John Meyer, M., & Teasell, R. (2010). Neurorrehabilitación tras el ictus. *Neurología*, 25(3), 189-196. [https://doi.org/10.1016/S0213-4853\(10\)70008-6](https://doi.org/10.1016/S0213-4853(10)70008-6)
- Mustafa, R. A., Santesso, N., Brozek, J., Akl, E. A., Walter, S. D., Norman, G., Kulasegaram, M., Christensen, R., Guyatt, G. H., Falck-Ytter, Y., Chang, S., Murad, M. H., Vist, G. E., Lasserson, T., Gartlehner, G., Shukla, V., Sun, X., Whittington, C., Post, P. N., ... Schünemann, H. J. (2013). The GRADE approach is reproducible in assessing the quality of evidence of quantitative evidence syntheses. *Journal of Clinical Epidemiology*, 66(7), 736-742; quiz 742.e1-5. <https://doi.org/10.1016/j.jclinepi.2013.02.004>
- Nakibuuka, J., Sajatovic, M., Nankabirwa, J., Furlan, A. J., Kayima, J., Ddumba, E., Katabira, E., & Byakika-Tusiime, J. (2015). Stroke-Risk Factors Differ between Rural and Urban Communities: Population Survey in Central Uganda. *Neuroepidemiology*, 44(3), 156-165. <https://doi.org/10.1159/000381453>
- Naveiro Soneira, J., Castillo, J., Suárez, P., Aldrey Vázquez, J. M., Lema Bouzas, M., & Noya García, M. (1998). Tiempo para el daño cerebral por hiperglucemia en la isquemia aguda. *Revista de Neurología*, 26(153), 790. <https://doi.org/10.33588/rn.26153.981014>
- Ng, S. S. M., Chen, P., Chang, H.-S., Chun, W.-K., Kong, T.-H., Lam, Y., Lau, W.-Y., Liu, T.-W., Ho, L. Y. W., Lai, C. Y. Y., & Tse, M. M. Y. (2023). Psychometric properties of Upper-body Dressing Scale in people with stroke. *Journal of Rehabilitation Medicine*, 55, jrm00391. <https://doi.org/10.2340/jrm.v55.5766>
- NICE. (2024). *National Institute for Health and Care Excellence*. <https://www.nice.org.uk/>
- Noce, M.-P., Valera-Gran, D., Hurtado-Pomares, M., Serrano-Reina, E., Soler-Pons, C., & Navarrete-Muñoz, E.-M. (2024). Spanish translation and cross-cultural adaptation of the Box and Block Test: A pilot study in adults with chronic acquired brain injury. *Brain Impairment: A Multidisciplinary Journal of the Australian Society for the Study of Brain Impairment*, 25, IB24014. <https://doi.org/10.1071/IB24014>

- Noé Sebastián, E., Balasch Bernat, M., Colomer Font, C., Moliner Muñoz, B., Rodríguez Sánchez-Leiva, C., Ugart, P., Llorens, R., & Ferri Campos, J. (2017). Ictus y discapacidad: Estudio longitudinal en pacientes con discapacidad moderada-grave tras un ictus incluidos en un programa de rehabilitación multidisciplinar. *Revista de Neurología*, 64(09), 385. <https://doi.org/10.33588/rn.6409.2016527>
- Norlander, A., Carlstedt, E., Jönsson, A.-C., Lexell, E. M., Ståhl, A., Lindgren, A., & Iwarsson, S. (2016). Long-Term Predictors of Social and Leisure Activity 10 Years after Stroke. *PloS One*, 11(2), e0149395. <https://doi.org/10.1371/journal.pone.0149395>
- Norman, G. R. (1989). Issues in the use of change scores in randomized trials. *Journal of Clinical Epidemiology*, 42(11), 1097-1105. [https://doi.org/10.1016/0895-4356\(89\)90051-6](https://doi.org/10.1016/0895-4356(89)90051-6)
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed). McGraw-Hill.
- Odgaard-Jensen, J., Vist, G. E., Timmer, A., Kunz, R., Akl, E. A., Schünemann, H., Briel, M., Nordmann, A. J., Pregno, S., & Oxman, A. D. (2011). Randomisation to protect against selection bias in healthcare trials. *The Cochrane Database of Systematic Reviews*, 2011(4), MR000012. <https://doi.org/10.1002/14651858.MR000012.pub3>
- Onose, G., Anghelescu, A., Ionescu, A., Tataranu, L. G., Spînu, A., Bumbea, A. M., Toader, C., Tuță, S., Carare, R. O., Popescu, C., Munteanu, C., Collaborative Working Group, & Daia, C. (2023). Translation of the Fugl-Meyer assessment into Romanian: Transcultural and semantic-linguistic adaptations and clinical validation. *Frontiers in Neurology*, 13, 1022546. <https://doi.org/10.3389/fneur.2022.1022546>
- Ortiz, M. S., & Fernández-Pera, M. (2018). Modelo de Ecuaciones Estructurales: Una guía para ciencias médicas y ciencias de la salud. *Terapia Psicológica*, 36(1), 51-57. <https://doi.org/10.4067/s0718-48082017000300047>
- Oviedo, H. C. & Campo-Arias, A. (2005). Aproximación al uso del coeficiente alfa de Cronbach. *Revista Colombiana de Psiquiatría*, 34(4), 572-580.
- Pacheco-Barrios, K., Giannoni-Luza, S., Navarro-Flores, A., Rebello-Sanchez, I., Parente, J., Balbuena, A., de Melo, P. S., Otiniano-Sifuentes, R., Rivera-Torrejón, O., Abanto, C., Alva-Díaz, C., Musolino, P. L., & Fregni, F. (2022). Burden of Stroke and Population-Attributable Fractions of Risk Factors in Latin America and the Caribbean. *Journal of the American Heart Association*, 11(21), e027044. <https://doi.org/10.1161/JAHA.122.027044>

- Page, S. J., Levine, P., Sisto, S. A., & Johnston, M. V. (2001). Mental practice combined with physical practice for upper-limb motor deficit in subacute stroke. *Physical Therapy*, 81(8), 1455-1462. <https://doi.org/10.1093/ptj/81.8.1455>
- Patrick, D. L., Burke, L. B., Gwaltney, C. J., Leidy, N. K., Martin, M. L., Molsen, E., & Ring, L. (2011a). Content validity--establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: part 1--eliciting concepts for a new PRO instrument. *Value in Health: The Journal of the International Society for Pharmacoeconomics and Outcomes Research*, 14(8), 967-977. <https://doi.org/10.1016/j.jval.2011.06.014>
- Patrick, D. L., Burke, L. B., Gwaltney, C. J., Leidy, N. K., Martin, M. L., Molsen, E., & Ring, L. (2011b). Content validity--establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO Good Research Practices Task Force report: part 2--assessing respondent understanding. *Value in Health: The Journal of the International Society for Pharmacoeconomics and Outcomes Research*, 14(8), 978-988. <https://doi.org/10.1016/j.jval.2011.06.013>
- Pedretti, L. W. (Ed.). (2001). *Occupational therapy: Practice skills for physical dysfunction* (5. ed). Mosby.
- Peña-Casanova, J. (2014). *Neurología de la conducta y neuropsicología* ([Versión electrónica de la 1a. edición])
- Piernik-Yoder, B., & Beck, A. (2012). The use of standardized assessments in occupational therapy in the United States. *Occupational Therapy in Health Care*, 26(2-3), 97-108. <https://doi.org/10.3109/07380577.2012.695103>
- Planton, M., Peiffer, S., Albucher, J. F., Barbeau, E. J., Tardy, J., Pastor, J., Januel, A. C., Bezy, C., Lemesle, B., Puel, M., Demonet, J. F., Chollet, F., & Pariente, J. (2012). Neuropsychological outcome after a first symptomatic ischaemic stroke with «good recovery». *European Journal of Neurology*, 19(2), 212-219. <https://doi.org/10.1111/j.1468-1331.2011.03450.x>
- Polit & Hungler. (2000). *Investigación Científica en Ciencias de la Salud*.
- Pollock, A., Farmer, S. E., Brady, M. C., Langhorne, P., Mead, G. E., Mehrholz, J., & Van Wijck, F. (2014). Interventions for improving upper limb function after stroke. *Cochrane Database of Systematic Reviews*, 2014(11). <https://doi.org/10.1002/14651858.CD010820.pub2>

- Pollock, N. (1993). Client-centered assessment. *The American Journal of Occupational Therapy: Official Publication of the American Occupational Therapy Association*, 47(4), 298-301. <https://doi.org/10.5014/ajot.47.4.298>
- Prieto, L., Lamarca, R., & Casado, A. (1998). [Assessment of the reliability of clinical findings: The intraclass correlation coefficient]. *Medicina Clínica*, 110(4), 142-145.
- Prinsen, C. a. C., Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., de Vet, H. C. W., & Terwee, C. B. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, 27(5), 1147-1157. <https://doi.org/10.1007/s11136-018-1798-3>
- Prinsen, C. A. C., Vohra, S., Rose, M. R., Boers, M., Tugwell, P., Clarke, M., Williamson, P. R., & Terwee, C. B. (2016a). How to select outcome measurement instruments for outcomes included in a “Core Outcome Set” – a practical guideline. *Trials*, 17(1), 449. <https://doi.org/10.1186/s13063-016-1555-2>
- Pohjasvaara, T., Erkinjuntti, T., Vataja, R., & Kaste, M. (1997). Comparison of stroke features and disability in daily life in patients with ischemic stroke aged 55 to 70 and 71 to 85 years. *Stroke*, 28(4), 729-735. <https://doi.org/10.1161/01.str.28.4.729>
- Polonio López, B., & Romero Ayuso, D. M. (2010). *Terapia Ocupacional aplicada al Daño Cerebral Adquirido* (Editorial Médica Panamericana). https://www.medicapanamericana.com/international/libros/terapia-ocupacional-aplicada-al-dano-cerebral-adquirido-coleccion-terapia-ocupacional?srsId=AfmBOogaRHSWtFluAjbUJ4GFHh9HNZnXjdxjW_jHWyCa09_BOb_PKDDp
- Purroy, F., Vena, A., Forné, C., de Arce, A. M., Dávalos, A., Fuentes, B., Arenillas, J. F., Krupinski, J., Gómez-Choco, M., Palomeras, E., Martí-Fábregas, J., Castillo, J., Ustrell, X., Tejada, J., Masjuan, J., Garcés, M., Benabdelhak, I., & Serena, J. (2019). Age- and Sex-Specific Risk Profiles and In-Hospital Mortality in 13,932 Spanish Stroke Patients. *Cerebrovascular Diseases (Basel, Switzerland)*, 47(3-4), 151-164. <https://doi.org/10.1159/000500205>
- Purves, D. (2016). *Neurociencia, 5a. Edición*. Editorial Médica Panamericana.
- Rabadi, M. H., & Rabadi, F. M. (2006). Comparison of the Action Research Arm Test and the Fugl-Meyer Assessment as Measures of Upper-Extremity Motor Weakness After Stroke. *Archives of Physical Medicine and Rehabilitation*, 87(7), 962-966. <https://doi.org/10.1016/j.apmr.2006.02.036>

- Rafsten, L., Danielsson, A., & Sunnerhagen, K. (2018). Anxiety after stroke: A systematic review and meta-analysis. *Journal of Rehabilitation Medicine*, 50(9), 769-778. <https://doi.org/10.2340/16501977-2384>
- Ramada-Rodilla, J. M., Serra-Pujadas, C., & Delclós-Clanchet, G. L. (2013). [Cross-cultural adaptation and health questionnaires validation: Revision and methodological recommendations]. *Salud Publica De Mexico*, 55(1), 57-66. <https://doi.org/10.1590/s0036-36342013000100009>
- Rand, D., Eng, J. J., Tang, P.-F., Hung, C., & Jeng, J.-S. (2010). Daily physical activity and its contribution to the health-related quality of life of ambulatory individuals with chronic stroke. *Health and Quality of Life Outcomes*, 8, 80. <https://doi.org/10.1186/1477-7525-8-80>
- Rathore, S. S., Hinn, A. R., Cooper, L. S., Tyroler, H. A., & Rosamond, W. D. (2002). Characterization of incident stroke signs and symptoms: Findings from the atherosclerosis risk in communities study. *Stroke*, 33(11), 2718-2721. <https://doi.org/10.1161/01.str.0000035286.87503.31>
- Redolar Ripoll, D. (2023). *Neurociencia cognitiva* (2. edición). Editorial Médica Panamericana.
- Reed, K. L., & Sanderson, S. N. (1999). *Concepts of occupational therapy* (4th ed). Lippincott Williams & Wilkins.
- Revet, M., Immerzeel, J., Voogt, L., & Paulis, W. (2021). Patients with neuropsychological disorders short after stroke have worse functional outcome: A systematic review and meta-analysis. *Disability and Rehabilitation*, 43(16), 2233-2252. <https://doi.org/10.1080/09638288.2019.1693642>
- Revicki, D., Hays, R. D., Cella, D., & Sloan, J. (2008). Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *Journal of Clinical Epidemiology*, 61(2), 102-109. <https://doi.org/10.1016/j.jclinepi.2007.03.012>
- Ribeiro De Souza, F., Sales, M., Rabelo Laporte, L., Melo, A., & Manoel Da Silva Ribeiro, N. (2023). Body structure/function impairments and activity limitations of post-stroke that predict social participation: A systematic review. *Topics in Stroke Rehabilitation*, 30(6), 589-602. <https://doi.org/10.1080/10749357.2022.2095086>
- Rochmah, T. N., Rahmawati, I. T., Dahlui, M., Budiarto, W., & Bilqis, N. (2021). Economic Burden of Stroke Disease: A Systematic Review. *International Journal of Environmental Research and Public Health*, 18(14), 7552. <https://doi.org/10.3390/ijerph18147552>

- Rogers, J. C., & Holm, M. B. (1991). Occupational therapy diagnostic reasoning: A component of clinical reasoning. *The American Journal of Occupational Therapy: Official Publication of the American Occupational Therapy Association*, 45(11), 1045-1053. <https://doi.org/10.5014/ajot.45.11.1045>
- Romero Ayuso & Moruno Miralles. (2003). *Terapia Ocupacional: Teorías y Técnicas*. Masson.
- Ropper, A. H., Samuels, M. A., & Klein, J. P. (2017). *Adams y Victor. Principios de Neurología* (10a ed). McGraw Hill.
- Rothman, M., Burke, L., Erickson, P., Leidy, N. K., Patrick, D. L., & Petrie, C. D. (2009). Use of existing patient-reported outcome (PRO) instruments and their modification: The ISPOR Good Research Practices for Evaluating and Documenting Content Validity for the Use of Existing Instruments and Their Modification PRO Task Force Report. *Value in Health: The Journal of the International Society for Pharmacoeconomics and Outcomes Research*, 12(8), 1075-1083. <https://doi.org/10.1111/j.1524-4733.2009.00603.x>
- Sacco, R. L., Kasner, S. E., Broderick, J. P., Caplan, L. R., Connors, J. J. B., Culebras, A., Elkind, M. S. V., George, M. G., Hamdan, A. D., Higashida, R. T., Hoh, B. L., Janis, L. S., Kase, C. S., Kleindorfer, D. O., Lee, J.-M., Moseley, M. E., Peterson, E. D., Turan, T. N., Valderrama, A. L., ... Council on Nutrition, Physical Activity and Metabolism. (2013). An updated definition of stroke for the 21st century: A statement for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke*, 44(7), 2064-2089. <https://doi.org/10.1161/STR.0b013e318296aeca>
- Sackett, D. L., Rosenberg, W. M., Gray, J. A., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: What it is and what it isn't. *BMJ (Clinical Research Ed.)*, 312(7023), 71-72. <https://doi.org/10.1136/bmj.312.7023.71>
- Sánchez Blanco, I., Ochoa Sangrador, C., & Izquierdo Sánchez, M. (2000). *Pronóstico de recuperación funcional en personas supervivientes de un ictus*. 34(6), 412-422.
- Sánchez Cabeza, Á. (2013). *Terapia ocupacional basada en la evidencia para la rehabilitación del daño cerebral adquirido*. Extra 8. <https://dialnet.unirioja.es/servlet/articulo?codigo=4641369>
- Sánchez Pedraza, R. & Gómez Restrepo, C. I. (1998). Conceptos básicos sobre validación de escalas. *Revista Colombiana de Psiquiatría*, 27(2), 121-130.
- Sánchez, R., & Echeverry, J. (2004). [Validating scales used for measuring factors in medicine]. *Revista De Salud Publica (Bogota, Colombia)*, 6(3), 302-318.

- Sanford, J., Moreland, J., Swanson, L. R., Stratford, P. W., & Gowland, C. (1993). Reliability of the Fugl-Meyer Assessment for Testing Motor Performance in Patients Following Stroke. *Physical Therapy*, 73(7), 447-454. <https://doi.org/10.1093/ptj/73.7.447>
- Santisteban, L., Térémetz, M., Bleton, J.-P., Baron, J.-C., Maier, M. A., & Lindberg, P. G. (2016). Upper Limb Outcome Measures Used in Stroke Rehabilitation Studies: A Systematic Literature Review. *PLOS ONE*, 11(5), e0154792. <https://doi.org/10.1371/journal.pone.0154792>
- Saris, W. E., Satorra, A., & Van Der Veld, W. M. (2009). Testing Structural Equation Models or Detection of Misspecifications? *Structural Equation Modeling: A Multidisciplinary Journal*, 16(4), 561-582. <https://doi.org/10.1080/10705510903203433>
- Sarker, S.-J., Rudd, A. G., Douiri, A., & Wolfe, C. D. A. (2012). Comparison of 2 Extended Activities of Daily Living Scales With the Barthel Index and Predictors of Their Outcomes: Cohort Study Within the South London Stroke Register (SLSR). *Stroke*, 43(5), 1362-1369. <https://doi.org/10.1161/STROKEAHA.111.645234>
- Schmitt, T. A. (2011). Current Methodological Considerations in Exploratory and Confirmatory Factor Analysis. *Journal of Psychoeducational Assessment*, 29(4), 304-321. <https://doi.org/10.1177/0734282911406653>
- Schrage, T., Thomalla, G., Härter, M., Lebherz, L., Appelbohm, H., Rimmele, D. L., & Kriston, L. (2023). Predictors of Discharge Destination After Stroke. *Neurorehabilitation and Neural Repair*, 37(5), 307-315. <https://doi.org/10.1177/15459683231166935>
- Schumacker, R. E., & Lomax, R. G. (2004). *A Beginner's Guide to Structural Equation Modeling* (0 ed.). Psychology Press. <https://doi.org/10.4324/9781410610904>
- Schünemann, H. J., Fretheim, A., & Oxman, A. D. (2006b). Improving the use of research evidence in guideline development: 9. Grading evidence and recommendations. *Health Research Policy and Systems*, 4, 21. <https://doi.org/10.1186/1478-4505-4-21>
- Schünemann, H. J., Fretheim, A., Oxman, A. D., & WHO Advisory Committee on Health Research. (2006a). Improving the use of research evidence in guideline development: 1. Guidelines for guidelines. *Health Research Policy and Systems*, 4, 13. <https://doi.org/10.1186/1478-4505-4-13>
- Schünemann, H. J., Oxman, A. D., Brozek, J., Glasziou, P., Jaeschke, R., Vist, G. E., Williams, J. W., Kunz, R., Craig, J., Montori, V. M., Bossuyt, P., Guyatt, G. H., & GRADE Working Group. (2008). Grading quality of evidence and strength of recommendations for diagnostic tests and

- strategies. *BMJ (Clinical Research Ed.)*, 336(7653), 1106-1110. <https://doi.org/10.1136/bmj.39500.677199.AE>
- See, J., Dodakian, L., Chou, C., Chan, V., McKenzie, A., Reinkensmeyer, D. J., & Cramer, S. C. (2013). A Standardized Approach to the Fugl-Meyer Assessment and Its Implications for Clinical Trials. *Neurorehabilitation and Neural Repair*, 27(8), 732-741. <https://doi.org/10.1177/1545968313491000>
- Selles, R. W., Andrinopoulou, E.-R., Nijland, R. H., van der Vliet, R., Slaman, J., van Wegen, E. E., Rizopoulos, D., Ribbers, G. M., Meskers, C. G., & Kwakkel, G. (2021). Computerised patient-specific prediction of the recovery profile of upper limb capacity within stroke services: The next step. *Journal of Neurology, Neurosurgery, and Psychiatry*, 92(6), 574-581. <https://doi.org/10.1136/jnnp-2020-324637>
- SEN. (2019). *El atlas del ictus en España*. Sociedad Española de Neurología (SEN). <https://www.sen.es/component/content/article/91-articulos/2617-el-atlas-del-ictus?Itemid=437>
- Servicio Andaluz de Salud. (2012). *Niveles de Evidencia*. Hospital Universitario Reina Sofía. Conserjería de Salud y Consumo. Servicio Andaluz de Salud. https://www.sspa.juntadeandalucia.es/servicioandaluzdesalud/hrs3/index.php?id=manual_procedimientos_niveles
- Shanks, M. J., & Byblow, W. D. (2025). Corticomotor pathway function and recovery after stroke: A look back and a way forward. *The Journal of Physiology*, 603(3), 651-662. <https://doi.org/10.1113/JP285562>
- Shi, Y., Yang, D., Zeng, Y., & Wu, W. (2017). Risk Factors for Post-stroke Depression: A Meta-analysis. *Frontiers in Aging Neuroscience*, 9, 218. <https://doi.org/10.3389/fnagi.2017.00218>
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420-428. <https://doi.org/10.1037//0033-2909.86.2.420>
- Siemieniuk, R., & Guyatt, G. (2019). *What is GRADE*. BMJ Best Practice. <https://dev-bestpractice.bmjgroup.com/info/us/toolkit/learn-ebm/what-is-grade/>
- Sign. (2025). *SIGN 50: A guideline developer's handbook*. <https://www.sign.ac.uk/our-guidelines/sign-50-a-guideline-developers-handbook/>
- Sijtsma, K. (2009). On the Use, the Misuse, and the Very Limited Usefulness of Cronbach's Alpha. *Psychometrika*, 74(1), 107-120. <https://doi.org/10.1007/s11336-008-9101-0>

- Singam, A., Ytterberg, C., Tham, K., & von Koch, L. (2015). Participation in Complex and Social Everyday Activities Six Years after Stroke: Predictors for Return to Pre-Stroke Level. *PloS One*, 10(12), e0144344. <https://doi.org/10.1371/journal.pone.0144344>
- Siotto, M., Germanotta, M., Santoro, M., Cipollini, V., Guardati, G., Papadopoulou, D., Bray, E., Mastrosera, A., & Aprile, I. (2021). Serotonin Levels and Cognitive Recovery in Patients with Subacute Stroke after Rehabilitation Treatment. *Brain Sciences*, 11(5), 642. <https://doi.org/10.3390/brainsci11050642>
- Smith, M.-C., Barber, A. P., Scrivener, B. J., & Stinear, C. M. (2022). The TWIST Tool Predicts When Patients Will Recover Independent Walking After Stroke: An Observational Study. *Neurorehabilitation and Neural Repair*, 36(7), 461-471. <https://doi.org/10.1177/15459683221085287>
- Soleimani, M., Ghazisaeedi, M., & Heydari, S. (2024). The efficacy of virtual reality for upper limb rehabilitation in stroke patients: A systematic review and meta-analysis. *BMC Medical Informatics and Decision Making*, 24(1), 135. <https://doi.org/10.1186/s12911-024-02534-y>
- Soler Cárdenas, S. F. (2008). Coeficientes de confiabilidad de instrumentos escritos en el marco de la teoría clásica de los test. *Educación médica superior*, 22(2). http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-21412008000200006
- Soler Cárdenas, S. F. & Soler Pons, L. (2012). Usos del coeficiente alfa de Cronbach en el análisis de instrumentos escritos. *Revista Médica Electrónica*, 34(1), 1-6.
- Spaccavento, S., Marinelli, C. V., Nardulli, R., Macchitella, L., Bivona, U., Piccardi, L., Zoccolotti, P., & Angelelli, P. (2019). Attention Deficits in Stroke Patients: The Role of Lesion Characteristics, Time from Stroke, and Concomitant Neuropsychological Deficits. *Behavioural Neurology*, 2019, 1-12. <https://doi.org/10.1155/2019/7835710>
- Sridharan, S. E., Unnikrishnan, J. P., Sukumaran, S., Sylaja, P. N., Nayak, S. D., Sarma, P. S., & Radhakrishnan, K. (2009). Incidence, types, risk factors, and outcome of stroke in a developing country: The Trivandrum Stroke Registry. *Stroke*, 40(4), 1212-1218. <https://doi.org/10.1161/STROKEAHA.108.531293>
- Stinear, C. M., Byblow, W. D., Ackerley, S. J., Smith, M.-C., Borges, V. M., & Barber, P. A. (2017). PREP2: A biomarker-based algorithm for predicting upper limb function after stroke. *Annals of Clinical and Translational Neurology*, 4(11), 811-820. <https://doi.org/10.1002/acn3.488>

- Stinear, C. M., Smith, M.-C., & Byblow, W. D. (2019). Prediction Tools for Stroke Rehabilitation. *Stroke*, 50(11), 3314-3322. <https://doi.org/10.1161/STROKEAHA.119.025696>
- Stratford, P. W., & Riddle, D. L. (2005). Assessing sensitivity to change: Choosing the appropriate change coefficient. *Health and Quality of Life Outcomes*, 3, 23. <https://doi.org/10.1186/1477-7525-3-23>
- Strauss, M. E. (2005). Introduction to the special section on construct validity of psychological tests: 50 years after Cronbach and Meehl (1955). *Psychological Assessment*, 17(4), 395. <https://doi.org/10.1037/1040-3590.17.4.395>
- Strauss, M. E., & Smith, G. T. (2009). Construct validity: Advances in theory and methodology. *Annual Review of Clinical Psychology*, 5, 1-25. <https://doi.org/10.1146/annurev.clinpsy.032408.153639>
- Streiner, D. L. (1993). A checklist for evaluating the usefulness of rating scales. *Canadian Journal of Psychiatry. Revue Canadienne De Psychiatrie*, 38(2), 140-148. <https://doi.org/10.1177/070674379303800214>
- Streiner, D. L. (1994). Figuring out factors: The use and misuse of factor analysis. *Canadian Journal of Psychiatry. Revue Canadienne De Psychiatrie*, 39(3), 135-140. <https://doi.org/10.1177/070674379403900303>
- Streiner, D. L. (2003). Starting at the Beginning: An Introduction to Coefficient Alpha and Internal Consistency. *Journal of Personality Assessment*, 80(1), 99-103. https://doi.org/10.1207/S15327752JPA8001_18
- Streiner, D. L., & Norman, G. R. (2008). *Health Measurement Scales*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199231881.001.0001>
- Streiner, D. L., Norman, G. R., & Cairney, J. (2015). *Health Measurement Scales* (Vol. 1). Oxford University Press. <https://doi.org/10.1093/med/9780199685219.001.0001>
- Strilciuc, S., Grad, D. A., Radu, C., Chira, D., Stan, A., Ungureanu, M., Gheorghe, A., & Muresanu, F.-D. (2021). The economic burden of stroke: A systematic review of cost of illness studies. *Journal of Medicine and Life*, 14(5), 606-619. <https://doi.org/10.25122/jml-2021-0361>
- Stucki, G., Liang, M. H., Fossel, A. H., & Katz, J. N. (1995). Relative responsiveness of condition-specific and generic health status measures in degenerative lumbar spinal stenosis. *Journal of Clinical Epidemiology*, 48(11), 1369-1378. [https://doi.org/10.1016/0895-4356\(95\)00054-2](https://doi.org/10.1016/0895-4356(95)00054-2)

- Sullivan, K. J., Tilson, J. K., Cen, S. Y., Rose, D. K., Hershberg, J., Correa, A., Gallichio, J., McLeod, M., Moore, C., Wu, S. S., & Duncan, P. W. (2011). Fugl-Meyer assessment of sensorimotor function after stroke: Standardized training procedure for clinical practice and clinical trials. *Stroke*, 42(2), 427-432. <https://doi.org/10.1161/STROKEAHA.110.592766>
- Sun, Y., Boots, J., & Zehr, E. P. (2015). The lingering effects of a busted myth—False time limits in stroke rehabilitation. *Applied Physiology, Nutrition, and Metabolism = Physiologie Appliquee, Nutrition Et Metabolisme*, 40(8), 858-861. <https://doi.org/10.1139/apnm-2014-0523>
- Sunderland, A., Tinson, D., Bradley, L., & Hewer, R. L. (1989). Arm function after stroke. An evaluation of grip strength as a measure of recovery and a prognostic indicator. *Journal of Neurology, Neurosurgery, and Psychiatry*, 52(11), 1267-1272. <https://doi.org/10.1136/jnnp.52.11.1267>
- Sunderland, A., Walker, C. M., & Walker, M. F. (2006). Action errors and dressing disability after stroke: An ecological approach to neuropsychological assessment and intervention. *Neuropsychological Rehabilitation*, 16(6), 666-683. <https://doi.org/10.1080/09602010500204385>
- Suzuki, M., Yamada, S., Omori, M., Hatakeyama, M., Sugimura, Y., Matsushita, K., & Tagawa, Y. (2008). Development of the upper-body dressing scale for a buttoned shirt: A preliminary correlational study. *American Journal of Physical Medicine & Rehabilitation*, 87(9), 740-749. <https://doi.org/10.1097/PHM.0b013e31818378b0>
- Svensson, J. S., Westerlind, E., Persson, H. C., & Sunnerhagen, K. S. (2019). Occupational gaps 5 years after stroke. *Brain and Behavior*, 9(3), e01234. <https://doi.org/10.1002/brb3.1234>
- Szekeres, M. (2018). Clinical relevance commentary in response to: The validity and clinical utility of the Disabilities of the Arm Shoulder and Hand questionnaire for hand injuries in developing country contexts: A systematic review. *Journal of Hand Therapy: Official Journal of the American Society of Hand Therapists*, 31(1), 91-92. <https://doi.org/10.1016/j.jht.2017.12.002>
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6. ed., international ed). Pearson.
- Talavera-Valverde, M.-Á., Souto Gómez*, A. I., & Moruno-Miralles, P. (2022). Evaluación ocupacional y el marco de trabajo para la práctica de la terapia ocupacional: Occupational evaluation and the working framework for the practice of occupational therapy. *Revista Terapia Ocupacional Galicia*, 19(1E), S62-S66.
- Tangarife, J. L. & Arias, J. A. C. (2015). Construcción y validación de escalas de medición en salud: Revisión de propiedades psicométricas. *Archivos de*

Medicina, 11(3). <https://www.archivosdemedicina.com/medicina-de-familia/construccin-y-validacin-de-escalas-de-medicin-en-salud-revisin-de-propiedades-psicomtricas.pdf>

Tauchi, Y., Kyougoku, M., Takahashi, K., Okita, Y., & Takebayashi, T. (2022). Dimensionality and item-difficulty hierarchy of the Fugl-Meyer assessment of the upper extremity among Japanese patients who have experienced stroke. *Topics in Stroke Rehabilitation*, 29(8), 579-587. <https://doi.org/10.1080/10749357.2021.1965797>

Tauchi, Y., Kyougoku, M., Okita, Y., & Takebayashi, T. (2023). Structural validity and internal consistency of a hypothesized factor structure of the Fugl-Meyer Assessment of the upper extremity. *Topics in Stroke Rehabilitation*, 30(5), 501-511. <https://doi.org/10.1080/10749357.2022.2070361>

Teasell, R., Salbach, N. M., & Members of the Canadian Stroke Best Practice Recommendations. (2019). *Canadian Stroke Best Practice Recommendations*. Canadian Stroke Best Practice. <https://www.heartandstroke.ca/-/media/1-stroke-best-practices/rehabilitation-nov2019/2019-csbpr6-rehabrecovery-module-eng-final-dec2019.ashx?rev=d9be6748ea0945368a0733e6b26423ae>

Terwee, C. B., Bot, S. D. M., de Boer, M. R., van der Windt, D. A. W. M., Knol, D. L., Dekker, J., Bouter, L. M., & de Vet, H. C. W. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, 60(1), 34-42. <https://doi.org/10.1016/j.jclinepi.2006.03.012>

Terwee, C. B., Dekker, F. W., Wiersinga, W. M., Prummel, M. F., & Bossuyt, P. M. M. (2003). On assessing responsiveness of health-related quality of life instruments: Guidelines for instrument evaluation. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, 12(4), 349-362. <https://doi.org/10.1023/a:1023499322593>

Terwee, C. B., Prinsen, C. a. C., Chiarotto, A., Westerman, M. J., Patrick, D. L., Alonso, J., Bouter, L. M., de Vet, H. C. W., & Mokkink, L. B. (2018a). COSMIN methodology for evaluating the content validity of patient-reported outcome measures: A Delphi study. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, 27(5), 1159-1170. <https://doi.org/10.1007/s11136-018-1829-0>

Terwee, C. B., Prinsen, C. a. C., Ricci Garotti, M. G., Suman, A., de Vet, H. C. W., & Mokkink, L. B. (2016). The quality of systematic reviews of health-related outcome measurement instruments. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, 25(4), 767-779. <https://doi.org/10.1007/s11136-015-1122-4>

- Thielbar, K. O., Triandafilou, K. M., Fischer, H. C., O'Toole, J. M., Corrigan, M. L., Ochoa, J. M., Stoykov, M. E., & Kamper, D. G. (2017). Benefits of Using a Voice and EMG-Driven Actuated Glove to Support Occupational Therapy for Stroke Survivors. *IEEE Transactions on Neural Systems and Rehabilitation Engineering: A Publication of the IEEE Engineering in Medicine and Biology Society*, 25(3), 297-305. <https://doi.org/10.1109/TNSRE.2016.2569070>
- Tornimbeni, S., Pérez, E., & Olaz, F. (2008). *Introducción a la Psicometría* (1.^a ed.). Paidós.
- Torres Ávila, I. (2014). *Evidencia del tratamiento desde terapia ocupacional en actividades de la vida diaria en pacientes con accidente cerebrovascular*. 11(19). <https://dialnet.unirioja.es/servlet/articulo?codigo=4710575>
- Tovar, A., de Oliveira Sousa, S. L., León Garzón, M. C., & González Carrillo, M. J. (2021). *Efectos del entrenamiento muscular inspiratorio sobre la función respiratoria y el equilibrio en supervivientes de ictus*. 72(4), 112-120.
- Tovar, J. (2007). *Psicometría: Test psicométricos, confiabilidad y validez*. 8, 85-108.
- Trombly, C. (1993). Anticipating the future: Assessment of occupational function. *The American Journal of Occupational Therapy: Official Publication of the American Occupational Therapy Association*, 47(3), 253-257. <https://doi.org/10.5014/ajot.47.3.253>
- Trombly, C. A., & Radomski, M. V. (Eds.). (2002). *Occupational therapy for physical dysfunction* (5. ed). Lippincott Williams & Wilkins.
- Tsouna-Hadjis, E., Vemmos, K. N., Zakopoulos, N., & Stamatelopoulos, S. (2000). First-stroke recovery process: The role of family social support. *Archives of Physical Medicine and Rehabilitation*, 81(7), 881-887. <https://doi.org/10.1053/apmr.2000.4435>
- Tyson, S. F., Hanley, M., Chillala, J., Selley, A. B., & Tallis, R. C. (2008). Sensory Loss in Hospital-Admitted People With Stroke: Characteristics, Associated Factors, and Relationship With Function. *Neurorehabilitation and Neural Repair*, 22(2), 166-172. <https://doi.org/10.1177/1545968307305523>
- Unnithan, A. K. A., Das, J. M., & Mehta, P. (2025). Hemorrhagic Stroke. En *StatPearls*. StatPearls Publishing. <http://www.ncbi.nlm.nih.gov/books/NBK559173/>
- U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research, U.S. Department of Health and Human Services FDA Center for Biologics Evaluation and Research, & U.S. Department of Health and Human Services FDA Center for Devices and Radiological Health. (2006). Guidance for industry: Patient-reported outcome measures: use in

medical product development to support labeling claims: draft guidance. *Health and Quality of Life Outcomes*, 4, 79. <https://doi.org/10.1186/1477-7525-4-79>

Ustrell-Roig, X., & Serena-Leal, J. (2007). Ictus. Diagnóstico y tratamiento de las enfermedades cerebrovasculares. *Revista Española de Cardiología*, 60(7), 753-769. <https://doi.org/10.1157/13108281>

Van der Lee, J. H., Beckerman, H., Lankhorst, G. J., & Bouter, L. M. (2001). The responsiveness of the Action Research Arm test and the Fugl-Meyer Assessment scale in chronic stroke patients. *Journal of Rehabilitation Medicine*, 33(3), 110-113. <https://doi.org/10.1080/165019701750165916>

Van der Lee, J. H., De Groot, V., Beckerman, H., Wagenaar, R. C., Lankhorst, G. J., & Bouter, L. M. (2001). The intra- and interrater reliability of the action research arm test: A practical test of upper extremity function in patients with stroke. *Archives of Physical Medicine and Rehabilitation*, 82(1), 14-19. <https://doi.org/10.1053/apmr.2001.18668>

Veerbeek, J. M., Kwakkel, G., van Wegen, E. E. H., Ket, J. C. F., & Heymans, M. W. (2011). Early prediction of outcome of activities of daily living after stroke: A systematic review. *Stroke*, 42(5), 1482-1488. <https://doi.org/10.1161/STROKEAHA.110.604090>

Veerbeek, J. M., Pohl, J., Luft, A. R., & Held, J. P. O. (2022). External validation and extension of the Early Prediction of Functional Outcome after Stroke (EPOS) prediction model for upper limb outcome 3 months after stroke. *PloS One*, 17(8), e0272777. <https://doi.org/10.1371/journal.pone.0272777>

Verdejo García, A., & Bechara, A. (2010). *Neuropsicología de las funciones ejecutivas*. 22(2), 227-235.

Viktorisson, A., Reinholdsson, M., Danielsson, A., Palstam, A., & Sunnerhagen, K. S. (2022). Pre-stroke physical activity in relation to post-stroke outcomes - linked to the International Classification of Functioning, Disability and Health (ICF): A scoping review. *Journal of Rehabilitation Medicine*, 54, jrm00251. <https://doi.org/10.2340/jrm.v53.51>

Viosca, E., Lafuente, R., Martínez, J. L., Almagro, P. L., Gracia, A., & González, C. (2005). Walking recovery after an acute stroke: Assessment with a new functional classification and the Barthel Index. *Archives of Physical Medicine and Rehabilitation*, 86(6), 1239-1244. <https://doi.org/10.1016/j.apmr.2004.11.015>

Virani, S. S., Alonso, A., Benjamin, E. J., Bittencourt, M. S., Callaway, C. W., Carson, A. P., Chamberlain, A. M., Chang, A. R., Cheng, S., Delling, F. N., Djousse, L., Elkind, M. S. V., Ferguson, J. F., Fornage, M., Khan, S. S., Kissela, B. M., Knutson, K. L., Kwan, T. W., Lackland, D. T., ... American

Heart Association Council on Epidemiology and Prevention Statistics Committee and Stroke Statistics Subcommittee. (2020). Heart Disease and Stroke Statistics-2020 Update: A Report From the American Heart Association. *Circulation*, 141(9), e139-e596. <https://doi.org/10.1161/CIR.0000000000000757>

von Groote, P. M., Bickenbach, J. E., & Gutenbrunner, C. (2011). The World Report on Disability – implications, perspectives and opportunities for physical and rehabilitation medicine (PRM). *Journal of Rehabilitation Medicine*, 43(10), 869-875. <https://doi.org/10.2340/16501977-0872>

Wade, D. T., & Hower, R. L. (1987). Motor loss and swallowing difficulty after stroke: Frequency, recovery, and prognosis. *Acta Neurologica Scandinavica*, 76(1), 50-54. <https://doi.org/10.1111/j.1600-0404.1987.tb03543.x>

Wafa, H. A., Wolfe, C. D. A., Emmett, E., Roth, G. A., Johnson, C. O., & Wang, Y. (2020). Burden of Stroke in Europe: Thirty-Year Projections of Incidence, Prevalence, Deaths, and Disability-Adjusted Life Years. *Stroke*, 51(8), 2418-2427. <https://doi.org/10.1161/STROKEAHA.120.029606>

Wagner, J. M., Lang, C. E., Sahrman, S. A., Edwards, D. F., & Dromerick, A. W. (2007). Sensorimotor impairments and reaching performance in subjects with poststroke hemiparesis during the first few months of recovery. *Physical Therapy*, 87(6), 751-765. <https://doi.org/10.2522/ptj.20060135>

Walker, C. M., Sunderland, A., Sharma, J., & Walker, M. F. (2004). The impact of cognitive impairment on upper body dressing difficulties after stroke: A video analysis of patterns of recovery. *Journal of Neurology, Neurosurgery, and Psychiatry*, 75(1), 43-48.

Walker, C. M., Walker, M. F., & Sunderland, A. (2003). Dressing after a Stroke: A Survey of Current Occupational Therapy Practice. *British Journal of Occupational Therapy*, 66(6), 263-268. <https://doi.org/10.1177/030802260306600605>

Walker, M. (1992). Dressing ability after stroke. *Nursing Times*, 88(1), 51.

Walker, M. F., & Lincoln, N. B. (1990). Reacquisition of dressing skills after stroke. *International Disability Studies*, 12(1), 41-43. <https://doi.org/10.3109/03790799009166603>

Walker, M. F., & Lincoln, N. B. (1991). Factors influencing dressing performance after stroke. *Journal of Neurology, Neurosurgery, and Psychiatry*, 54(8), 699-701. <https://doi.org/10.1136/jnnp.54.8.699>

- Walker, M. F., Sunderland, A., Fletcher-Smith, J., Drummond, A., Logan, P., Edmans, J. A., Garvey, K., Dineen, R. A., Ince, P., Horne, J., Fisher, R. J., & Taylor, J. L. (2012). The DRESS trial: A feasibility randomized controlled trial of a neuropsychological approach to dressing therapy for stroke inpatients. *Clinical Rehabilitation*, 26(8), 675-685. <https://doi.org/10.1177/0269215511431089>
- Ween, J. E., Alexander, M. P., D'Esposito, M., & Roberts, M. (1996). Factors predictive of stroke outcome in a rehabilitation setting. *Neurology*, 47(2), 388-392. <https://doi.org/10.1212/wnl.47.2.388>
- Wei, X.-J., Tong, K., & Hu, X. (2011). The responsiveness and correlation between Fugl-Meyer Assessment, Motor Status Scale, and the Action Research Arm Test in chronic stroke with upper-extremity rehabilitation robotic training. *International Journal of Rehabilitation Research*, 34(4), 349-356. <https://doi.org/10.1097/MRR.0b013e32834d330a>
- Weinstock-Zlotnick, G., & Hinojosa, J. (2004). Bottom-up or top-down evaluation: Is one better than the other? *The American Journal of Occupational Therapy: Official Publication of the American Occupational Therapy Association*, 58(5), 594-599. <https://doi.org/10.5014/ajot.58.5.594>
- WHO. (2016). *Framework on integrated, people-centred health services: Report by the Secretariat*. World Health Organization. <https://iris.who.int/handle/10665/252698>
- Wild, D., Grove, A., Martin, M., Eremenco, S., McElroy, S., Verjee-Lorenz, A., Erikson, P., & ISPOR Task Force for Translation and Cultural Adaptation. (2005). Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: Report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value in Health: The Journal of the International Society for Pharmacoeconomics and Outcomes Research*, 8(2), 94-104. <https://doi.org/10.1111/j.1524-4733.2005.04054.x>
- Winstein, C. J., Stein, J., Arena, R., Bates, B., Cherney, L. R., Cramer, S. C., Deruyter, F., Eng, J. J., Fisher, B., Harvey, R. L., Lang, C. E., MacKay-Lyons, M., Ottenbacher, K. J., Pugh, S., Reeves, M. J., Richards, L. G., Stiers, W., Zorowitz, R. D., & American Heart Association Stroke Council, Council on Cardiovascular and Stroke Nursing, Council on Clinical Cardiology, and Council on Quality of Care and Outcomes Research. (2016). Guidelines for Adult Stroke Rehabilitation and Recovery: A Guideline for Healthcare Professionals From the American Heart Association/American Stroke Association. *Stroke*, 47(6), e98-e169. <https://doi.org/10.1161/STR.0000000000000098>
- Wolf, S. L., Catlin, P. A., Ellis, M., Archer, A. L., Morgan, B., & Piacentino, A. (2001). Assessing Wolf motor function test as outcome measure for

- research in patients after stroke. *Stroke*, 32(7), 1635-1639. <https://doi.org/10.1161/01.str.32.7.1635>
- Woodbury, M. L., Velozo, C. A., Richards, L. G., Duncan, P. W., Studenski, S., & Lai, S.-M. (2007). Dimensionality and Construct Validity of the Fugl-Meyer Assessment of the Upper Extremity. *Archives of Physical Medicine and Rehabilitation*, 88(6), 715-723. <https://doi.org/10.1016/j.apmr.2007.02.036>
- Wright, J. G., & Young, N. L. (1997). A comparison of different indices of responsiveness. *Journal of Clinical Epidemiology*, 50(3), 239-246. [https://doi.org/10.1016/s0895-4356\(96\)00373-3](https://doi.org/10.1016/s0895-4356(96)00373-3)
- Yao, M., Ren, Y., Jia, Y., Xu, J., Wang, Y., Zou, K., & Sun, X. (2023). Projected burden of stroke in China through 2050. *Chinese Medical Journal*, 136(13), 1598-1605. <https://doi.org/10.1097/CM9.0000000000002060>
- Yi, J. H., Jeon, J., Kwon, H., Cho, E., Yun, J., Lee, Y. C., Ryu, J. H., Park, S. J., Cho, J. H., & Kim, D. H. (2020). Rubrofusarin Attenuates Chronic Restraint Stress-Induced Depressive Symptoms. *International Journal of Molecular Sciences*, 21(10), 3454. <https://doi.org/10.3390/ijms21103454>
- Yost, K. J., Eton, D. T., Garcia, S. F., & Cella, D. (2011). Minimally important differences were estimated for six Patient-Reported Outcomes Measurement Information System-Cancer scales in advanced-stage cancer patients. *Journal of Clinical Epidemiology*, 64(5), 507-516. <https://doi.org/10.1016/j.jclinepi.2010.11.018>
- Youngstrom, M. J. (2002). The Occupational Therapy Practice Framework: The evolution of our professional language. *The American Journal of Occupational Therapy: Official Publication of the American Occupational Therapy Association*, 56(6), 607-608. <https://doi.org/10.5014/ajot.56.6.607>
- Yozbatiran, N., Der-Yeghiaian, L., & Cramer, S. C. (2008). A standardized approach to performing the action research arm test. *Neurorehabilitation and Neural Repair*, 22(1), 78-90. <https://doi.org/10.1177/1545968307305353>
- Yu, B., Kaku, A., Liu, K., Parnandi, A., Fokas, E., Venkatesan, A., Pandit, N., Ranganath, R., Schambra, H., & Fernandez-Granda, C. (2024). Quantifying impairment and disease severity using AI models trained on healthy subjects. *NPJ Digital Medicine*, 7(1), 180. <https://doi.org/10.1038/s41746-024-01173-x>
- Zhao, K., He, C., Xiang, W., Zhou, Y., Zhang, Z., Li, J., & Scano, A. (2023). Evidence of synergy coordination patterns of upper-limb motor control in stroke patients with mild and moderate impairment. *Frontiers in Physiology*, 14, 1214995. <https://doi.org/10.3389/fphys.2023.1214995>

- Zou, G. Y. (2005). Quantifying responsiveness of quality of life measures without an external criterion. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, 14(6), 1545-1552. <https://doi.org/10.1007/s11136-004-0027-4>

12. ANEXOS

12.1. Anexo 1: Fugl-Meyer Assessment de Miembro Superior

PROTOCOLO FMA -ES

Traducción de la versión original sueca del FMA, Universidad de Gotemburgo, Suecia
www.neurophys.gu.se/sektioner/klinisk-neurovetenskap/forskning/rehab_med/fugl-meyer

VALORACIÓN DE FUGL-MEYER

EXTREMIDAD SUPERIOR (FMA-ES)

Valoración de la función sensoriomotora

Identificación:

Fecha:

Examinador:

Fugl-Meyer AR, Jääskö L, Leyman I, Olsson S, Steglind S. The post-stroke hemiplegic patient. A method for evaluation of physical performance. *Scandinavian Journal of Rehabilitation Medicine* 1975, 7:13-31.

A. EXTREMIDAD SUPERIOR, posición sedente					
I. Actividad refleja			ning.	puede ser provocada	
Flexores: Bíceps y flexores de los dedos (al menos uno)			0	2	
Extensores: Tríceps			0	2	
Subtotal I (máx. 4)					
II. Movimiento voluntario dentro de sinergias, sin ayuda gravitacional			ning.	parcial	total
Sinergia flexora: Mano desde rodilla contralateral hasta oído ipsilateral. Desde la sinergia extensora (aducción de hombro/rotación interna, extensión del codo, pronación del antebrazo) hasta la sinergia flexora (abducción del hombro/rotación externa, flexión del codo, supinación del antebrazo).	Hombro	Retracción	0	1	2
		Elevación	0	1	2
		Abducción (90°)	0	1	2
	Codo	Rotación externa	0	1	2
		Flexión	0	1	2
		Antebrazo	Supinación	0	1
Sinergia extensora: Mano desde el oído ipsilateral hasta la rodilla contralateral	Hombro	Aducción/rotac. inter	0	1	2
	Codo	Extensión	0	1	2
	Antebrazo	Pronación	0	1	2
Subtotal II (máx. 18)					
III. Movimiento voluntario mezclando sinergias, sin compensación			ning.	parcial	total
Mano hasta la columna lumbar	No puede realizar, mano en frente a espina iliaca antero-superior		0	1	2
Mano sobre regazo	Mano detrás de espina iliaca antero-superior (sin compensación)				
	Mano hasta la columna lumbar (sin compensación)				
Flexión de hombro 0°-90°	Abducción inmediata o flexión de codo		0	1	2
Codo a 0°	Abducción o flexión de codo durante movimiento				
Pronación-supinación 0°	90° de flexión, no abducción de hombro ni flexión de codo				
Pronación-supinación	No pronación/supinación, imposible posición inicio		0	1	2
Codo a 90°	Pronación/supinación limitada, mantiene posición de inicio				
Hombro a 0°	Pronación/supinación completa, mantiene posición de inicio				
Subtotal III (máx. 6)					
IV. Movimiento voluntario con poca o ninguna sinergia			ning.	parcial	total
Abducción de hombro 0°-90°	Supinación inmediata o flexión de codo		0	1	2
Codo a 0°	Supinación o flexión de codo durante movimiento				
Antebrazo pronado	90° de abducción, mantiene extensión y pronación				
Flexión de hombro 90°-180°	Abducción inmediata o flexión de codo		0	1	2
Codo a 0°	Abducción o flexión de codo durante movimiento				
Pronación-supinación 0°	Flexión de 180°, no abducción de hombro o flexión de codo				
Pronación/supinación	No pronación/supinación, imposible posición inicio		0	1	2
Codo a 0°	Pronación/supinación limitada, mantiene posición de inicio				
Hombro a flexión de 30°-90°	Pronación/supinación completa, mantiene posición de inicio				
Subtotal IV (máx. 6)					
V. Actividad refleja normal evaluada solo si se logra puntaje total de 6 en parte IV					
Bíceps, Tríceps, Flexores de dedos	0 puntos en parte IV o 2 de 3 reflejos marcadamente hiperactivos		0	1	2
	1 reflejo marcadamente hiperactivo o al menos 2 reflejos enérgicos				
	Máximo de 1 reflejo enérgico, ninguno hiperactivo				
Subtotal V (máx. 2)					
Total A. EXTREMIDAD SUPERIOR (máx. 36)					

Version española: Hospital Militar Central, Universidad Nacional de Colombia
Barbosa NE, Forero SM, Galeano CP, Hernández ED, Landinez NS

2017-03-04

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

PROTOCOLO FMA -ES

Traducción de la versión original sueca del FMA, Universidad de Gotemburgo, Suecia
www.neurophys.gu.se/sektioner/klinisk-neurovetenskap/forskning/rehab_med/fugl-meyer

B. MUÑECA se puede dar apoyo en el codo para adoptar o mantener la posición, no apoyo en muñeca, verifique rango pasivo de movimiento antes de realizar prueba		ning.	parcial	total
Estabilidad a flexión dorsal de 15° Codo a 90°, antebrazo pronado Hombro a 0°	Flexión dorsal activa menor de 15° 15° de Flexión dorsal, no tolera resistencia Mantiene flexión dorsal contra resistencia	0	1	2
Flexión dorsal/volar repetida Codo a 90°, antebrazo pronado Hombro a 0° leve (flexión de los dedos)	No puede realizar voluntariamente Rango de movimiento activo limitado Rango de movimiento activo completo, fluido	0	1	2
Estabilidad a flexión dorsal de 15° Codo a 0°, antebrazo pronado Leve flexión/abducción de hombro	Flexión dorsal activa menor de 15° 15° de flexión dorsal, sin resistencia Mantiene posición contra resistencia	0	1	2
Flexión dorsal/volar repetida Codo a 0°, antebrazo pronado Leve flexión/abducción de hombro	No puede realizar voluntariamente Rango de movimiento activo limitado Rango de movimiento activo completo, fluido	0	1	2
Circunducción Codo a 90°, antebrazo pronado, hombro a 0°	No puede realizar voluntariamente Movimiento brusco o incompleto Circunducción completa y suave	0	1	2
Total B (máx. 10)				
C. MANO se puede dar apoyo en el codo para mantener flexión de 90°, no apoyo en la muñeca, compare con mano no afectada, los objetos están interpuestos, agarre activo		ning.	parcial	total
Flexión en masa	Desde extensión total activa o pasiva	0	1	2
Extensión en masa	Desde flexión total activa o pasiva	0	1	2
AGARRE				
a. Agarre de gancho flexión en IFP y IFD (dígitos II – V) Extensión en MCF II-V	No puede realizar Puede mantener posición pero débil Mantiene posición contra resistencia	0	1	2
b. Aducción de pulgar 1er CMC, MCF, IFP a 0°, trozo de papel Entre pulgar y 2da articulación MCF	No puede realizar Puede sostener papel pero no contra tirón Puede sostener papel contra tirón	0	1	2
c. Agarre tipo pinza, oposición Pulpejo del pulgar, contra pulpejo del 2 do dedo, se tira o hala el lápiz hacia arriba	No puede realizar Puede sostener lápiz pero no contra tirón Puede sostener lápiz contra tirón	0	1	2
d. Agarre cilíndrico Objeto en forma cilíndrica (pequeña lata) Se tira o hala hacia arriba con oposición en dígitos I y II	No puede realizar Puede sostener cilindro pero no contra tirón Puede sostener cilindro contra tirón	0	1	2
e. Agarre esférico Dedos en abducción/flexión, pulgar opuesto, bola de tenis	No puede realizar Puede sostener bola pero no contra tirón Puede sostener bola contra tirón	0	1	2
Total C (máx. 14)				
D. COORDINACIÓN/VELOCIDAD después de una prueba con ambos brazos, con los ojos vendados, punta del dedo índice desde la rodilla hasta la nariz, 5 veces tan rápido como sea posible		marcado	leve	ninguno
Temblor	Al menos 1 movimiento completo	0	1	2
Dismetria	Pronunciada o asistemática Leve y sistemática No dismetria	0	1	2
		> 6s	2 - 5s	< 2s
Tiempo Inicio y final con la mano sobre la rodilla	Al menos 6 seg. más lento que el lado no afectado 2-5 seg. más lento que el lado no afectado Menos de 2 segundos de diferencia	0	1	2
Total D (máx. 6)				
Total A-D (máx.6)				

Version española: Hospital Militar Central, Universidad Nacional de Colombia
Barbosa NE, Forero SM, Galeano CP, Hernández ED, Landinez NS

2017-03-04

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

PROTOCOLO FMA -ES

Traducción de la versión original sueca del FMA, Universidad de Gotemburgo, Suecia
www.neurophys.gu.se/sektioner/klinisk-neurovetenskap/forskning/rehab_med/fugl-meyer

H. SENSACIÓN, extremidad superior con los ojos vendados, comparado con el lado no afectado		anestesia	hipoestesia disestesia	normal
Tacto Suave	Brazo, antebrazo, superficie palmar de mano	0 0	1 1	2 2
		ausencia menos de ¾ correcto	¾ correcto considerable diferencia	correcto 100% poca o no diferencia
Posición Pequeña alteración en la posición	Hombro	0	1	2
	Codo	0	1	2
	Muñeca	0	1	2
	Pulgar (articulación - IF)	0	1	2
Total H. (máx. 12)				

I. MOVIMIENTO ARTICULAR PASIVO, extremidad superior				J. DOLOR ARTICULAR durante movimiento pasivo, extremidad superior		
Posición sedente, compare con lado no afectado	solo pocos grados (menos de 10° en hombro)	disminuido	normal	dolor constante pronunciado durante o al final del movimiento o dolor muy marcado al final del movimiento	algún dolor	no dolor
Hombro						
Flexión (0°-180°)	0	1	2	0	1	2
Abducción (0°-90°)	0	1	2	0	1	2
Rotación externa	0	1	2	0	1	2
Rotación interna	0	1	2	0	1	2
Codo						
Flexión	0	1	2	0	1	2
Extensión	0	1	2	0	1	2
Antebrazo						
Pronación	0	1	2	0	1	2
Supinación	0	1	2	0	1	2
Muñeca						
Flexión	0	1	2	0	1	2
Extensión	0	1	2	0	1	2
Dedos						
Flexión	0	1	2	0	1	2
Extensión	0	1	2	0	1	2
Total I (máx. 24)				Total J(max. 24)		

A. EXTREMIDAD SUPERIOR	/36
B. MUÑECA	/10
C. MANO	/14
D. COORDINACIÓN/VELOCIDAD	/6
TOTAL A - D (función motora)	/66

H. SENSACION	/12
I. MOVIMIENTO ARTICULAR PASIVO	/24
J. DOLOR ARTICULAR	/24

Version española: Hospital Militar Central, Universidad Nacional de Colombia
Barbosa NE, Forero SM, Galeano CP, Hernández ED, Landinez NS

2017-03-04

12.2. Anexo 2: Fugl-Meyer Assessment de Miembro Inferior

PROTOCOLO FMA -EI

Traducción de la versión original sueca del FMA, Universidad de Gotemburgo, Suecia
www.neurophys.gu.se/sektioner/klinisk-neurovetenskap/forskning/rehab_med/fugl-meyer

VALORACIÓN DE FUGL-MEYER EXTREMIDAD INFERIOR (FMA-EI)

Identificación:

Fecha:

Valoración de la función sensoriomotora Examinador:

Fugl-Meyer AR, Jaaskö L, Leyman I, Olsson S, Steglind S. The post-stroke hemiplegic patient. A method for evaluation of physical performance. Scandinavian Journal of Rehabilitation Medicine 1975, 7:13-31.

E. EXTREMIDAD INFERIOR					
I. Actividad refleja, posición supina			ning.	puede ser provocada	
Flexores: Flexores de rodilla			0	2	
Extensores: Reflejo Patelar y Aquiliano (al menos uno)			0	2	
Subtotal I (máx. 4)					
II. Movimiento voluntario dentro de sinergias, posición supina			ning.	parcial	total
Sinergia flexora: Flexión de cadera máxima (abducción/rotación externa), máxima flexión en rodilla y articulación de tobillo (palpar tendones distales para asegurar flexión activa de rodilla)	Cadera	Flexión	0	1	2
	Rodilla	Flexión	0	1	2
	Tobillo	Flexión dorsal	0	1	2
Sinergia extensora: Desde la sinergia flexora hasta la aducción/ extensión de la cadera, extensión de la rodilla y flexión plantar de tobillo. Se aplica resistencia para asegurar movimiento activo, evaluar movimiento y fortaleza (compare con el lado no afectado)	Cadera	Extensión	0	1	2
	Rodilla	Aducción	0	1	2
	Tobillo	Extensión	0	1	2
Subtotal II (máx. 14)					
III. Movimiento voluntario mezclado con sinergias, posición sentado, rodilla a 10 cm del borde de la silla/cama			ning.	parcial	total
Flexión de rodilla desde rodilla extendida activa o pasivamente	No movimiento activo		0	1	2
	Flexión no activa menor de 90°, palpar tendones isquiotibiales				
Flexión dorsal de tobillo	Flexión activa más de 90°		0	1	2
	Comparar con lado no afectado				
Subtotal III (máx. 4)					
IV. Movimiento voluntario con poca o ninguna sinergia, posición de pie, cadera a 0°			ning.	parcial	total
Flexión de rodilla a 90°	Movimiento no activo o inmediato, flexión de cadera simultánea		0	1	2
	Cadera a 0°, puede sostenerse para equilibrio				
Flexión dorsal de tobillo	Flexión de rodilla de al menos 90° o flexión de cadera durante movimiento		0	1	2
	Flexión de rodilla de al menos 90° sin flexión de cadera simultánea.				
Comparar con lado no afectado	No movimiento activo		0	1	2
	Flexión dorsal limitada				
Subtotal IV (máx.4)					
V. Actividad refleja normal posición supina, se evalúa solo si se logra el puntaje total de 4 puntos en la primera parte IV, compare con lado no afectado			ning.	parcial	total
Actividad refleja	0 puntos en parte IV o 2 de 3 reflejos marcadamente hiperactivos		0	1	2
	Flexores de rodilla, tendón Aquiliano y Patelar				
Subtotal V (máx. 2)					
Total E: EXTREMIDAD INFERIOR (máx. 28)					

Version española: Hospital Militar Central, Universidad Nacional de Colombia
Barbosa NE, Forero SM, Galeano CP, Hernández ED, Landínez NS

2017-03-04

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

PROTOCOLO FMA -EI

Traducción de la versión original sueca del FMA, Universidad de Gotemburgo, Suecia
www.neurophys.gu.se/sektioner/klinisk-neurovetenskap/forskning/rehab_med/fugl-meyer

F. COORDINACIÓN/ VELOCIDAD posición supina, después de una prueba con ambas piernas, con los ojos vendados, talón a la patela de la pierna opuesta, 5 veces tan rápido como sea posible.		marcado	leve	ninguno
Temblor	Al menos 1 movimiento completo	0	1	2
Dismetria	Pronunciada o asistemática	0		
	Leve y sistemática		1	
	No dismetria			2
		>6s	2-5s	<2s
Tiempo	Al menos 6 seg. más lento que el lado no afectado	0		
	2-5 seg. más lento que el lado no afectado		1	
	Menos de 2 seg. de diferencia			2
Total F (máx. 6)				

H. SENSACIÓN , extremidad inferior, ojos vendados, compare con el lado no afectado		anestesia	hipoestesia Disestesia	normal
Tacto Suave	Pierna	0	1	2
	Planta del pie	0	1	2
		menos de ¼ correcto o Ausencia	¼ correcto o considerable diferencia	correcto 100% poca o ninguna diferencia
Posición Pequeña alteración en la posición	Cadera	0	1	2
	Rodilla	0	1	2
	Tobillo	0	1	2
	Dedo gordo del pie (articulación - IF)	0	1	2
Total H. (máx. 12)				

I. MOVIMIENTO ARTICULAR PASIVO, extremidad inferior				J. DOLOR ARTICULAR durante movimiento pasivo, extremidad inferior			
compare con lado no afectado		solo pocos grados	disminuido	normal	dolor severo durante el movimiento o dolor muy marcado al final del movimiento	algún dolor	no dolor
Cadera	Flexión	0	1	2	0	1	2
	Abducción	0	1	2	0	1	2
	Rotación externa	0	1	2	0	1	2
	Rotación interna	0	1	2	0	1	2
Rodilla	Flexión	0	1	2	0	1	2
	Extensión	0	1	2	0	1	2
Tobillo	Flexión dorsal	0	1	2	0	1	2
	Flexión plantar	0	1	2	0	1	2
Pie	Pronación	0	1	2	0	1	2
	supinación	0	1	2	0	1	2
Total (máx. 20)					Total (máx. 20)		

E. EXTREMIDAD INFERIOR	/28
F. COORDINACIÓN/ VELOCIDAD	/6
TOTAL E-F (función motora)	/34

H. SENSACION	/12
I. MOVIMIENTO ARTICULAR PASIVO	/20
J. DOLOR ARTICULAR	/20

Version española: Hospital Militar Central, Universidad Nacional de Colombia
Barbosa NE, Forero SM, Galeano CP, Hernández ED, Landinez NS

2017-03-04

12.3. Anexo 3: NSDA-R Instructions for the therapist

NOTTINGHAM STROKE DRESSING ASSESSMENT-REVISED (NSDA-R) GENERAL INSTRUCTIONS FOR THE THERAPIST

Area of assessment: PADL/dressing.

Content and purpose:

a) **Nottingham Stroke Dressing Assessment-Revised (NSDA-R)** - evaluates the independence level in performing dressing tasks and identifies problem areas.

b) **Action Error Analysis-Revised** - identifies action errors due to cognitive impairment that interfere with dressing performance.

Diagnosis: People with stroke in the acute, subacute, or chronic phase.

Age: Suitable for all ages.

Required training: Read the instructions carefully.

How to administer the instrument:

- Each NSDA-R item is scored on a 3-point ordinal scale ranging from 0 to 2 points:
 - 2 points: The person performs the action independently.
 - 1 point: The person is exclusively dependent on hands-off assistance (for example: the therapist provides verbal step-by-step instructions, encourages, points to a garment, or demonstrates how an action must be performed).
 - 0 points: The person is dependent on hands-on assistance (for example: the therapist provides hand-over-hand guidance or takes over the task). Hands-off assistance may additionally be needed.
 - N/A: Not applicable (for example: the person does not use that garment, or the action was not observed).
- All NSDA-R items of relevance to the person should be attempted.
- The person may use assistive devices and/or other compensatory strategies that are familiar to them. If an action is performed using assistive devices and/or other compensatory strategies, but without any type of assistance from the therapist, a score of 2 points is given (independent).
- If the assessment is conducted in hospital, use the person's own clothes.
- Place the garments randomly within the person's vision field and within easy reach. If the placement of garments is part of a compensatory strategy the person follows, place the garments accordingly.
- When the objective of the assessment is to evaluate changes in independence over time, the same garments as in the initial evaluation must be used.
- Starting position: The person can choose to be sitting or standing. An upright chair with armrests (for example, dining chair, wheelchair, bath

chair, bath stool) should be available for the person in the assessment area. Example: If the person, despite impaired balance, attempts to put on shoes in standing position and requires assistance from the therapist, provide the necessary assistance and score accordingly.

- Score only what the person does, not what you think the person is able to do.
- Provide hands-off assistance as the first option, if possible. If hands-off assistance is not sufficient, move on to hands-on assistance.
- If the person continues having difficulties after several attempts, is visibly distressed, or performs the task unacceptably slowly (≥ 5 min.), provide the necessary level of assistance, and score accordingly.
- If the person is independent in putting on the garment, but requires assistance to adjust it, score accordingly. Example: If the person manages to put on a sock, but the heel is on the front side of the foot and hands-on assistance is required to adjust the sock, a score of 0 points is given.
- Complete the Action Error Analysis-Revised based on the observations of the NSDA-R items.

Calculation of the NSDA-R score:

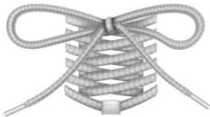
The NSDA-R score is a percentage (0-100%) and it is calculated by dividing the total score achieved for the attempted items by the maximum possible score for those items (number of attempted items x 2 maximum points). N/A items and the Action Error Analysis-Revised are not included in the score calculation. Example: If the person attempted 15 items and got a total score of 20 points for those observed items, then the NSDA-R score is: $20 / (15 \times 2 \text{ maximum points}) = 0.66 = 66\%$.

12.4. Anexo 4: NSDA-R Scale

NO.	ITEMS	INSTRUCTIONS FOR THE THERAPIST	SCORE (2, 1, 0) or N/A	COMMENTS (for example: starting position, assistive devices, other compensatory strategies, or action errors)
1.	FASTENINGS/FASTENERS	Starting position: The person can choose to be sitting or standing. An upright chair with armrests (for example: dining chair, wheelchair, bath chair, or bath stool) should be available for the person in the assessment area. Assess only fastenings relevant for the person in question. If the person uses the same type of fastening on different garments, register the lowest score. Example: If the person can button small buttons on a shirt, but not on trousers, then register the score for the buttons on the trousers.		
1.1	Zip 	2: The person must be able to raise/lower the zip; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
1.2	Large button 	2: The person must be able to button/unbutton the button; no assistance needed. 1: Hands-off assistance needed.		

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

		0: Hands-on assistance needed.		
1.3	Small button 	2: The person must be able to button/unbutton the button; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
1.4	Velcro 	2: The person must be able to fasten/unfasten the velcro; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
1.5	Hook and eye/clasp 	2: The person must be able to hook/unhook the hook and eye; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
1.6	Press stud/popper 	2: The person must be able to fasten/unfasten the press stud; no assistance needed. 1: Hands-off assistance needed.		

		0: Hands-on assistance needed.		
1.7	Bow/knot 	2: The person must be able to tie/untie the bow; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
1.8	Buckle 	2: The person must be able to buckle/unbuckle the buckle; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
1.9	Trouser clip/trouser clasp 	2: The person must be able to hook/unhook the trouser clip; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
2.	BRA 	Starting position: The person can choose to be sitting or standing. An upright chair with armrests (for example: dining chair, wheelchair, bath chair, or bath stool) should be available for the person in the assessment area. Assess only the dressing strategy (2a, 2b or 2c) chosen by the person in question. The use of fastenings is scored in section 1.		

				
2a	Dressing strategy 2a: the person puts on the bra unfastened, with fastening on the back.			
2a.1	Selecting the correct opening and strap	2: The person must be able to select and put both arms through the correct openings and strap; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
2a.2	Adjusting	2: The person must be able to adjust the position of the bra; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
2b	Dressing strategy 2b: the person puts on the bra unfastened, with fastening in the front.			
2b.1	Placing bra around the upper body	2: The persons must be able to place the bra around the upper body with fastening in the front; no assistance		

		needde. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
2b.2	Turning the bra around the body	2: The person must be able to turn the bra around the body so that the fastening is on the back; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
2b.3	Raising the straps to the shoulders and adjusting.	2: The person must be able to raise the straps to shoulders and adjust the position of the bra; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
2c	Dressing strategy 2c: the person puts on a bra without fastening or an already fastened bra.			
2c.1	Selecting the correct opening and strap	2: The person must be able to select and put both arms through the correct openings and straps; no assistance needed.		

		1: Hands-off assistance needed. 0: Hands-on assistance needed.		
2c.2	Pulling over the head	2: The person must be able to push the bra over the elbows and pull it over the head; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
2c.3	Pulling down and adjusting	2: The person must be able to pull down the bra and adjust its position; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
3	TANK TOP, SLIP 	Starting position: The person can choose to be sitting or standing. An upright chair with armrests (for example: dining chair, wheelchair, bath chair, or bath stool) should be available for the person in the assessment area.		
3.1	Selecting the correct opening	2: The person must be able to select and put both arms through the correct openings; no assistance		

		needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
3.2	Pulling over head	2: The person must be able to push the garment over elbows and pull it over head; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
3.3	Pulling down and adjusting	2: The person must be able to pull down the garment and adjust it; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
4.	T-SHIRT, SHORT-SLEEVED JUMPER 	Starting position: The person can choose to be sitting or standing. An upright chair with armrests (for example: dining chair, wheelchair, bath chair, or bath stool) should be available for the person in the assessment area. The use of fastenings is scored in section 1.		
4.1	Selecting the correct opening	2: The person must be able to select the correct openings and introduce		

		both arms through the corresponding sleeves; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
4.2	Pulling over head	2: The person must be able to push the garment over elbows and pull it over head; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
4.3	Pulling down and adjusting	2: The person must be able to pull down the garment and adjust its position no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
5.	DRESS 	Starting position: The person can choose to be sitting or standing. An upright chair with armrests (for example: dining chair, wheelchair, bath chair, or bath stool) should be available for the person in the assessment area. The use of fastenings is scored in section 1.		
5.1	Selecting the correct opening	2: The person must be able to select the correct openings and introduce		

		both arms through the corresponding sleeves/openings; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
5.2	Pulling over head.	2: The person must be able to push the dress over elbows and pull it over the head; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
5.3	Pulling down and adjusting	2: The person must be able to pull down the dress and adjust its position; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
6.	SHIRT, CARDIGAN, JACKET 	Starting position: The person can choose to be sitting or standing. An upright chair with armrests (for example: dining chair, wheelchair, bath chair, or bath stool) should be available for the person in the assessment area. The use of fastenings is scored in section 1.		

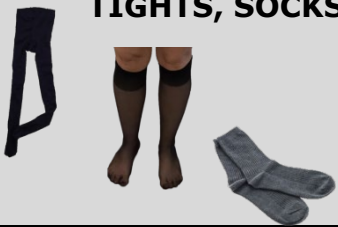
6.1	Selecting correct opening	2: The person must be able to select the correct openings and introduce both arms through the corresponding sleeves; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
6.2	Pulling the garment around the back, around the back at waist height, or over the head	2: After putting in the first arm, the person must be able to pull the garment around the back (at waist height as an option) or over the head to introduce the other arm; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
6.3	Pulling down and adjusting	2: The person must be able to push the garment over shoulders, pull down to their waist and adjust its position; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		

7.	LONG-SLEEVED JUMPER 	Starting position: The person can choose to be sitting or standing. An upright chair with armrests (for example: dining chair, wheelchair, bath chair, or bath stool) should be available for the person in the assessment area. The use of fastenings is scored in section 1.		
7.1	Selecting correct opening	2: The person must be able to select the correct openings and introduce both arms through the corresponding sleeves; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
7.2	Pulling over head.	2: The person must be able to push the jumper over elbows and pull it over the head; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
7.3	Pulling down and adjusting.	2: The person must be able to pull down the jumper and adjust its position; no assistance needed. 1: Hands-off assistance needed.		

		0: Hands-on assistance needed.		
8.	SKIRT, WAIST SLIP 	Starting position: The person can choose to be sitting or standing. An upright chair with armrests (for example: dining chair, wheelchair, bath chair, or bath stool) should be available for the person in the assessment area. Assess only the dressing strategy (8A or 8B) chosen by the person in question. The use of fastenings is scored in section 1.		
8a	Dressing strategy 8a: the person introduces the legs through the waist band.			
8a.1	Putting legs through the skirt	2: The person must be able to put both legs through waist band; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
8a.2	Pulling up and adjusting	2: The person must be able to pull up the garment to the waist and adjust its position; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
8b	Dressing strategy 8b: the person introduces the arms through the skirt.			
8b.1	Putting arms through the skirt	2: The person must be able to put both arms through correct opening;		

		no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
8b.2	Pulling over head	2: The person must be able to pull the garment over the head; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
8b.3	Pulling down and adjusting	2: The person must be able to pull down the garment to the waist and adjust its position; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
9.	PANTS/PANTIES/UNDERPANTS 	Starting position: The person can choose to be sitting or standing. An upright chair with armrests (for example: dining chair, wheelchair, bath chair, or bath stool) should be available for the person in the assessment area.		
9.1	Selecting correct	2: The person must be able to		

	opening	introduce both legs through the corresponding openings; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
9.2	Pulling down and adjusting.	2: The person must be able to pull garment up fully to cover the bottom and adjust the garment's position; no assistance needed 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
10.	TROUSERS, LEGGINGS 	Starting position: The person can choose to be sitting or standing. An upright chair with armrests (for example: dining chair, wheelchair, bath chair, or bath stool) should be available for the person in the assessment area. The use of fastenings is scored in section 1.		
10.1	Selecting the correct opening	2: The person must be able to select the correct opening of the garment		

		and put the legs through it; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
10.2	Pulling up and adjusting	2: The person must be able to pull up garment to fully cover the bottom and adjust the garment's position; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
11.	STOCKINGS, TIGHTS, SOCKS 	Starting position: The person can choose to be sitting or standing. An upright chair with armrests (for example: dining chair, wheelchair, bath chair, or bath stool) should be available for the person in the assessment area.		
11.1	Pulling up over toes/heels and adjusting	2: In the case of tights, the person must be able to select the correct opening. For all garments, the		

		person must be able to place the garment over the toes and heels and adjust the garment's position; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
11.2	Pulling up and adjusting.	2: The person must be able to pull up the garment (tights must cover the bottom fully) and adjust its position; no assistance needed. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
12.	SHOES 	Starting position: The person can choose to be sitting or standing. An upright chair with armrests (for example: dining chair, wheelchair, bath chair, or bath stool) should be available for the person in the assessment area.		
12.1	Selecting the correct shoe	2: The person must be able to select the correct shoe for each foot; no assistance needed.		

		1: Hands-off assistance needed. 0: Hands-on assistance needed.		
12.2	Put on the shoes	2: The person must be able to put the shoes on both feet; no assistance needed. The foot should be fitted comfortably in the shoe. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
12.3	Lacing/fastening shoes	2: The person must be able to tie the laces/fasten both shoes; no assistance needed. It may include different strategies and fastenings like tying them with one hand, elastic laces, or velcro. 1: Hands-off assistance needed. 0: Hands-on assistance needed.		
CALCULATION OF NSDA-R SCORE				
Total score achieved for the attempted items (a):				
Number of attempted items (b):				

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

NSDA-R score: $a / (b \times 2)$	

12.5 Anexo 5: NSDA-R Action error analysis

ACTION ERROR ANALYSIS-REVISED

NO.	TYPE OF ACTION ERROR	EXAMPLES	TICK CORRESPONDING BOX
1	Problems with initiation	Shows difficulties in starting to get dressed or moving on to the next garment.	⊗ Action error present ⊗ Action error absent
2	Problems with sustaining attention	Distracted. Does not concentrate (the sufficient time) on the dressing task.	⊗ Action error present ⊗ Action error absent
3	Mental slowness	Completes the dressing task, but the performance is slow and requires additional time (for example, prolonged searching for the garment openings). Exclude motor and sensory problems.	⊗ Action error present ⊗ Action error absent
4	Problems with locating objects in parts of the visual field	Cannot find garments placed to the left/right.	⊗ Action error present ⊗ Action error absent
5	Inattention to one side of the body	Does not complete the dressing on the left/right side of the body.	⊗ Action error present ⊗ Action error absent
6	Problems with ordering/sequencing	Performs steps in the wrong order (for example, puts on shoes before socks).	⊗ Action error present ⊗ Action error absent
7	Perseverance	Repeats an already completed step.	⊗ Action error present ⊗ Action error absent
8	Problems with selecting correct objects	Selects the incorrect garment (for example, takes the trousers when they are dressing the upper body).	⊗ Action error present ⊗ Action error absent
9	Disorganisation, lack of clear strategy	Shows unmistakable signs of not knowing what to do with a garment (for example, handles a shirt without seeming to know what to do with it).	⊗ Action error present ⊗ Action error absent
10	Impaired self-monitoring	Problem with monitoring own task performance and correcting errors which occur.	⊗ Action error present ⊗ Action error absent
11	Spatial confusion	Cannot find or select the correct garment openings. Puts the clothes on backwards. Puts the back part of a garment on the front of the body	⊗ Action error present ⊗ Action error absent

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

		(with the neck label in the front).	
12	Clumsiness	Difficulties to grasp the garment or to manage fastenings. Exclude motor and sensory problems.	☐ Action error present ☐ Action error absent
ADDITIONAL OBSERVATIONS:			

12.6 Anexo 6: ENVI-R Instrucciones para el terapeuta

ESCALA NOTTINGHAM DEL VESTIDO EN ICTUS-REVISADA (ENVI-R)

INSTRUCCIONES GENERALES PARA EL TERAPEUTA

Área de evaluación: AVD Básica/Vestido.

Contenido y objetivo:

a) **Escala Nottingham de Vestido en Ictus-Revisada (ENVI-R)** - evalúa el nivel de independencia en el desempeño de la actividad de vestido e identifica áreas problemáticas.

b) **Análisis de los Errores de Acción-Revisado** - identifica errores en las acciones debidos a déficits cognitivos que interfieren a la hora de vestirse.

Diagnóstico: Personas con ictus en cualquier fase de evolución.

Edad: Aplicable a cualquier edad.

Formación requerida: Lea las instrucciones detenidamente.

Cómo administrar la escala:

- Cada ítem de *ENVI-R* se puntúa en una escala de 3 puntos, en un rango de 0 a 2 puntos:
 - 2 puntos: La persona realiza la acción de manera independiente.
 - 1 punto: La persona sólo depende de otro tipo de asistencia no física (por ejemplo: el terapeuta da instrucciones verbales, anima, señala una prenda o demuestra cómo se debe realizar una acción).
 - 0 puntos: La persona necesita asistencia física (por ejemplo: el terapeuta le guía con las manos o realiza una acción por la persona). También se proporciona otro tipo de asistencia no física.
 - N/A: No aplicable (por ejemplo: la persona no usa esa prenda, o la acción no está presente y no se puede observar).
- Se deben evaluar todas las prendas que sean significativas para la persona.
- La persona puede utilizar cualquier producto de apoyo u otra estrategia compensatoria que use de manera cotidiana y que le sean familiares. Si la acción es llevada a cabo usando un producto de apoyo u otra estrategia compensatoria, pero sin ningún tipo de asistencia del terapeuta, la puntuación obtenida será de 2 puntos (independiente).
- Si la evaluación se lleva a cabo en el hospital, usar la ropa propia de la persona.
- Coloque las prendas de forma aleatoria dentro del campo de visión de la persona y al alcance de su mano. Si la colocación de las prendas es parte de una estrategia compensatoria de la persona, dispóngalas de esa forma.

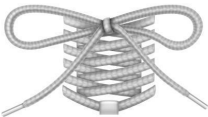
- Cuando el objetivo de la evaluación sea documentar cambios en la independencia a lo largo del tiempo, se deben usar las mismas prendas que en la evaluación inicial.
- Posición inicial: La persona puede elegir estar sentada o de pie. Debe haber un asiento con reposabrazos disponible para la persona en el lugar de la evaluación (por ejemplo: una silla de comedor, una silla de ruedas, una silla de baño, taburete de baño). Por ejemplo, si la persona intenta ponerse los zapatos de pie a pesar de existir alteraciones de equilibrio, proporcione la asistencia necesaria y puntúe acorde a ello.
- Puntúe sólo lo que la persona hace, no lo que usted crea que la persona es capaz de hacer.
- Proporcione otro tipo de asistencia (no física) como primera opción si es posible. Si esta asistencia es insuficiente, ofrezca asistencia física.
- Si la persona sigue teniendo dificultades tras varios intentos, está notablemente angustiada/agobiada/estresada, o el desempeño de la tarea es excesivamente lento (≥ 5 min.), proporcione la asistencia necesaria y puntúe acorde a ello.
- Si la persona se pone la prenda de manera independiente, pero requiere asistencia para colocarla correctamente, puntúe acorde a ello. Por ejemplo: si la persona logra ponerse el calcetín, pero el talón del calcetín está en la parte delantera del pie y se proporciona asistencia física, se puntúa como 0 puntos.
- Si un ítem aplica a varias prendas, se evaluará en las diferentes ocasiones de aparición y se registrarán las puntuaciones acordes. Por ejemplo, puede aparecer el lazo/nudo en la cintura del pantalón y también en las zapatillas, por lo que el ítem de lazo/nudo debe evaluarse 2 veces, ya tenga la misma puntuación, o puntuaciones diferentes; y registrar ambas puntuaciones. Lo mismo puede ocurrir con la camisa y la chaqueta, porque ambas se evalúan dentro de la categoría 6 (Camisa/Chaqueta/Abrigo), y se puntuará el desempeño del vestido en ambas prendas.
- Complete el documento *Análisis de los Errores de Acción-Revisado* basándose en las observaciones de los ítems de *ENVI-R*.

Cálculo de la puntuación de ENVI-R:

La puntuación de la escala *ENVI-R* es un porcentaje (0-100%) que se obtiene dividiendo la puntuación total lograda en los ítems intentados, entre la puntuación máxima posible de esos ítems (número de ítems que se han intentado x2 puntos máximo). Los ítems N/A y el *Análisis de los Errores de Acción-Revisado* no se incluyen en el cálculo de la puntuación. Ejemplo: Si la persona intentó 15 ítems y obtuvo una puntuación de 20 puntos en estos ítems observados, la puntuación de *ENVI-R* es: $20 / (15 \times 2 \text{ puntos máximo}) = 0.66 = 66\%$.

12.7. Anexo 7: ENVI-R Escala

NO.	ITEMS	INSTRUCCIONES PARA EL TERAPEUTA	PUNTUACIÓN (2, 1, 0) o N/A	COMENTARIOS (por ejemplo: posición inicial, dispositivos de asistencia, estrategias compensatorias, o errores de acción)
1.	CIERRES	Posición inicial: La persona puede elegir estar sentado o de pie. Debe haber un asiento con reposabrazos disponible para la persona en el lugar de la evaluación (por ejemplo: una silla de comedor, una silla de ruedas, una silla de baño, un taburete de baño). Evalúe sólo los cierres relevantes para la persona. Si la persona utiliza el mismo cierre en diferentes prendas, registre la puntuación más baja (por ejemplo: si la persona puede abrochar botones pequeños en una camisa, pero no en los pantalones, registre la puntuación de los botones del pantalón).		
1.1	Cremallera 	2: La persona sube/baja la cremallera sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
1.2	Botón grande	2: La persona abrocha/desabrocha el		

		botón sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
1.3	Botón pequeño 	2: La persona abrocha/desabrocha el botón sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
1.4	Corchete (de gancho y ojo) 	2: La persona engancha/desengancha el corchete sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
1.5	Lazo/nudo 	2: La persona hace/deshace el lazo/nudo sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		

1.6	Hebilla 	2: La persona abrocha/desabrocha la hebilla sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
1.7	Corchete de pantalón 	2: La persona engancha/desengancha el corchete sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
2.	SUJETADOR/TOP 	Posición inicial: La persona puede elegir estar sentado o de pie. Debe haber un asiento con reposabrazos disponible para la persona en el lugar de la evaluación (por ejemplo: una silla de comedor, una silla de ruedas, una silla de baño, un taburete de baño). Evalúe sólo la estrategia de vestido (2a, 2b o 2c) elegida por la persona. La utilización de los cierres se puntúa en la sección 1.		
2a	Estrategia de vestido 2a: la persona se pone el sujetador desabrochado, con el cierre en la espalda.			

2a.1	Seleccionar el orificio y el tirante correcto con ambos brazos.	2: La persona selecciona y pasa el brazo por el orificio y tirante correctos sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
2a.2	Colocar correctamente.	2: La persona se coloca correctamente el sujetador sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
2b	Estrategia de vestido 2b: la persona se pone el sujetador desabrochado, con el cierre por delante.			
2b.1	Posicionar el sujetador alrededor de la parte superior del cuerpo.	2: La persona se coloca el sujetador alrededor de la parte superior del cuerpo con el cierre por delante sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		

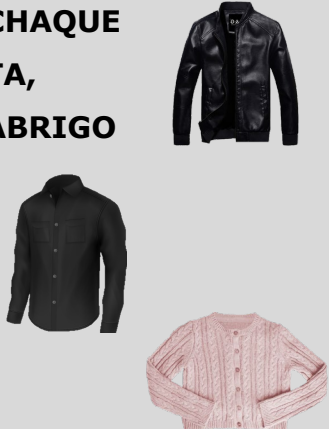
2b.2	Girar el sujetador, de manera que el cierre quede en la espalda.	2: La persona gira el sujetador alrededor del cuerpo, de manera que el cierre se queda en la espalda, sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
2b.3	Subir los tirantes hasta los hombros y colocar el sujetador correctamente.	2: La persona sube los tirantes hasta los hombros y se coloca el sujetador sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
2c	Estrategia de vestido 2c: la persona se pone un sujetador sin cierre o un sujetador ya abrochado.			
2c.1	Seleccionar el orificio y el tirante correcto con ambos brazos.	2: La persona selecciona y pasa el brazo por el orificio y tirante correctos con ambos brazos sin asistencia. 1: Necesita otro tipo de asistencia (no física).		

		0: Necesita asistencia física.		
2c.2	Pasar el sujetador por encima de la cabeza.	2: La persona sube los tirantes por encima de los codos y pasa el sujetador por encima de la cabeza sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
2c.3	Bajar el sujetador y colocarlo correctamente.	2: La persona baja el sujetador y lo coloca correctamente sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
3.	CAMISETA DE TIRANTES 	Posición inicial: La persona puede elegir estar sentado o de pie. Debe haber un asiento con reposabrazos disponible para la persona en el lugar de la evaluación (por ejemplo: una silla de comedor, una silla de ruedas, una silla de baño, un taburete de baño).		
3.1	Seleccionar el orificio correcto con ambos brazos.	2: La persona selecciona los orificios correctos e introduce los brazos por los orificios correspondientes sin		

		asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
3.2	Pasar la camiseta por encima de la cabeza.	2: La persona sube la prenda por encima de los codos y la pasa por encima de la cabeza sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
3.3	Bajar la camiseta y colocarla correctamente.	2: La persona se baja la prenda y se la coloca correctamente sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
4.	CAMISETA, JERSEY DE MANGA CORTA	Posición inicial: La persona puede elegir estar sentado o de pie. Debe haber un asiento con reposabrazos disponible para la persona en el lugar de la evaluación (por ejemplo: una silla de comedor, una silla de ruedas, una silla de baño, un taburete de baño). La utilización de los cierres se puntúa en la sección 1.		

				
4.1	Seleccionar el orificio correcto con ambos brazos.	2: La persona selecciona los orificios correctos e introduce los brazos por los orificios correspondientes sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
4.2	Pasar la camiseta/jersey por encima de la cabeza.	2: La persona sube la prenda por encima de los codos y la pasa por encima de la cabeza sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
4.3	Bajar la camiseta y colocarla correctamente.	2: La persona se baja la prenda y se la coloca correctamente sin asistencia. 1: Necesita otro tipo de asistencia		

		(no física). 0: Necesita asistencia física.		
5.	VESTIDO 	Posición inicial: La persona puede elegir estar sentado o de pie. Debe haber un asiento con reposabrazos disponible para la persona en el lugar de la evaluación (por ejemplo: una silla de comedor, una silla de ruedas, una silla de baño, un taburete de baño). La utilización de los cierres se puntúa en la sección 1.		
5.1	Seleccionar el orificio correcto con ambos brazos.	2: La persona selecciona los orificios correctos e introduce los brazos por los orificios correspondientes sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
5.2	Pasar el vestido por encima de la cabeza.	2: La persona sube la prenda por encima de los codos y la pasa por encima de la cabeza sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
5.3	Bajar el vestido y	2: La persona se baja la prenda y se		

	colocarlo correctamente.	la coloca correctamente sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
6.	CAMISA, CHAQUETA, ABRIGO 	Posición inicial: La persona puede elegir estar sentado o de pie. Debe haber un asiento con reposabrazos disponible para la persona en el lugar de la evaluación (por ejemplo: una silla de comedor, una silla de ruedas, una silla de baño, un taburete de baño). La utilización de los cierres se puntúa en la sección 1.		
6.1	Seleccionar el orificio correcto.	2: La persona selecciona el orificio correcto e introduce el brazo por el orificio correspondiente sin asistencia. 1: Necesita otro tipo de asistencia (no física).		

		0: Necesita asistencia física.		
6.2	Pasar la prenda por la espalda/por la espalda a la altura de la cintura/por encima de la cabeza.	2: Tras vestir el primer brazo, la persona se pasa la prenda por la espalda (a la altura de la cintura, o por encima de la cabeza) para colocarse la manga del otro brazo, sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
6.3	Bajar la prenda y colocarla correctamente.	2: La persona se coloca la prenda sobre los hombros, se baja la prenda hacia la cintura y se la coloca correctamente sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
7.	CAMISETA, JERSEY DE MANGA LARGA	Posición inicial: La persona puede elegir estar sentado o de pie. Debe haber un asiento con reposabrazos disponible para la persona en el lugar de la evaluación (por ejemplo: una silla de comedor, una silla de ruedas, una silla de baño, un taburete de baño). La utilización de los		

		cierres se puntúa en la sección 1.		
7.1	Seleccionar el orificio correcto con ambos brazos.	2: La persona selecciona los orificios correctos e introduce los brazos por los orificios correspondientes sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
7.2	Pasar el jersey por encima de la cabeza.	2: La persona sube la prenda por encima de los codos y la pasa por encima de la cabeza sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
7.3	Bajar el jersey y colocarlo correctamente.	2: La persona se baja la prenda y se la coloca correctamente sin asistencia. 1: Necesita otro tipo de asistencia		

		(no física). 0: Necesita asistencia física.		
8.	FALDA 	Posición inicial: La persona puede elegir estar sentado o de pie. Debe haber un asiento con reposabrazos disponible para la persona en el lugar de la evaluación (por ejemplo: una silla de comedor, una silla de ruedas, una silla de baño, un taburete de baño). Evalúe sólo la estrategia de vestido (8a o 8b) elegida por la persona. La utilización de los cierres se puntúa en la sección 1.		
8a	Estrategia de vestido 8a: la persona introduce las piernas a través de la falda.			
8a.1	Pasar ambas piernas a través de la falda.	2: La persona introduce ambas piernas a través del orificio de la falda sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
8a.2	Subir la falda y colocarla correctamente.	2: La persona se sube la falda hasta la cintura y se la coloca correctamente sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		

8b	Estrategia de vestido 8b: la persona introduce los brazos y la cabeza a través de la falda.			
8b.1	Pasar ambos brazos a través de la falda.	2: La persona introduce ambos brazos por el lado correcto del orificio de la falda sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
8b.2	Pasar la falda sobre la cabeza.	2: La persona introduce la cabeza por el orificio de la falda sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
8b.3	Bajar la falda y colocarla correctamente.	2: La persona baja la falda hasta la cintura y la coloca correctamente sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
9.	BRAGAS/CALZONCIL	Posición inicial: La persona puede elegir estar sentado o de pie. Debe haber un asiento con		

	LOS 	reposabrazos disponible para la persona en el lugar de la evaluación (por ejemplo: una silla de comedor, una silla de ruedas, una silla de baño, un taburete de baño).		
9.1	Seleccionar el orificio correcto con ambas piernas.	2: La persona selecciona los orificios correctos e introduce las piernas por los orificios correspondientes sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
9.2	Subir la prenda y colocarla correctamente.	2: La persona se sube la prenda hasta cubrir las nalgas y los genitales y se la coloca correctamente sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		

10.	PANTALONES, LEGGINGS 	Posición inicial: La persona puede elegir estar sentado o de pie. Debe haber un asiento con reposabrazos disponible para la persona en el lugar de la evaluación (por ejemplo: una silla de comedor, una silla de ruedas, una silla de baño, un taburete de baño). La utilización de los cierres se puntúa en la sección 1.		
10.1	Seleccionar el orificio correcto con ambas piernas.	2: La persona selecciona los orificios correctos e introduce las piernas por los orificios correspondientes sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
10.2	Subir la prenda y colocarla correctamente.	2: La persona se sube la prenda hasta la cintura y se la coloca correctamente sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		

11.	MEDIAS, PANTYS, CALCETINES 	Posición inicial: La persona puede elegir estar sentado o de pie. Debe haber un asiento con reposabrazos disponible para la persona en el lugar de la evaluación (por ejemplo: una silla de comedor, una silla de ruedas, una silla de baño, un taburete de baño).		
11.1	Subir la prenda por los pies y tobillos y colocarla correctamente.	2: En el caso de los pantys, la persona selecciona el orificio correcto. Para las demás prendas, la persona pasa la prenda por los pies y los tobillos y la coloca correctamente sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
11.2	Subir la prenda por las piernas y colocarla	2: La persona sube la prenda por las piernas (en el caso de los pantys		

	correctamente.	deben subirse hasta la cintura) y colocarlas correctamente sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
12.	CALZADO 	Posición inicial: La persona puede elegir estar sentado o de pie. Debe haber un asiento con reposabrazos disponible para la persona en el lugar de la evaluación (por ejemplo: una silla de comedor, una silla de ruedas, una silla de baño, un taburete de baño). La utilización de los cierres se puntúa en la sección 1.		
12.1	Seleccionar el zapato correcto para cada pie.	2: La persona selecciona el zapato correcto para cada pie sin asistencia. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
12.2	Ponerse los zapatos en ambos pies.	2. La persona se pone los zapatos en ambos pies sin asistencia. El pie debe quedar cómodamente colocado y encajado en el zapato. 1: Necesita otro tipo de asistencia		

		(no física). 0: Necesita asistencia física.		
12.3	Atar los cordones/abrocharse ambos zapatos.	2: La persona se ata los cordones/se abrocha ambos zapatos, sin asistencia. Puede incluir diferentes estrategias y cierres, como abrochárselo con una mano, cordones elásticos, velcros o cierres para cordones. 1: Necesita otro tipo de asistencia (no física). 0: Necesita asistencia física.		
CÁLCULO DE LA PUNTUACIÓN DE ENVI-R				
Puntuación total de los ítems intentados (a):				
Número de ítems intentados (b):				
Puntuación de ENVI-R: $a / (b \times 2)$				

12.8. Anexo 8: ENVI-R Análisis de los errores de acción

ANÁLISIS DE LOS ERRORES DE ACCIÓN

NO.	TIPO DE ERROR DE ACCIÓN	EJEMPLOS	MARQUE LA CASILLA CORRESPONDIENTE
1	Problemas con la iniciación	Muestra dificultades para empezar a vestirse o para pasar a la siguiente prenda.	☐ Error de acción presente ☐ Error de acción ausente
2	Problemas para mantener la atención	Distraído. No se concentra (o no el tiempo suficiente) en la tarea de vestido.	☐ Error de acción presente ☐ Error de acción ausente
3	Enlentecimiento cognitivo	Completa la tarea del vestido, pero el desempeño es lento y requiere de tiempo adicional (por ejemplo, mucho tiempo buscando el orificio correcto de la prenda). Excluya los problemas motores y sensoriales.	☐ Error de acción presente ☐ Error de acción ausente
4	Problemas para localizar objetos en algunas partes del campo visual	No puede encontrar las prendas situadas a su izquierda/derecha.	☐ Error de acción presente ☐ Error de acción ausente
5	Inatención a un lado del cuerpo	No completa el vestido del lado izquierdo/derecho de su cuerpo.	☐ Error de acción presente ☐ Error de acción ausente
6	Problemas con el orden/secuenciación	Hace los pasos del vestido en un orden incorrecto (por ejemplo, se pone los zapatos antes que los calcetines).	☐ Error de acción presente ☐ Error de acción ausente
7	Perseveración	Repite un paso ya completado.	☐ Error de acción presente ☐ Error de acción ausente
8	Problemas para seleccionar los objetos correctos	Selecciona la prenda incorrecta (por ejemplo, coge los pantalones cuando se va a vestir la parte superior del cuerpo).	☐ Error de acción presente ☐ Error de acción ausente
9	Desorganización, ausencia de una estrategia clara	Muestra signos inconfundibles de no saber qué hacer con una prenda (por ejemplo, toquetea una	☐ Error de acción presente

		camisa sin parecer saber qué hacer con ella).	☐ Error de acción ausente
10	Deficiente autosupervisión de la realización de la tarea	Problemas con la autosupervisión del desempeño de la tarea y la resolución de problemas que surjan.	☐ Error de acción presente ☐ Error de acción ausente
11	Confusión espacial	No puede encontrar o seleccionar la apertura correcta de la prenda. Se pone la prenda del revés. Se coloca la parte trasera de la prenda, en la parte delantera del cuerpo (con la etiqueta del cuello en la parte delantera).	☐ Error de acción presente ☐ Error de acción ausente
12	Torpeza	Dificultades para agarrar las prendas o para utilizar los cierres. Excluya los problemas motores y sensoriales.	☐ Error de acción presente ☐ Error de acción ausente
OBSERVACIONES ADICIONALES:			

12.9. Anexo 9: Licencia de uso de la figura 1

25/9/25, 14:06

RightsLink Printable License

ELSEVIER LICENSE TERMS AND CONDITIONS

Sep 25, 2025

This Agreement between Pablo de Blas Zamorano ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	6115901092065
License date	Sep 25, 2025
Licensed Content Publisher	Elsevier
Licensed Content Publication	Cognitive Systems Research
Licensed Content Title	Automatic detection of ischemic stroke using higher order spectra features in brain MRI images
Licensed Content Author	U. Rajendra Acharya,Kristen M. Meiburger,Oliver Faust,Joel En Wei Koh,Shu Lih Oh,Edward J. Ciaccio,Asit Subudhi,V. Jahmunah,Sukanta Sabut
Licensed Content Date	Dec 1, 2019
Licensed Content Volume	58
Licensed Content Issue	n/a
Licensed Content Pages	9
Start Page	134
End Page	142
Type of Use	reuse in a thesis/dissertation

<https://s100.copyright.com/AppDispatchServlet>

1/7

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

25/9/25, 14:06

RightsLink Printable License

Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Title of new work	Análisis psicométrico y validación de instrumentos de evaluación funcional y de desempeño ocupacional desde terapia ocupacional en personas que han sufrido un ictus
Institution name	Universidad de Málaga
Expected presentation date	Dec 2025
Portions	Figure 1
The Requesting Person / Organization to Appear on the License	Pablo de Blas Zamorano
Requestor Location	Mr. Pablo de Blas Zamorano Calle José Paredes, 25 Murcia, 30150 Spain
Publisher Tax ID	GB 494 6272 12
Total	0.00 EUR
Terms and Conditions	

INTRODUCTION

<https://s100.copyright.com/AppDispatchServlet>

2/7

25/9/25, 14:06

RightsLink Printable License

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your RightsLink account and that are available at any time at <https://myaccount.copyright.com>).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given. The material may not be reproduced or used in any other way, including use in combination with an artificial intelligence tool (including to train an algorithm, test, process, analyse, generate output and/or develop any form of artificial intelligence tool), or to create any derivative work and/or service (including resulting from the use of artificial intelligence tools).

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier's permissions helpdesk [here](#)). No modifications can be made to any Lancet figures/tables and they must be reproduced in full.

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

<https://s100.copyright.com/AppDispatchServlet>

3/7

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. **Translation:** This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article.

16. **Posting licensed content on any Website:** The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at <http://www.sciencedirect.com/science/journal/xxxxx> or the Elsevier homepage for books at <http://www.elsevier.com>; Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at <http://www.elsevier.com>. All content posted to the web site must maintain the copyright information line on the bottom of each image.

Posting licensed content on Electronic reserve: In addition to the above the following clauses are applicable: The web site must be password-protected and made available only

to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

17. For journal authors: the following clauses are applicable in addition to the above:

Preprints:

A preprint is an author's own write-up of research results and analysis, it has not been peer-reviewed, nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancement etc.).

Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).

If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.

Accepted Author Manuscripts: An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes author-incorporated changes suggested during submission, peer review and editor-author communications.

Authors can share their accepted author manuscript:

- immediately
 - via their non-commercial person homepage or blog
 - by updating a preprint in arXiv or RePEc with the accepted manuscript
 - via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group
 - directly by providing copies to their students or to research collaborators for their personal use
 - for private scholarly sharing as part of an invitation-only work group on commercial sites with which Elsevier has an agreement
- After the embargo period
 - via non-commercial hosting platforms such as their institutional repository
 - via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:

- link to the formal publication via its DOI
- bear a CC-BY-NC-ND license - this is easy to do
- if aggregated with other manuscripts, for example in a repository or other site, be shared in alignment with our hosting policy not be added to or enhanced in any way to appear more like, or to substitute for, the published journal article.

Published journal article (PJA): A published journal article (PJA) is the definitive final record of published research that appears or will appear in the journal and embodies all value-adding publishing activities including peer review co-ordination, copy-editing, formatting, (if relevant) pagination and online enrichment.

Policies for sharing publishing journal articles differ for subscription and gold open access articles:

Subscription Articles: If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help your users to find, access, cite, and use the best available version.

Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

If you are affiliated with a library that subscribes to ScienceDirect you have additional private sharing rights for others' research accessed under that agreement. This includes use for classroom teaching and internal training at the institution (including use in course packs and courseware programs), and inclusion of the article for grant funding purposes.

Gold Open Access Articles: May be shared according to the author-selected end-user license and should contain a [CrossMark logo](#), the end user license, and a DOI link to the formal publication on ScienceDirect.

Please refer to Elsevier's [posting policy](#) for further information.

18. **For book authors** the following clauses are applicable in addition to the above: Authors are permitted to place a brief summary of their work online only. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. **Posting to a repository:** Authors are permitted to post a summary of their chapter only in their institution's repository.

19. **Thesis/Dissertation:** If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for Proquest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

Elsevier Open Access Terms and Conditions

You can publish open access with Elsevier in hundreds of open access journals or in nearly 2000 established subscription journals that support open access publishing. Permitted third party re-use of these open access articles is defined by the author's choice of Creative Commons user license. See our [open access license policy](#) for more information.

Terms & Conditions applicable to all Open Access articles published with Elsevier:

Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author's honour or reputation. If any changes have been made, such changes must be clearly indicated.

The author(s) must be appropriately credited and we ask that you include the end user license and a DOI link to the formal publication on ScienceDirect.

If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the

25/9/25, 14:06

RightsLink Printable License

user to ensure their reuse complies with the terms and conditions determined by the rights holder.

Additional Terms & Conditions applicable to each Creative Commons user license:

CC BY: The CC-BY license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article and to make commercial use of the Article (including reuse and/or resale of the Article by commercial entities), provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. The full details of the license are available at <http://creativecommons.org/licenses/by/4.0>.

CC BY NC SA: The CC BY-NC-SA license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article, provided this is not done for commercial purposes, and that the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. Further, any new works must be made available on the same conditions. The full details of the license are available at <http://creativecommons.org/licenses/by-nc-sa/4.0>.

CC BY NC ND: The CC BY-NC-ND license allows users to copy and distribute the Article, provided this is not done for commercial purposes and further does not permit distribution of the Article if it is changed or edited in any way, and provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, and that the licensor is not represented as endorsing the use made of the work. The full details of the license are available at <http://creativecommons.org/licenses/by-nc-nd/4.0>. Any commercial reuse of Open Access articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee.

Commercial reuse includes:

- Associating advertising with the full text of the Article
- Charging fees for document delivery or access
- Article aggregation
- Systematic distribution via e-mail lists or share buttons

Posting or linking by commercial companies for use by customers of those companies.

20. Other Conditions:

v1.11

Questions? customer care@copyright.com.

<https://s100.copyright.com/AppDispatchServlet>

7/7

12.10. Anexo 10: Licencia de uso de la figura 3

25/9/25, 13:47

RightsLink Printable License

WOLTERS KLUWER HEALTH, INC. LICENSE
TERMS AND CONDITIONS

Sep 25, 2025

This Agreement between Pablo de Blas Zamorano ("You") and Wolters Kluwer Health, Inc. ("Wolters Kluwer Health, Inc.") consists of your license details and the terms and conditions provided by Wolters Kluwer Health, Inc. and Copyright Clearance Center.

License Number	6115891402691
License date	Sep 25, 2025
Licensed Content Publisher	Wolters Kluwer Health, Inc.
Licensed Content Publication	Stroke
Licensed Content Title	Global, Regional, and National Burden of Stroke, 1990–2021: A Systematic Analysis for Global Burden of Disease 2021
Licensed Content Author	Qiang He, Wenjing Wang, Yangchang Zhang, Yang Xiong, et al.
Licensed Content Date	Oct 17, 2024
Licensed Content Volume	55
Licensed Content Issue	12
Type of Use	Dissertation/Thesis
Requestor type	University/College
Sponsorship	No Sponsorship
Format	Print and electronic

<https://s100.copyright.com/AppDispatchServlet>

1/5

Análisis Psicométrico y Validación de Instrumentos de Evaluación Funcional y de Desempeño desde Terapia Ocupacional en Personas que han sufrido un Ictus

25/9/25, 13:47

RightsLink Printable License

Will this be posted online?	Yes, on a secure website
Portion	Figures/tables/illustrations
Number of figures/tables/illustrations	1
Author of this Wolters Kluwer article	No
Will you be translating?	No
Intend to modify/change the content	No
Title of new work	Análisis psicométrico y validación de instrumentos de evaluación funcional y de desempeño ocupacional desde terapia ocupacional en personas que han sufrido un ictus
Institution name	Universidad de Málaga
Expected presentation date	Dec 2025
Portions	Figure 1. Global distribution of stroke burden in 2021. A, Age-standardized prevalence rates
The Requesting Person / Organization to Appear on the License	Pablo de Blas Zamorano
Requestor Location	Calle José Paredes, 25 Murcia, 30150 Spain
Publisher Tax ID	EU826013006
Billing Type	Invoice

<https://s100.copyright.com/AppDispatchServlet>

2/5

25/9/25, 13:47

RightsLink Printable License

Billing Address
Mr. Pablo de Blas Zamorano
Calle José Paredes, 25

Murcia, Spain 30150

Total 0.00 EUR

Terms and Conditions

Wolters Kluwer Health Inc. Terms and Conditions

1. **Duration of License:** Permission is granted for a one time use only. Rights herein do not apply to future reproductions, editions, revisions, or other derivative works. This permission shall be effective as of the date of execution by the parties for the maximum period of 12 months and should be renewed after the term expires.
 - i. When content is to be republished in a book or journal the validity of this agreement should be the life of the book edition or journal issue.
 - ii. When content is licensed for use on a website, internet, intranet, or any publicly accessible site (not including a journal or book), you agree to remove the material from such site after 12 months, or request to renew your permission license
2. **Credit Line:** A credit line must be prominently placed and include: For book content: the author(s), title of book, edition, copyright holder, year of publication; For journal content: the author(s), titles of article, title of journal, volume number, issue number, inclusive pages and website URL to the journal page; If a journal is published by a learned society the credit line must include the details of that society.
3. **Warranties:** The requestor warrants that the material shall not be used in any manner which may be considered derogatory to the title, content, authors of the material, or to Wolters Kluwer Health, Inc.
4. **Indemnity:** You hereby indemnify and hold harmless Wolters Kluwer Health, Inc. and its respective officers, directors, employees and agents, from and against any and all claims, costs, proceeding or demands arising out of your unauthorized use of the Licensed Material
5. **Geographical Scope:** Permission granted is non-exclusive and is valid throughout the world in the English language and the languages specified in the license.
6. **Copy of Content:** Wolters Kluwer Health, Inc. cannot supply the requestor with the original artwork, high-resolution images, electronic files or a clean copy of content.
7. **Validity:** Permission is valid if the borrowed material is original to a Wolters Kluwer Health, Inc. imprint (J.B Lippincott, Lippincott-Raven Publishers, Williams & Wilkins, Lea & Febiger, Harwal, Rapid Science, Little Brown & Company, Harper & Row Medical, American Journal of Nursing Co, and Urban & Schwarzenberg - English Language, Raven Press, Paul Hoeber, Springhouse, Ovid), and the Anatomical Chart Company
8. **Third Party Material:** This permission does not apply to content that is credited to publications other than Wolters Kluwer Health, Inc. or its Societies. For images credited to non-Wolters Kluwer Health, Inc. books or journals, you must obtain permission from the source referenced in the figure or table legend or credit line before making any use of the image(s), table(s) or other content.
9. **Adaptations:** Adaptations are protected by copyright. For images that have been adapted, permission must be sought from the rightsholder of the original material and the rightsholder of the adapted material.

<https://s100.copyright.com/AppDispatchServlet>

3/5

10. **Modifications:** Wolters Kluwer Health, Inc. material is not permitted to be modified or adapted without written approval from Wolters Kluwer Health, Inc. with the exception of text size or color. The adaptation should be credited as follows: Adapted with permission from Wolters Kluwer Health, Inc.: [the author(s), title of book, edition, copyright holder, year of publication] or [the author(s), titles of article, title of journal, volume number, issue number, inclusive pages and website URL to the journal page].
11. **Full Text Articles:** Reproduction of full articles in English is prohibited.
12. **Branding and Marketing:** No drug name, trade name, drug logo, or trade logo can be included on the same page as material borrowed from *Diseases of the Colon & Rectum*, *Plastic Reconstructive Surgery*, *Obstetrics & Gynecology (The Green Journal)*, *Critical Care Medicine*, *Pediatric Critical Care Medicine*, *the American Heart Association publications* and *the American Academy of Neurology publications*.
13. **Open Access:** Unless you are publishing content under the same Creative Commons license, the following statement must be added when reprinting material in Open Access journals: "The Creative Commons license does not apply to this content. Use of the material in any format is prohibited without written permission from the publisher, Wolters Kluwer Health, Inc. Please contact permissions@lww.com for further information."
14. **Translations:** The following disclaimer must appear on all translated copies: Wolters Kluwer Health, Inc. and its Societies take no responsibility for the accuracy of the translation from the published English original and are not liable for any errors which may occur.
15. **Published Ahead of Print (PAP):** Articles in the PAP stage of publication can be cited using the online publication date and the unique DOI number.
 - i. Disclaimer: Articles appearing in the PAP section have been peer-reviewed and accepted for publication in the relevant journal and posted online before print publication. Articles appearing as PAP may contain statements, opinions, and information that have errors in facts, figures, or interpretation. Any final changes in manuscripts will be made at the time of print publication and will be reflected in the final electronic version of the issue. Accordingly, Wolters Kluwer Health, Inc., the editors, authors and their respective employees are not responsible or liable for the use of any such inaccurate or misleading data, opinion or information contained in the articles in this section.
16. **Termination of Contract:** Wolters Kluwer Health, Inc. must be notified within 90 days of the original license date if you opt not to use the requested material.
17. **Waived Permission Fee:** Permission fees that have been waived are not subject to future waivers, including similar requests or renewing a license.
18. **Contingent on payment:** You may exercise these rights licensed immediately upon issuance of the license, however until full payment is received either by the publisher or our authorized vendor, this license is not valid. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of Wolters Kluwer Health, Inc.'s other billing and payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.
19. **STM Signatories Only:** Any permission granted for a particular edition will apply to subsequent editions and for editions in other languages, provided such editions are for the work as a whole in situ and do not involve the separate exploitation of the permitted illustrations or excerpts. Please view: [STM Permissions Guidelines](#)
20. **Warranties and Obligations:** LICENSOR further represents and warrants that, to the best of its knowledge and belief, LICENSEE's contemplated use of the Content as represented to LICENSOR does not infringe any valid rights to any third party.

25/9/25, 13:47

RightsLink Printable License

21. **Breach:** If LICENSEE fails to comply with any provisions of this agreement, LICENSOR may serve written notice of breach of LICENSEE and, unless such breach is fully cured within fifteen (15) days from the receipt of notice by LICENSEE, LICENSOR may thereupon, at its option, serve notice of cancellation on LICENSEE, whereupon this Agreement shall immediately terminate.
22. **Assignment:** License conveyed hereunder by the LICENSOR shall not be assigned or granted in any manner conveyed to any third party by the LICENSEE without the consent in writing to the LICENSOR.
23. **Governing Law:** The laws of The State of New York shall govern interpretation of this Agreement and all rights and liabilities arising hereunder.
24. **Unlawful:** If any provision of this Agreement shall be found unlawful or otherwise legally unenforceable, all other conditions and provisions of this Agreement shall remain in full force and effect.

For Copyright Clearance Center / RightsLink Only:

1. **Service Description for Content Services:** Subject to these terms of use, any terms set forth on the particular order, and payment of the applicable fee, you may make the following uses of the ordered materials:
 - i. **Content Rental:** You may access and view a single electronic copy of the materials ordered for the time period designated at the time the order is placed. Access to the materials will be provided through a dedicated content viewer or other portal, and access will be discontinued upon expiration of the designated time period. An order for Content Rental does not include any rights to print, download, save, create additional copies, to distribute or to reuse in any way the full text or parts of the materials.
 - ii. **Content Purchase:** You may access and download a single electronic copy of the materials ordered. Copies will be provided by email or by such other means as publisher may make available from time to time. An order for Content Purchase does not include any rights to create additional copies or to distribute copies of the materials

Other Terms and Conditions:

v1.18

Questions? customercare@copyright.com.

<https://s100.copyright.com/AppDispatchServlet>

5/5

12.11. Anexo 11: Modelo de Hoja informativa, Consentimiento Informado y Hoja de revocación.

CONSENTIMIENTO INFORMADO DE PARTICIPACIÓN VOLUNTARIA EN ESTUDIO DE INVESTIGACIÓN

Le invitamos a participar en el estudio que se detalla a continuación.

Título del estudio: “Traducción, adaptación transcultural y validación de la escala “Nottingham Stroke Dressing Assessment” (NSDA) y evaluación de las propiedades psicométricas de la escala Fugl-Meyer Assessment en población española con ictus”.

Antes de proceder a la firma de este consentimiento informado, lea atentamente la información que a continuación se le facilita, y realice las preguntas que considere oportunas.

Justificación y Objetivos del Estudio: Debido a la escasa disposición de escalas para el estudio de la actividad de la vida diaria del vestido, se ha considerado necesario la traducción, adaptación transcultural y validación de la escala “Nottingham Stroke Dressing Scale” (NSDA).

Dada la amplia variedad de versiones de la escala Fugl-Meyer Assessment, se ha hallado que la versión original de 50 ítems, es la más adecuada para evaluar la función sensoriomotora del miembro superior, y se pretende estudiar sus propiedades.

Diseño y métodos del Estudio: El presente estudio, es un estudio de validación de una escala para el estudio de la actividad de vestirse, y de la función del miembro superior de personas que han sufrido un ictus. Este estudio se realizará en los centros centro IntegraDC/Dacemur/Aspaym/Dismo/FundaciónAisse/Caser/Grupo5CIAN/Neuroavanzo/Neurovilla/Ceneri/Clínica Ribera Salud Puerto Lumbreras, contando con la autorización de la directora del centro, y con la conformidad del Comité de Ética de la Investigación detallado al final del documento. En una primera sesión, se le pedirá que se vista con su ropa habitual, y que realice una serie de movimientos con los brazos, para puntuar su capacidad de vestirse y su función sensoriomotora del miembro superior afectado; es decir, un total de 2 escalas. Además, se registrarán sus datos personales y clínicos. Esta sesión se llevará a cabo en su horario habitual de estancia en el centro.

A las dos semanas de la primera sesión, se le volverá a evaluar de la misma manera con estas 2 escalas.

Privacidad y Confidencialidad: Se garantizará la protección y confidencialidad de la información, ya que el tratamiento, la comunicación y la cesión de los datos de carácter personal de todos los sujetos participantes se ajustará a lo dispuesto en la Ley Orgánica 3/2018, del 5 de diciembre, de protección de datos personales y garantía de los derechos digitales, y en su reglamento de desarrollo de datos. Se hará un tratamiento confidencial y

numérico de la información, asignándose un código o bien utilizando el número de historia clínica y en ningún caso, figurará su nombre y apellidos, además de otros datos que comprometan su intimidad. Los datos recogidos para el estudio estarán identificados mediante un código y sólo el investigador y colaboradores podrán relacionar dichos datos con usted y con su historia clínica. Por lo tanto, su identidad no será revelada a persona alguna salvo en caso de urgencia médica, requerimiento de la administración sanitaria o requerimiento legal. Además, sólo se transmitirán a terceros y a otros países, previa notificación a la Agencia Española de Protección de Datos, los datos recogidos para el estudio que en ningún caso contendrán información que pueda identificarlo directamente, como nombre y apellidos, iniciales, dirección, nº de la seguridad social, etc. En el caso de que se produzca esta cesión, será para los mismos fines del estudio descrito y garantizando la confidencialidad como mínimo con el nivel de protección de la legislación vigente en nuestro país. El acceso a la información personal quedará restringido al personal investigador, autoridades sanitarias, al Comité de Ética de la Investigación de ambos hospitales incluidos en el estudio y al personal autorizado, cuando lo precisen para comprobar los datos y procedimientos del estudio, pero siempre manteniendo la confidencialidad de los mismos de acuerdo a la legislación vigente.

Aspectos éticos: Este estudio de investigación clínica se realizará siguiendo las recomendaciones de la Declaración de Helsinki y la normativa legal vigente en nuestro país en materia de investigación clínica, especialmente la Ley 14/2007, de 3 de julio, de Investigación Biomédica. Este proyecto cuenta además con la conformidad del Comité de Ética de la Investigación de los hospitales incluidos en el estudio.

HOJA DE CONSENTIMIENTO INFORMADO

Yo:_____. DNI/Pasaporte:_____.

He leído la hoja informativa que me ha sido entregada.

He tenido oportunidad de efectuar preguntas sobre el estudio. He recibido respuestas satisfactorias.

He recibido suficiente información en relación con el estudio. He hablado con el Dr./Investigador: Pablo de Blas Zamorano

Entiendo que la participación es voluntaria y que soy libre de participar o no en el estudio.

También he sido informado de forma clara, precisa y suficiente de los siguientes extremos que afectan a los datos personales que se contienen en este consentimiento y en la ficha o expediente que se abra para la investigación:

- Que estos datos serán tratados y custodiados con respeto a mi intimidad y a la vigente normativa de protección de datos (Ley Orgánica 3/2018, del 5 de diciembre, de protección de datos personales y garantía de los derechos digitales), por la que debe garantizarse la confidencialidad de los mismos.
- Sobre estos datos me asisten los derechos de acceso, rectificación, cancelación y oposición que podré ejercitar mediante solicitud ante el investigador responsable en la dirección de contacto que figura en este documento.
- Estos datos no podrán ser cedidos sin mi consentimiento expreso y no lo otorgo en este acto. Entiendo que puedo abandonar el estudio y retirar mi consentimiento:
 - Cuando lo desee.
 - Sin que tenga que dar explicaciones.
 - Sin que ello afecte a mis cuidados sanitarios

Por ello presto libremente mi CONFORMIDAD para participar en este proyecto de investigación sobre "Traducción, adaptación transcultural y validación de la escala "Nottingham Stroke Dressing Assessment" (NSDA) y evaluación de las propiedades psicométricas de la escala Fugl-Meyer Assessment en población española con ictus" hasta que decida lo contrario. Al firmar este consentimiento no renuncio a ninguno de mis derechos. Recibiré una copia de este consentimiento para guardarlo y poder consultarlo en el futuro. Nombre del paciente o sujeto colaborador:_____.

Indique el nombre del centro en el que se llevará a cabo:

Firma:

Fecha:.

Nombre del investigador: Pablo de Blas Zamorano

DNI: 06026210-A

Centro Asistencial:

Dirección de contacto: pablodeblasz@gmail.com

Firma:

HOJA DE REVOCACIÓN DE PARTICIPACIÓN.

Yo:_____. DNI/Pasaporte:_____.

Comprendo que los datos aportados son exclusivamente con fines científicos, quedando, con su guarda y custodia hasta la finalización del estudio que serán destruidos por el investigador principal, pudiendo ejercer sus derechos de modificación o anulación cuando desee en virtud de la Ley Orgánica 3/2018, del 5 de diciembre, de protección de datos personales y garantía de los derechos digitales).

Libre y conscientemente RETIRO mi consentimiento previamente dado para participar en este estudio. Pongo en conocimiento esta decisión a Pablo de Blas Zamorano (Investigador Principal), que no muestra oposición, y firmo para que conste.

(Fecha) (Nombre y firma del participante). (Fecha) (Nombre y firma del investigador).

12.12. Anexo 12: Autorización para investigar en Centro Asistencial 1

AUTORIZACIÓN PARA REALIZAR INVESTIGACIÓN EN CENTRO ASISTENCIAL

Yo, Dña. Doña Rosa María Delgado con DNI 8512831 como directora/presidenta del Centro Dismo, confirmo que he recibido toda la información y resuelto todas las dudas respecto al proyecto de investigación "Traducción, adaptación transcultural y validación de la escala "Nottingham Stroke Dressing Assessment" (NSDA) y evaluación de las propiedades psicométricas de la escala Fugl-Meyer Assessment en población española con ictus", llevado a cabo por el investigador principal Pablo de Blas Zamorano, con DNI 06026210-A, como proyecto de su Tesis Doctoral de la Universidad de Málaga.

Por lo cual, autorizo a Pablo de Blas Zamorano, a reclutar muestra en el centro asistencial Dismo, con domicilio en Avda. de Pío X, 1; con fines únicamente investigadores.

Firmado:

Pablo de Blas Zamorano

Odismo
ASOCIACIÓN DE PERSONAS CON DISCAPACIDAD
DE MOLINA DE SEGURA Y VEGA ALTA "ODISMO"
ICLIF: G-30100569
Avda. de Pío X, 1 - La Malineta
30300 MOLINA DE SEGURA - Tel: 968 64 41 97

12.13. Anexo 13: Autorización para investigar en Centro Asistencial 2

AUTORIZACIÓN PARA REALIZAR INVESTIGACIÓN EN CENTRO ASISTENCIAL

Yo, Dña, INMACULADA VICO DÍAZ con DNI 75107743R, como directora/presidenta de FUNDACIÓN AISSE con NIF G19679273, confirmo que he recibido toda la información y resuelto todas las dudas respecto al proyecto de investigación "Traducción, adaptación transcultural y validación de la escala "Nottingham Stroke Dressing Assessment" (NSDA) y evaluación de las propiedades psicométricas de la escala Fugl-Meyer Assessment en población española con ictus", llevado a cabo por el investigador principal Pablo de Blas Zamorano, con DNI 06026210-A, como proyecto de su Tesis Doctoral de la Universidad de Málaga.

Por lo cual, autorizo a Pablo de Blas Zamorano, a reclutar muestra en el centro asistencial SINERGIA de Fundación AISSE, con domicilio en Calle Pintor Manuel Maldonado, 14 del 18007 Granada; con fines únicamente investigadores.

Firmado:

Firmado por 75107743R INMACULADA VICO
(R: G19679273) el día 12/06/2024
con un certificado emitido por AC
Representación

Pablo de Blas Zamorano

Firmado por PABLO DE
BLAS ZAMORANO con un
certificado emitido por
06026210A

12.14. Anexo 14: Autorización para investigar en Centro Asistencial 3

AUTORIZACIÓN PARA REALIZAR INVESTIGACIÓN EN CENTRO ASISTENCIAL

Yo, Dña. Milagros Muñoz Pagan, con DNI 23276121-Y, como directora del Centro INTEGRA DAÑO CEREBRAL, confirmo que he recibido toda la información y resuelto todas las dudas respecto al proyecto de investigación "Traducción, adaptación transcultural y validación de la escala "Nottingham Stroke Dressing Assessment" (NSDA) y evaluación de las propiedades psicométricas de la escala Fugl-Meyer Assessment en población española con ictus", llevado a cabo por el investigador principal Pablo de Blas Zamorano, con DNI 06026210-A, como proyecto de su Tesis Doctoral de la Universidad de Málaga.

Por lo cual, autorizo a Pablo de Blas Zamorano, a reclutar muestra en el centro asistencial INTEGRA DAÑO CEREBRAL, con domicilio en Calle Claveles, nº6, Murcia; con fines únicamente investigadores.

Firmado:


C/ de los Claveles nº6 - 30007 (Murcia)
CIF B73731416 - Telf. 868 07 90 50

Milagros Muñoz Pagan

Firmado por PABLO DE BLAS
ZAMORANO con un certificado
emitido por 06026210A

Pablo de Blas Zamorano

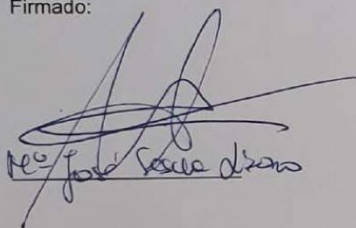
12.15. Anexo 15: Autorización para investigar en Centro Asistencial 4

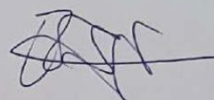
AUTORIZACIÓN PARA REALIZAR INVESTIGACIÓN EN CENTRO ASISTENCIAL

Yo, Dña. M^{te} José Sesma Lizaso, con DNI 34.830345-6, como directora/presidenta del Centro Neuroavento, confirmo que he recibido toda la información y resuelto todas las dudas respecto al proyecto de investigación "Traducción, adaptación transcultural y validación de la escala "Nottingham Stroke Dressing Assessment" (NSDA) y evaluación de las propiedades psicométricas de la escala Fugl-Meyer Assessment en población española con ictus", llevado a cabo por el investigador principal Pablo de Blas Zamorano, con DNI 06026210-A, como proyecto de su Tesis Doctoral de la Universidad de Málaga.

Por lo cual, autorizo a Pablo de Blas Zamorano, a reclutar muestra en el centro asistencial Neuroavento con domicilio en C/Geta, nº 14, 30.007, con fines únicamente investigadores.

Firmado:


M^{te} José Sesma Lizaso



Pablo de Blas Zamorano

12.16. Anexo 16: Autorización para investigar en Centro Asistencial 5

AUTORIZACIÓN PARA REALIZAR INVESTIGACIÓN EN CENTRO ASISTENCIAL

Yo, Dña. María Sáez Pajares, con DNI 11854713F, como directora/presidenta del Centro Neuron Las Mercedes, confirmo que he recibido toda la información y resuelto todas las dudas respecto al proyecto de investigación "Traducción, adaptación transcultural y validación de la escala "Nottingham Stroke Dressing Assessment" (NSDA) y evaluación de las propiedades psicométricas de la escala Fugl-Meyer Assessment en población española con ictus", llevado a cabo por el investigador principal Pablo de Blas Zamorano, con DNI 06026210-A, como proyecto de su Tesis Doctoral de la Universidad de Málaga.

Por lo cual, autorizo a Pablo de Blas Zamorano, a reclutar muestra en el centro asistencial Neuron Las Mercedes, con domicilio en Calle Cigoitia 13; con fines únicamente investigadores.

Firmado:

Firmado por PABLO DE BLAS
ZAMORANO con un certificado
emitido por Pablo de Blas Zamorano

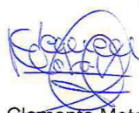
12.17. Anexo 17: Autorización para investigar en Centro Asistencial 6

AUTORIZACIÓN PARA REALIZAR INVESTIGACIÓN EN CENTRO ASISTENCIAL

Yo, Dña. Clementa Mota, con DNI 04619307-X, como directora del Centro Neuron Madrid Rio, confirmo que he recibido toda la información y resuelto todas las dudas respecto al proyecto de investigación "Traducción, adaptación transcultural y validación de la escala "Nottingham Stroke Dressing Assessment" (NSDA) y evaluación de las propiedades psicométricas de la escala Fugl-Meyer Assessment en población española con ictus", llevado a cabo por el investigador principal Pablo de Blas Zamorano, con DNI 06026210-A, como proyecto de su Tesis Doctoral de la Universidad de Málaga.

Por lo cual, autorizo a Pablo de Blas Zamorano, a reclutar muestra en el centro asistencial Neuron Madrid Rio, con domicilio en C/ Anracita 7; con fines únicamente investigadores.

Firmado:



Clementa Mota Valbuena

Firmado por PABLO DE BLAS
ZAMORANO con un certificado
emitido por 06026210A

Pablo de Blas Zamorano

12.18. Anexo 18: Autorización para investigar en Centro Asistencial 7

AUTORIZACIÓN PARA REALIZAR INVESTIGACIÓN EN CENTRO ASISTENCIAL

Yo, Dña. ESTEFANÍA GARCÍA ZAMORA, con DNI 50767345-C, como subdirectora del Centro NEURON NUEVOS MINISTERIOS, confirmo que he recibido toda la información y resuelto todas las dudas respecto al proyecto de investigación "Traducción, adaptación transcultural y validación de la escala "Nottingham Stroke Dressing Assessment" (NSDA) y evaluación de las propiedades psicométricas de la escala Fugl-Meyer Assessment en población española con ictus", llevado a cabo por el investigador principal Pablo de Blas Zamorano, con DNI 06026210-A, como proyecto de su Tesis Doctoral de la Universidad de Málaga.

Por lo cual, autorizo a Pablo de Blas Zamorano, a reclutar muestra en el centro asistencial _NEURON NUEVOS MINISTERIOS_____, con domicilio en C/ AGUSTÍN DE BETANCOURT, 17, 28003 MADRID; con fines únicamente investigadores.

Firmado:

ESTEFANÍA GARCÍA ZAMORA

Firmado por PABLO DE BLAS
ZAMORANO con un certificado
emitido por 06026210A

Pablo de Blas Zamorano

12.19. Anexo 19: Autorización para investigar en Centro Asistencial 8

AUTORIZACIÓN PARA REALIZAR INVESTIGACIÓN EN CENTRO ASISTENCIAL

Yo, D./Dña. ROSALBA VILLANUEVA con DNI 3177586-E, como director/a del Centro Clínica Ribera Salud Puerto Lumbreras, confirmo que he recibido toda la información y resuelto todas las dudas respecto al proyecto de investigación "Traducción, adaptación transcultural y validación de la escala "Nottingham Stroke Dressing Assessment" (NSDA) y evaluación de las propiedades psicométricas de la escala Fugl-Meyer Assessment en población española con ictus", llevado a cabo por el investigador principal Pablo de Blas Zamorano, con DNI 06026210-A, como proyecto de su Tesis Doctoral de la Universidad de Málaga.

Por lo cual, autorizo a Pablo de Blas Zamorano, a reclutar muestra en el centro asistencial Clínica Ribera Salud Puerto Lumbreras, con domicilio en Av. Pedro García Rubio, 55, 30890; con fines únicamente investigadores.

Firmado:

Rosalba Villanueva

Firmado por PABLO DE BLAS ZAMORANO con un certificado emitido por 06026210A

Pablo de Blas Zamorano

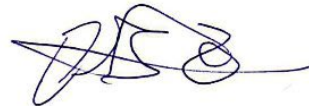
12.20. Anexo 20: Autorización para investigar en Centro Asistencial 9

AUTORIZACIÓN PARA REALIZAR INVESTIGACIÓN EN CENTRO ASISTENCIAL

Yo, Dña. Eulalia López Bastida, con DNI 27482256-Q, como directora/presidenta del Centro DACEMUR, confirmo que he recibido toda la información y resuelto todas las dudas respecto al proyecto de investigación "Traducción, adaptación transcultural y validación de la escala "Nottingham Stroke Dressing Assessment" (NSDA) y evaluación de las propiedades psicométricas de la escala Fugl-Meyer Assessment en población española con ictus", llevado a cabo por el investigador principal Pablo de Blas Zamorano, con DNI 06026210-A, como proyecto de su Tesis Doctoral de la Universidad de Málaga.

Por lo cual, autorizo a Pablo de Blas Zamorano, a reclutar muestra en el centro asistencial DACEMUR, con domicilio en Avenida de la Ñora, nº96, Murcia; con fines únicamente investigadores.

Firmado:



DACEMUR

Pablo de Blas Zamorano

Firmado por ***8225** EULALIA LOPEZ
(R: ****7556*) el día 21/02/2024
con un certificado emitido por AC
Representación

12.21. Anexo 21: Autorización para investigar en Centro Asistencial 10

AUTORIZACIÓN PARA REALIZAR INVESTIGACIÓN EN CENTRO ASISTENCIAL

Yo, Dña. Ascensión Martínez Sánchez, con DNI 27483904P, como representante del Centro Caser Residencial Santo Ángel, confirmo que he recibido toda la información y resuelto todas las dudas respecto al proyecto de investigación "Traducción, adaptación transcultural y validación de la escala "Nottingham Stroke Dressing Assessment" (NSDA) y evaluación de las propiedades psicométricas de la escala Fugl-Meyer Assessment en población española con ictus", llevado a cabo por el investigador principal Pablo de Blas Zamorano, con DNI 06026210-A, como proyecto de su Tesis Doctoral de la Universidad de Málaga.

Por lo cual, autorizo a Pablo de Blas Zamorano, a reclutar muestra en el centro asistencial Caser Residencial Santo Ángel, con domicilio en C/ Madre María Séiquer, 48; con fines únicamente investigadores.

Firmado:



Ascensión Martínez Sánchez



Pablo de Blas Zamorano

12.22. Anexo 22: Autorización para investigar en Centro Asistencial 11



MEMORIA DE SOLICITUD DE PROYECTOS DE INVESTIGACIÓN

FPR-PE02-07.01
Versión 01
Página 1 de 5

TÍTULO:
Traducción, adaptación transcultural y validación de la escala "Nottingham Stroke Dressing Assessment" (NSDA) y evaluación de las propiedades psicométricas de la escala Fugl-Meyer Assessment en población española con ictus.
Responsable del proyecto (incluir teléfono, email):
Pablo de Blas Zamorano Pablobz.terapeuta@gmail.com 618976458
Centro de trabajo del responsable del proyecto:
Universidad de Málaga
Equipo investigador (especificar si alguien tiene experiencia en investigación):
Proyecto de Tesis Doctoral. Programa de Doctorado en Ciencias de la Salud. Universidad de Málaga. Investigador principal: Pablo de Blas Zamorano. Experiencia en investigación en grupos de investigación independientes, grupos de investigación del colegio de Terapeutas Ocupacionales de la Región de Murcia, y en el Máster de Investigación en ciencias Sanitarias. Investigadores adscritos al proyecto: David Pérez Cruzado: Profesor Doctor del Grado de Terapia Ocupacional, Facultad de Medicina. Universidad de Málaga. Amplia experiencia en Investigación en ámbitos nacional e internacional, con múltiples publicaciones de alto factor de impacto. Experiencia en docencia universitaria. José Antonio Merchán Baeza: Profesor Doctor del Grado de Terapia Ocupacional, Facultad de Ciencias de la Salud, Universitat de Vic. Universitat Central de Catalunya. Perteneciente a varios grupos de investigación nacionales e internacionales, con múltiple experiencia en investigación y desarrollo de proyectos nacionales e internacionales. Experiencia en docencia universitaria.
(Nombre, centro de trabajo):
(Nombre, centro de trabajo):
(Nombre, centro de trabajo):
(Nombre, centro de trabajo):
(Nombre, centro de trabajo):

Dispositivos de GRUPO 5 implicados:
Terapeutas Ocupacionales
Centro Integral de Atención Neurorehabilitadora, Grupo 5 CIAN. C/ Alejo Carpentier, 2, 28806 Alcalá de Henares. Madrid
Centro Integral de Atención Neurorehabilitadora, Grupo 5 CIAN. Plaza de Europa, 9. Ciudad del Transporte de Pamplona, 31119, Imaoain, Navarra.
Centro Integral de Atención Neurorehabilitadora, Grupo 5 CIAN. Paseo de ruiseñores, 6-8, Zaragoza.
Centro Integral de Atención Neurorehabilitadora, Grupo 5 CIAN. Calle Julián Sanz Ibáñez, 57, 50017, Zaragoza.
Otros organismos o entidades colaboradores:
Diversos centros de neurorrehabilitación colaboran con las personas usuarias de sus centros, para lograr alcanzar la muestra de población necesaria para llevar a cabo este estudio de investigación.
Centros colaboradores:
- Integra Daño Cerebral
- Dacemur
- Aspaym Murcia

- Dismo, Asociación de Discapacitados
- Fundación Aisse
- Caser, Murcia
- Clínica Ribera Salud Puerto Lumbreras
- Neuron Madrid
- Neurovila

Duración estimada del proyecto (en meses):

Dependiendo de la muestra que puedan aportar, los centros colaboradores, y el tiempo del que dispongan.

En administrar la escala Fugl-Meyer Assessment (solo la parte motora, no la de sensación, ni rangos articulares ni dolor), tanto de miembro superior como de miembro inferior, se tardan 30 minutos.

En administrar la escala ENVI-R se tardan 15 minutos aproximadamente.

En total, se necesitan entre 30 minutos (mínimo), y 50 minutos (máximo) por paciente. En mi experiencia, que, como investigador principal, llevo evaluados a 50 pacientes, tardo de media 40 minutos por persona aproximadamente.

Aproximadamente, en una semana como máximo, se puede evaluar a los usuarios del centro.

Se debe realizar una reevaluación tras un periodo mínimo de 2 semanas. Por lo que la duración estimada de la colaboración de Grupo 5 CIAN será de aproximadamente 1 mes o mes y medio, teniendo que evaluar a sus pacientes/usuarios en una semana, dejar pasar un periodo de 2 semanas entre medias, y posteriormente reevaluar a los pacientes/usuarios de nuevo, en una semana.

JUTIFICACIÓN DEL ESTUDIO (revisión bibliográfica, resultados obtenidos, impacto esperado)

La evaluación de la actividad de la vida diaria de vestirse resulta importante en personas que han sufrido un ictus, ya que debido a las secuelas causadas por dicha patología, la actividad del vestido se ve afectada. Esta actividad suele ser comúnmente trabajada desde Terapia Ocupacional, con el fin de aportar mayor autonomía e independencia a la persona; por ello es necesario contar con una escala específica del vestido, que sea sensible y registre todos los fallos y mejoras en los diferentes tipos de prendas. Es por ello que se ha seleccionado la escala NSDA para traducirla, adaptarla transculturalmente y validarla en población española.

Por otro lado, se ha identificado que la versión original de Fugl-Meyer Assessment (50 ítems) es la más adecuada para la evaluación sensoriomotora del miembro superior afecto. Esta variable suele ser frecuentemente evaluada y trabajada en la rehabilitación tras el ictus. Se ha identificado que esta escala ha sido validada en población española, pero sin embargo se desconocen sus propiedades psicométricas. Es por ello que en este proyecto de investigación se va a llevar a cabo un análisis de las propiedades psicométricas de esta versión original validada al castellano de la escala Fugl-Meyer Assessment.

RESUMEN (Objetivos, Diseño, Ámbito de estudio, Muestra, Intervenciones que se van a aplicar, instrumentos de evaluación, Dificultades y Limitaciones):

“Nottingham Stroke Dressing Assessment” (NSDA) es una escala que evalúa el desempeño en la actividad de vestirse en personas que han sufrido un ictus. El objetivo de este proyecto de investigación es realizar una traducción y adaptación transcultural de la escala NSDA al castellano y validación en la población española; así como evaluar las propiedades psicométricas de esta. A la misma vez, se evaluarán las propiedades psicométricas de la escala Fugl-Meyer Assessment (versión larga de 50 ítems); la cual, tras la realización de una búsqueda sistemática siguiendo la metodología COSMIN ha resultado ser la versión más recomendada de Fugl Meyer Assessment. Para ello, se empleará la metodología de traducción establecida en la guía ISPOR. Tras ello, se administrará la escala generada a una muestra mínima de 200 personas, distribuidos en diferentes centros colaboradores, y siempre respetando la intimidad y protección de datos de la persona. Para finalizar, se realizará un análisis estadístico psicométrico de los datos obtenidos en las escalas administradas a la muestra anteriormente seleccionada. Los datos estadísticos guiarán las modificaciones y/o adaptaciones que serán necesarias hacer para que la traducción y adaptación de la escala NSDA sea finalmente válida en población española; y para conocer las propiedades psicométricas de Fugl-Meyer Assessment.

PLAN DE TRABAJO (Etapas de desarrollo del estudio. Lugar de realización del estudio, incluir cronograma)

Objetivos del estudio:

- Traducción y adaptación transcultural de la escala “Nottingham Stroke Dressing Assessment” al castellano.
- Validación y evaluación de las propiedades psicométricas de la traducción de la escala “Nottingham Stroke Dressing Assessment” en población española que ha sufrido un ictus.
- Evaluación de las propiedades psicométricas de Fugl-Meyer Assessment (50 ítems) en personas que han sufrido un ictus.

Muestra:

Los participantes que se incluyan en el estudio, deberán cumplir los siguientes criterios de inclusión:

- Personas mayores de edad (18-99 años).
- Haber sufrido, al menos, un accidente cerebrovascular.
- Ser capaz de mover activamente (Daniels 4/5) el hemisferio “sano”.

También existen los siguientes criterios de exclusión:

- Ausencia de respuesta al medio (Glasgow inferior a 15).
- Incapacidad para comprender las instrucciones del terapeuta.
- Estar conectado a un dispositivo de soporte vital básico (por ejemplo, respirador, bomba de baclofeno, etc.).
- Agresividad o suponer un riesgo para sí mismo o los demás.

- Estar encamado.

Secuencia:

Se contactará con determinados centros y clínicas de neurorrehabilitación, en función de la muestra a la que se pueda tener acceso en ellos, para que autoricen al investigador principal (Pablo de Blas Zamorano) a realizar el estudio de investigación en dichos centros y concretar asuntos de logística (cantidad de muestra que cumpla con los criterios de inclusión y exclusión, horario para llevar a cabo la evaluación, espacio necesario, recursos disponibles, etc.) para llevar a cabo el estudio en el centro.

El investigador principal, realizará una videoconferencia de 1 hora aproximadamente con los terapeutas ocupacionales de los centros que vayan a participar en el estudio, para explicarles cómo administrar las escalas.

Se contactará con los pacientes que cumplan los requisitos para ser incluidos en el estudio, y en el caso de ser necesario con sus familiares también. Se les entregará un documento informativo, con los datos necesarios sobre el estudio de investigación, y lo que se le pedirá a los pacientes durante la realización del mismo.

Tras ello, se les entregará el Consentimiento Informado, para que quede constancia de la conformidad de participación. Deberá entregarse una copia al investigador principal o al centro de referencia, y los pacientes o sus familiares, quedarse con otra copia. En dicho documento, también aparecerá la hoja de revocación del estudio, para en el caso de que se desee, cesar la participación en la investigación.

Cuando se haya formalizado todo esto, se evaluará a los 200 pacientes en total, entre los diferentes centros colaboradores, con ambas escalas:

- ENVI-R para evaluar el vestido, observando mientras el paciente coge y se coloca la ropa, y otorgando una puntuación acorde.
- Fugl-Meyer Assessment (versión original de 50 ítems) para medir la función sensoriomotora del miembro superior, a través de la observación del paciente realizando los movimientos requeridos en la escala, y otorgando la puntuación obtenida.

Se evaluará en 2 ocasiones diferentes, con un intervalo de tiempo mínimo de 2 semanas y máximo de 2 meses.

Tras la obtención de los resultados, se enviarán al investigador principal, mediante email.

Una vez se haya evaluado a la muestra necesaria, comenzará el análisis estadístico de los datos con el programa informático SPSS y la extensión AMOS.

Protección de datos y buenas prácticas éticas:

A la hora de administrar estas escalas, se preservará el respeto y la intimidad de estas personas de la siguiente manera:

El investigador principal le explicará al paciente en qué consiste el proyecto y la evaluación con estas escalas, y en el caso de que la persona no quiera participar o ser evaluado con dichas escalas, no habrá ningún problema, se le retirará del estudio. De esta manera, sólo participarán en el estudio las personas que voluntariamente acepten, conociendo el procedimiento y la duración del mismo. A las personas que acepten participar, se les pedirá que firmen un consentimiento informado para vincularse al estudio. En el caso de que durante el desarrollo de esta investigación, la persona quisiera retirarse del estudio, lo podría hacer sin ningún problema.

A las personas evaluadas se les asignará un código numérico con el fin de preservar su intimidad y proteger sus datos, de esta manera no aparecerá el nombre, ni la edad ni ningún dato médico ni personal de la persona. Además se protegerán los datos de los participantes según la LOPD 3/2018, de 5 de diciembre; y en todo momento se actuará siguiendo las buenas prácticas encontradas en la declaración de Helsinki (Asociación Médica Mundial., 2017).

El único dato que aparecerá en las escalas, será el código asignado. De esta manera se garantizará la protección de datos de todos los participantes.

MEDIOS (Aquellos necesarios para la realización del proyecto: personal de GRUPO 5 implicado, recursos materiales y técnicos que se utilizarán durante la realización del estudio, recursos económicos necesarios, tamaño de la muestra)

Para la participación de Grupo 5 CIAN en este estudio, los recursos necesarios serán:

En primero lugar, la muestra de pacientes que cumplan los requisitos de inclusión previamente indicados, y que accedan voluntariamente a participar en el estudio mediante la firma del consentimiento informado.

Terapeutas Ocupacionales que evalúen a los pacientes con las 2 escalas. En el caso de no ser posible, las escalas podrían ser administradas por fisioterapeutas.

Un tiempo de aproximadamente 40-45 minutos por paciente para la administración de la escala.

Dispositivo tecnológico para realizar la videoconferencia de formación sobre las escalas, y para enviar los resultados de las evaluaciones al investigador principal.

UTILIDAD Y O RELEVANCIA PRÁCTICA (De los resultados para Grupo 5)

Grupo 5 podrá disponer de los resultados de las diferentes evaluaciones realizadas a los pacientes (función sensoriomotora de miembro inferior afecto, función sensoriomotora de miembro superior afecto, desempeño e independencia en la AVD del vestido), para poder conocer el estado y función de sus personas usuarias, así como poder ser utilizada como indicador de éxito del tratamiento, permitiendo comparar los resultados obtenidos en las evaluaciones con dichas escalas.

En la publicación del artículo científico, Grupo 5 CIAN aparecerá como centro colaborador y en agradecimientos.

En el momento en el que las escalas validadas se hagan públicas, el investigador principal las enviará de manera directa al Grupo 5, a modo de agradecimiento por su aportación a que las escalas se puedan validar y al desarrollo científico de los profesionales e investigadores involucrados.

A la Att. de Pablo de Blas Zamorano

En relación con la solicitud presentada a la entidad, y relativa a la investigación **Traducción, adaptación transcultural y validación de la escala "Nottingham Stroke Dressing Assessment" (NSDA) y evaluación de las propiedades psicométricas de la escala Fugl-Meyer Assessment en población española con ictus.**, el resultado de la evaluación es el siguiente:

DICTÁMEN DEL ÁREA TÉCNICA: FAVORABLE

La presente investigación reúne los requisitos de calidad necesarios para su realización: pertinencia de la metodología, interés científico del tema, factibilidad, utilidad y adecuación a los principios que rigen las líneas de interés de GRUPO 5 CIAN.

En base a lo arriba indicado, el resultado global es **FAVORABLE**, por lo tanto, **QUEDA AUTORIZADA** la realización de la investigación planteada.

A partir de este momento, cuenta con el visto bueno de Grupo 5 para iniciar su investigación, debiendo ser objeto de seguimiento durante su desarrollo. Del mismo modo, cualquier modificación durante la realización de la misma deberá ser notificada previamente a la ejecución de los cambios, para que desde Grupo 5 pueda evaluarse la conformidad de los mismos.]

Agradecemos su colaboración con nuestra entidad y el trabajo desarrollado hasta ahora, y animamos a seguir adelante con las labores de investigación, motor fundamental e indiscutible en la mejora de la intervención con personas.

En Madrid, a 10 de ABRIL de 2024:

Miguel Simón
Director del Departamento Técnico y de Calidad
Grupo 5

C/ Juan Esplandiú 11, planta 11
28007 Madrid
Tel.: + 34 91 402 60 12
Fax.: +34 91 401 68 06
contacto@grupo5.net
www.grupo5.net