

# Estudio de validación de la versión española de la escala Fahn-Tolosa-Marín para el temblor esencial

M. Victoria Castrillo-Fraile, Elena Casas-Peña, José M. Trejo-Gabriel y Galán, Esther Cubo

**Introducción.** Hasta ahora no existe una validación formal de la versión española de la escala Fahn-Tolosa-Marín (FTM), usada ampliamente para valorar la gravedad del temblor.

**Objetivo.** Analizar la validez y la fiabilidad de la versión en castellano de la escala FTM.

**Pacientes y métodos.** Estudio observacional transversal en pacientes diagnosticados de temblor esencial. Se evaluó la gravedad del temblor con la escala FTM; la discapacidad en el miembro superior, con la escala *Disabilities of the Arm, Shoulder and Hand* (DASH), y la calidad de vida relacionada con la salud, con el cuestionario *Short Form-36 Health Survey* (SF-36). El análisis estadístico incluyó descripción de la muestra, fiabilidad (alfa de Cronbach), validez convergente y capacidad discriminativa (curvas ROC).

**Resultados.** Se incluyó a 40 pacientes (22 mujeres y 18 hombres con temblor esencial), con una edad media de 65,8 ± 14,7 años (rango: 21-90 años). La fiabilidad de la FTM fue alta, con un alfa de Cronbach de 0,90 (subescala A: 0,85; subescala B: 0,91; subescala C: 0,77). Para evaluar la discapacidad originada por el temblor esencial, la validez convergente entre las escalas DAHS y FTM (subescala C) fue adecuada, con una capacidad diagnóstica aceptable: área bajo la curva de 0,86 (intervalo de confianza al 95%: 0,67-1,00), sensibilidad del 78% y especificidad del 75%, para un punto de corte mayor de 5,5.

**Conclusión.** La escala FTM es un instrumento fiable, válido y preciso para la valoración del temblor esencial en la población española adulta.

**Palabras clave.** Análisis clinimétrico. Discapacidad. Escala Fahn-Tolosa-Marín. Escalas de medida. Resultados. Temblor. Temblor esencial.

## Introducción

El temblor se caracteriza por ser un movimiento rítmico y oscilatorio de una parte del cuerpo, con una frecuencia relativamente constante y amplitud variable, debido a una falta de coordinación entre la musculatura agonista y la antagonista [1]. Para realizar el diagnóstico del temblor es necesaria una minuciosa exploración neurológica que permita caracterizar el tipo de movimiento en función de su frecuencia, distribución, amplitud, patrón de aparición y asociación a otras alteraciones de la exploración neurológica.

La escala Fahn-Tolosa-Marín (FTM) se publicó por primera vez en 1988 y se revisó en 1993 [2], y se usa frecuentemente en la práctica clínica habitual y como medida de eficacia en ensayos clínicos. Esta escala contiene tres secciones (A, para evaluar la amplitud del temblor de reposo, el postural y el cinético en localizaciones anatómicas específicas; B, para el temblor de escritura, dibujo y vertido; y C, para las actividades de la vida diaria), tiene una valoración global por el paciente y el examinador, y cada

ítem está calificado en una escala de 0 a 4. Las principales diferencias de la escala FTM revisada sobre la versión inicial son la inclusión del temblor ortostático del tronco y los miembros inferiores en bipedestación, la inclusión de una sección que evalúa el impacto del temblor en las actividades sociales y la definición de temblor grave extremo como > 4 cm de amplitud [3]. Las medidas para el temblor de la cara, la lengua, la voz, la cabeza, el tronco y el temblor ortostático se valoran como: ninguno, 0; leve, 1; moderado, 2; marcado, 3; y grave, 4; y se recomienda un entrenamiento previo en los evaluadores para lograr una adecuada valoración [3]. La definición de intensidad extrema de temblor (> 4 cm) conlleva ciertas dificultades, pues puede presentar un efecto techo problemático en los pacientes con temblor esencial avanzado, observado en los ensayos de estimulación cerebral profunda, y la falta de definición clara en las instrucciones a la hora de usar la escala.

Desde un punto de vista clinimétrico, la escala FTM está recomendada por la International Parkinson's Disease and Movement Disorder Society para su uso clínico y como medida de eficacia en

Facultad de Medicina y Ciencias de la Salud; Universidad de Valladolid; Valladolid (M.V. Castrillo-Fraile, J.M. Trejo-Gabriel y Galán, E. Cubo). Servicio de Rehabilitación (M.V. Castrillo-Fraile); Servicio de Neurología (E. Casas-Peña, J.M. Trejo-Gabriel y Galán, E. Cubo); Hospital Universitario de Burgos; Burgos, España.

### Correspondencia:

Dra. María Victoria Castrillo Fraile. Servicio de Rehabilitación. Hospital Universitario de Burgos. Avda. Islas Baleares, 3. E-09006 Burgos.

### E-mail:

vcastrillof@saludcastillayleon.es

### Financiación:

Beca de la Fundación Burgos por la Salud y la Investigación, año 2015-2016.

Acceptado tras revisión externa: 20.03.19.

### Cómo citar este artículo:

Castrillo-Fraile MV, Casas-Peña E, Trejo-Gabriel y Galán JM, Cubo E. Estudio de validación de la versión española de la escala Fahn-Tolosa-Marín para el temblor esencial. Rev Neurol 2019; 69: 265-70. doi: 10.33588/m.6907.2019067.

© 2019 Revista de Neurología

ensayos clínicos [4]. Sin embargo, a pesar del amplio uso de la versión española de la escala FTM en la práctica clínica habitual y en investigación [5], su validación al castellano no se ha estudiado todavía. En este sentido, está formalmente aceptado que la caracterización de las propiedades clinimétricas es un requisito obligatorio durante la validación de escalas clínicas como instrumento de medida, cuando éstas se aplican en una versión diferente, o más allá del uso habitual de la propia escala, en poblaciones diferentes a lo establecido en la escala original [6-10]. El objetivo de este estudio es, por tanto, caracterizar las propiedades clinimétricas de la versión en castellano de la escala de FTM para el temblor esencial con el fin de analizar su validez y fiabilidad.

## Pacientes y métodos

### Diseño y muestra

Estudio observacional, transversal. Se ha incluido una muestra consecutiva de 109 pacientes diagnosticados con temblor esencial mayores de 18 años, según criterios internacionales [11-13], procedentes de la consulta ambulatoria del Servicio de Neurología del Hospital Universitario de Burgos, entre los meses de mayo de 2015 y diciembre de 2016. El estudio fue aprobado por el comité de ética del complejo hospitalario. Los pacientes fueron informados y firmaron por escrito el consentimiento informado. Se excluyó a los pacientes diagnosticados con otras enfermedades neurológicas, alteraciones psiquiátricas, alcoholismo, deterioro cognitivo y presencia de comorbilidades importantes, y a los que habían recibido tratamiento quirúrgico con estimulación cerebral profunda para el control del temblor. Los pacientes suspendían la medicación para el temblor 24 horas antes de realizar las valoraciones, y tampoco habían tomado bebidas estimulantes ni fármacos con sospecha de interacción con el temblor.

### Evaluaciones

Se registraron datos demográficos (edad, sexo y población) y datos específicos de la enfermedad (edad de inicio, duración de la enfermedad y antecedentes familiares de temblor). Las diferentes escalas clínicas fueron administradas por los investigadores (M.V.C.F. y E.C.P.) tras un proceso de entrenamiento. Se administraron las siguientes escalas: la versión española de la FTM [14] se compone de tres subescalas A, B y C, como se ha descrito anteriormente, con una puntuación total máxima de 156; una mayor

puntuación indica una mayor gravedad [3,4,12,14]. Para valorar la calidad de vida genérica relacionada con la salud, se utilizó el cuestionario *Short Form-36 Health Survey* (SF-36), validado al castellano [15], que incluye 36 ítems de respuesta múltiple, donde se valoran ocho dominios: valoración de la salud, limitación en las actividades físicas, interferencia en el trabajo y en otras actividades diarias, interferencia de los problemas emocionales en el trabajo u otras actividades diarias, rol social o grado en que los problemas de salud física o emocional interfieren en la vida social habitual, intensidad del dolor y su efecto biopsicosocial, sentimiento de energía y vitalidad, y, por último, salud mental general. El rango de puntuación normalizada para la población general es de 0 a 100, y un mayor valor indica un mejor estado de salud. Los valores que se encuentran en la población general son superiores a 50 [16].

Para valorar la discapacidad en miembro superior se usó la escala *Disabilities of the Arm, Shoulder and Hand* (DASH) [17], un cuestionario autoadministrado validado al castellano, que valora el miembro superior como una unidad funcional y permite cuantificar la discapacidad. La escala incluye 30 ítems, puntuados de 1 a 5, y las puntuaciones más altas corresponden a mayor discapacidad.

### Análisis estadístico

Se analizó la distribución normal de las variables cuantitativas usando la prueba de Kolmogorov-Smirnov. Se realizó un análisis descriptivo donde se calcularon las medias  $\pm$  desviaciones estándares y porcentajes de frecuencia de las variables sociodemográficas y clínicas de la muestra incluida, así como de las puntuaciones obtenidas en los cuestionarios DASH, SF-36 y FTM. Se compararon las puntuaciones en las escalas DASH, SF-36 y FTM con el test de Mann-Whitney o *t* de Student según la distribución normal de las variables. Se analizaron los siguientes atributos psicométricos: aceptabilidad, validez de constructo, asunción escalar, fiabilidad y precisión. Los resultados obtenidos se interpretaron conforme a las recomendaciones publicadas para las herramientas de medición de la salud [18].

Para el análisis de la aceptabilidad definida como lo admisible que resulta un instrumento para la población diana y para la apropiada distribución de las puntuaciones de la escala, se tuvieron en cuenta los efectos suelo y techo del rango de las puntuaciones (tomando un máximo del 15% en el rango superior e inferior como satisfactorio), y la asimetría (*skewness*) (tomando como valores de referencia entre  $-1$  y  $+1$ ). Para analizar la validez de construc-

to se realizó un análisis factorial exploratorio de los ítems de la escala FTM, usando el método de componente principal con rotaciones Varimax, previa aplicación del test de esfericidad de Kayser-Meyer-Olkin (KMO) [19] de precisión de muestreo, que debe ser  $> 0,7$ . Además, se usó un análisis de conglomerados (*cluster*), que es una técnica multivariante que busca agrupar variables, tratando de lograr la máxima homogeneidad en cada grupo y la mayor diferencia entre los grupos, usando dendrogramas como representación gráfica. Para el análisis de la asunción escalar, definida como el correcto agrupamiento de los ítems de la escala, se calculó la correlación ítem total corregida (criterio de valor  $\geq 0,4$ ). Para el análisis de la fiabilidad, definida como hasta qué punto la escala está libre de error aleatorio, y existe una interrelación entre sus componentes y las puntuaciones son estables, se calcularon la consistencia interna, determinada según la homogeneidad de los ítems (promedio de los diferentes coeficientes de correlación entre los diferentes ítems, donde se considera aceptable un coeficiente medio de correlación con valor  $\geq 0,30$ ), y el coeficiente  $\alpha$  de Cronbach (donde se considera aceptable un valor  $\alpha \geq 0,70$ ). La precisión o sensibilidad (capacidad de la escala para detectar pequeñas diferencias) se evaluó usando el error estándar de la media [error estándar de la media = desviación estándar  $\times \sqrt{1 - \text{coeficiente de fiabilidad}}$ ], y se consideró adecuado cuando era menor de la mitad de la desviación estándar.

Para el análisis de la validez diagnóstica de discapacidad de la subescala C de la escala FTM con respecto al criterio de referencia (diagnóstico de discapacidad sí/no según el punto de corte establecido por la DASH) [17], se usó la curva ROC (*receiving operating curve*), que proporciona la sensibilidad y la especificidad de la escala FTM para determinar la discapacidad del miembro superior, y el área bajo la curva según la distribución de sus puntuaciones. Para su interpretación se tuvo en cuenta que los valores del área bajo la curva  $< 0,50$  podrían deberse a la casualidad,  $0,51-0,79$  son levemente aceptables,  $0,80-0,89$  son aceptables y  $\geq 0,90$  son excelentes [20]. Todos los análisis estadísticos se realizaron utilizando los programas SPSS v. 20 y MINITAB v. 15. El nivel estadístico de significación se fijó en  $\alpha = 0,05$ .

## Resultados

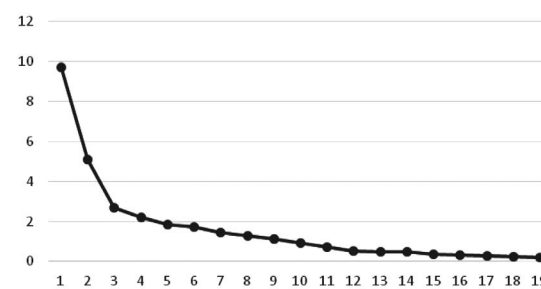
El tamaño muestral de este trabajo se determinó sobre la base de la prevalencia estimada de temblor

**Tabla.** Características demográficas y puntuaciones de las diferentes escalas (media  $\pm$  desviación estándar).

	Hombres	Mujeres	< 68 años	$\geq 68$ años	
Muestra (n = 40)	18 (45%)	22 (55%)	20 (50%)	20 (50%)	
Edad (años)	66,9 $\pm$ 10,45	65 $\pm$ 17,76	55 $\pm$ 13	76 $\pm$ 6,6	
FTM	FTM	23 $\pm$ 9,26	22,2 $\pm$ 13,47	19,5 $\pm$ 9,76	25,6 $\pm$ 12,77
	FTM-A	6,2 $\pm$ 2,83	8 $\pm$ 4,76	6,95 $\pm$ 3,17	7,5 $\pm$ 4,88
	FTM-B	11,5 $\pm$ 5,77	9,2 $\pm$ 5,38	8,5 $\pm$ 5,22	12 $\pm$ 5,54
	FTM-C	5,7 $\pm$ 3	5,1 $\pm$ 4,88	5 $\pm$ 4	6 $\pm$ 4,21
SF-36	Física	69,8 $\pm$ 17,3	65,9 $\pm$ 18,8	69,8 $\pm$ 17,3	65,4 $\pm$ 21,4
	Mental	73,5 $\pm$ 34,6	64,2 $\pm$ 42,6	64,7 $\pm$ 39,9	67,8 $\pm$ 38,6
DASH	Sin discapacidad	41,7 $\pm$ 6,6 (15; 83,3%)	40,1 $\pm$ 11,1 (16; 72,7%)	38,1 $\pm$ 6,9 (14; 70%)	43,9 $\pm$ 7,7 (17; 8%)
	Con discapacidad	79 $\pm$ 17,2 (3; 16,7%)	85,3 $\pm$ 14,9 (6; 27,3%)	82,3 $\pm$ 18,1 (6; 30%)	85 $\pm$ 10,4 (3; 15%)

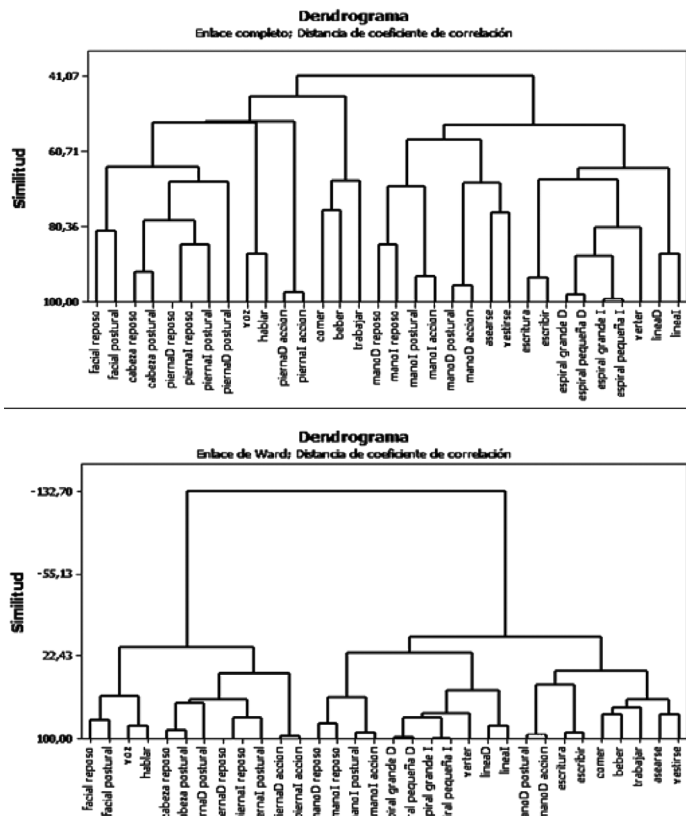
DASH: *Disabilities of the Arm, Shoulder and Hand*; FTM: Fahn-Tolosa-Marín; FTM-A: valoración de la amplitud del temblor; FTM-B: puntuación del temblor en el miembro superior en tareas de escritura, delinear y verter un líquido; FTM-C: puntuación de la incapacidad funcional por el temblor para realizar las actividades de la vida diaria; SF-36: *Short Form-36 Health Survey*. No se encontraron diferencias estadísticamente significativas cuando se compararon las puntuaciones de la FTM, la DASH y el SF-36 entre hombres y mujeres, y participantes  $< 68$  o  $\geq 68$  años.

**Figura 1.** Gráfico de sedimentación: muestra los valores propios (*eigenvalues*) con el análisis factorial de la escala Fahn-Tolosa-Marín.

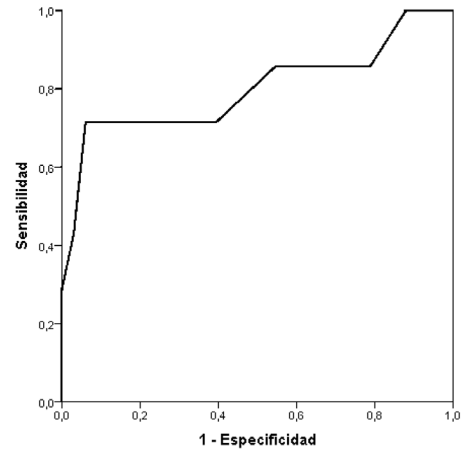


esencial en ancianos españoles, del 5% [21], y se limitó a la localización de los pacientes sujetos al estudio, dentro del área de influencia del Hospital Universitario de Burgos. El tamaño muestral de este estudio comprendió, por lo tanto, aproximadamente al 3% de la estimación de la población total afectada por temblor esencial en la provincia de Burgos. Este criterio de selección del tamaño muestral es si-

**Figura 2.** Dendrogramas que muestran agrupaciones entre los diferentes ítems de la escala Fahn-Tolosa-Marín.



**Figura 3.** Curva ROC de la escala Fahn-Tolosa-Marín, subescala C.



milar al de otros estudios incluidos para validación de escalas clínicas [22-24]. De los 109 pacientes diagnosticados de temblor esencial, se excluyó a 79 pacientes, 26 (24%) de ellos porque no cumplían los criterios de inclusión; 31 (28%) porque no fueron localizados y otros 12 (11%) porque no quisieron participar en el estudio una vez informados. Se incluyó a 40 pacientes con una edad media de  $65,8 \pm 14,7$  años (rango: 21-90 años), 22 (55%) mujeres y 18 (45%) hombres con temblor esencial. Se estratificó la muestra en dos grupos, según la mediana de edad:  $< 68$  años y  $\geq 68$  años. Las características sociodemográficas y clínicas se muestran en la tabla. Según la escala DASH, 31 sujetos (77,5%) no presentaban discapacidad en el miembro superior, y 9 (22,5%) presentaban discapacidad. No se encontraron valores perdidos en ninguna escala administrada. La descripción y la comparación de las principales características clínicas y demográficas de la población incluida se pueden ver en la tabla.

No se encontraron diferencias estadísticamente significativas en las puntuaciones totales de las escalas FTM, SF-36 y DASH entre hombres y mujeres, y participantes  $< 68$  años frente a  $\geq 68$  años en el test de Mann-Whitney ( $p > 0,05$ ). En términos de aceptabilidad para toda la muestra incluida, la escala FTM no mostró efecto suelo (2,5%) ni efecto techo (2,5%), con una puntuación media de  $22,5 \pm 11,4$ . La distribución de las puntuaciones resultó asimétrica, ligeramente fuera de rango (1,40).

El análisis factorial de la escala FTM identificó nueve factores que explicaban el 84% de la varianza, de los cuales los tres primeros explicaban el 55% según el gráfico de sedimentación (Fig. 1). El factor I agrupó los ítems con especial peso de los ítems de la subescala C, espirales y verter (subescala B), y temblor de la mano en reposo (subescala A). El factor II incluyó ítems del temblor procedentes de la subescala A, concretamente el temblor en la cabeza y las piernas tanto en reposo como postural y en acción. El factor III incluyó ítems de la subescala B y la mano derecha con temblor de acción (subescala A).

Los dendrogramas (Fig. 2) mostraron dos agrupaciones jerárquicas distinguibles a un nivel de similitud similar para ambas ramas: ítems relacionados con el temblor localizado en la cabeza y las extremidades inferiores (subescala A), y resto de ítems, que se dividieron a su vez en otras dos ramas que representan, por un lado, agrupaciones de ítems relacionados con la discapacidad (subescala C), y por otro, ítems relacionados con el temblor en el miembro superior dominante (ítem 5 de la subescala A) y la funcionalidad (subescala B).

La escala FTM total demostró una fiabilidad alta con un índice  $\alpha$  de Cronbach de 0,90, desglosada en sus tres subescalas: A ( $\alpha = 0,85$ ), B ( $\alpha = 0,91$ ) y C ( $\alpha = 0,77$ ). Asimismo, demostró una adecuada consistencia interna con índices de correlación  $r \geq 0,40$  entre las puntuaciones totales de las diferentes subescalas A, B y C de la escala FTM. Además, la escala FTM presentó una precisión del error en las medidas aceptable (error estándar de la media = 3,60) y una validez de contenido del test de esfericidad de KMO satisfactoria: 0,70 en los grupos A y B, y  $\geq 0,77$  en el apartado C.

La validez convergente entre las escalas DAHS y la subescala C de la escala FTM estableció una correlación significativa  $r \geq 0,69$ , y la escala FTM global con la escala DASH, de 0,37-0,86. La curva ROC mostró que la subescala C de la escala FTM permite diagnosticar una discapacidad con un área bajo la curva de 0,86 (intervalo de confianza al 95%: 0,67-1,00), estableciendo un punto de corte con una puntuación  $> 5,5$ , con una sensibilidad del 78% y una especificidad del 75% (Fig. 3).

## Discusión

La FTM es una escala de uso general ampliamente aceptada y utilizada en ensayos clínicos, recomendada por la International Parkinson's Disease and Movement Disorder Society [24-28]. Este estudio representa la primera evaluación formal de la escala FTM en España siguiendo los principios y métodos de la teoría clásica de la prueba [18]. Este estudio de validación de las tres subescalas que componen la FTM demostró que, desde un punto de vista clinimétrico, la escala en castellano es válida en términos de calidad y aceptabilidad de los datos, en una muestra representativa sin diferencias en términos de edad ni sexo. La consistencia interna fue satisfactoria dentro de los valores y criterios, similares al estudio original [3,4]. En general, la escala FTM mostró una adecuada interrelación entre los ítems de la escala, una falta de error aleatorio, y, por tanto, una alta precisión. En términos de validez de discapacidad, nuestros resultados apoyan que la subescala C con un punto de corte  $> 5,5$  permite identificar participantes con discapacidad del miembro superior con una sensibilidad y una especificidad aceptables.

Hemos usado la traducción al castellano de la escala FTM implementada por el Grupo de Estudios de Trastornos del Movimiento de la Sociedad Española de Neurología, pero la validez transcultural en otros países hispanoparlantes se desconoce.

En relación con las limitaciones de este estudio, hemos de reconocer que, al igual que en otros estudios [3,12,13], se ha incluido una muestra de pacientes predominantemente afectada por un temblor esencial leve-moderado, limitando por tanto la extrapolación de los resultados que se desprenden en este trabajo a los sujetos con temblor grave. Asimismo, no hemos analizado otras características clinimétricas importantes, como la estabilidad temporal, y la fiabilidad intraevaluador e interevaluador, características ya estudiadas en su versión inglesa [2].

En conclusión, la escala FTM es un instrumento fiable, válido y preciso para la valoración del temblor esencial en la población española adulta. No obstante, se necesitan más estudios con mayor tamaño muestral que confirmen los resultados de este trabajo en pacientes con temblor grave asociado a alta discapacidad.

## Bibliografía

1. Jankovic J, Fahn S. Physiologic and pathologic tremors. Diagnosis, mechanism and management. *Ann Intern Med* 1980; 93: 460.
2. Fahn S, Tolosa E, Marín C. Clinical rating scale for tremor. In Jankovic J, Tolosa E, eds. *Parkinson's disease and movement disorders*. 2 ed. Baltimore: Williams & Wilkins; 1993. p. 225-34.
3. Elble RJ, Bain P, Forjaz MJ, Haubenberger D, Testa C, Goetz CG, et al. Task force report: scales for screening and evaluating tremor: Critique and recommendations. *Mov Disord* 2013; 28: 1793-800.
4. Stacy MA, Elble RJ, Ondo WG, Wu SC, Hulihan J. Assessment of interrater and intrarater reliability of the Fahn-Tolosa-Marín Tremor Rating Scale in essential tremor. *Mov Disord* 2007; 30: 833-8.
5. Perestelo-Pérez L. Estimulación cerebral profunda en trastornos neurológicos y psiquiátricos. Informes de evaluación de tecnologías sanitarias SESCO. Sta. Cruz de Tenerife: Ministerio de Sanidad, Servicios Sociales e Igualdad. Servicio Canario de la Salud; 2013.
6. Carvajal A, Centeno C, Watson R, Martínez M, Rubiales AS. ¿Cómo validar un instrumento de medida de la salud? *An Sist Sanit Navar* 2011; 34: 63-72.
7. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol* 1993; 46: 1417-32.
8. Bullinger M, Alonso J, Apolone G, Leplege A, Sullivan M, Wood-Dauphinee S, et al. Translating health status questionnaires and evaluating their quality: the IQOLA Project approach. *International Quality of Life Assessment*. *J Clin Epidemiol* 1998; 51: 913-23.
9. Scientific Advisory Committee of the Medical Outcomes Trust. Assessing health status and quality-of-life instruments: attributes and review criteria. *Qual Life Res* 2002; 11: 193-205.
10. Badia X. Sobre la adaptación transcultural de medidas de la calidad de vida relacionada con la salud para su uso en España. *Med Clin (Barc)* 1995; 105: 56-8.
11. Deuschl G, Bain P, Brin M; Ad Hoc Scientific Committee. Consensus statement of the Movement Disorder Society on tremor. *Mov Disord* 1998; 13 (Suppl 3): S2-23.
12. Elble RJ. Diagnostic criteria for essential tremor and differential diagnosis. *Neurology* 2000; 54 (Suppl 4): S2-6.
13. Elble RJ. The Essential Tremor Rating Assessment Scale. *J Neurol Neuromed* 2016; 1: 34-8.

14. Bermejo-Pareja F. Más de cien escalas en neurología. 2 ed. Madrid: Aula Médica; 2008. p. 233-6.
15. Vilagut G, Ferrer M, Rajmil L, Rebollo P, Permanyer-Miralda G, Quintana MJ, et al. El cuestionario de salud SF-36 español: una década de experiencia y nuevos desarrollos. *Gac Sanit (Barc)* 2005; 19: 135-50.
16. Vilagut G, Valderas JM, Ferrer M, Garin O, López-García E, Alonso J. Interpretación de los cuestionarios de salud SF-36 y SF-12 en España: componentes físico y mental. *Med Clin (Barc)* 2008; 130: 726-35.
17. Hervás MT, Navarro-Collado MJ, Peiró S, Rodrigo-Pérez JL, López-Mateu P, Martínez-Tello I. Versión española del cuestionario DASH. Adaptación transcultural, fiabilidad, validez y sensibilidad a los cambios. *Med Clin (Barc)* 2006; 127: 441-7.
18. Martínez-Martín P, Cubo E. Scales to measure parkinsonism. *Handb Clin Neurol* 2007; 83: 289-327.
19. Álvarez-Cáceres R. Estadística aplicada a las ciencias de la salud. Barcelona: Díaz de Santos; 2013.
20. Tape TG. Interpreting diagnostic tests. University of Nebraska Medical Center; 2004. URL: <http://gim.unmc.edu/dxtests/Default.htm> [17.12.2018].
21. Labiano-Fontcuberta A, Benito-León J. Temblor esencial: una actualización. *Med Clin (Barc)* 2013; 140: 128-33.
22. Wenning GK, Tison F, Seppi K, Sampaio C, Diem A, Yekhelef F, et al. Development and validation of the Unified Multiple System Atrophy Rating Scale (UMSARS). *Mov Disord* 2004; 19: 1391-402.
23. Luvizutto GJ, Monteiro TA, Braga GL, Pontes-Neto OM, De Lima Resende LA, Bazan R. Validation of the Scandinavian Stroke Scale in a multicultural population in Brazil. *Cerebrovasc Dis Extra* 2012; 2: 121-6.
24. Gironell A, Martínez-Corral M, Pagonabarraga J, Kulisevsky J. The glass scale: a simple tool to determine severity in essential tremor. *Parkinsonism Relat Disord* 2010; 16: 412-4.
25. Hooper J, Taylor R, Pentland B, Whittle IR. Rater reliability of Fahn's tremor rating scale in patients with multiple sclerosis. *Arch Phys Med Rehabil* 1998; 79: 1076-9.
26. Eible RJ, Pullman SL, Matsumoto JY, Raethjen J, Deuschl G, Tintner R. Tremor amplitude is logarithmically related to 4- and 5-point tremor rating scales. *Brain* 2006; 129: 2660-6.
27. Haubenberger D, Kalowitz D, Nahab FB, Toro C, Ippolito D, Luckenbaugh DA, et al. Validation of digital spiral analysis as outcome parameter for clinical trials in essential tremor. *Mov Disord* 2011; 26: 2073-80.
28. Putzke JD, Whaley NR, Baba Y, Wszolek ZK, Uitti RJ. Essential tremor: predictors of disease progression in a clinical cohort. *J Neurol Neurosurg Psychiatry* 2006; 77: 1235-7.

### Validation study of the Spanish version of the Fahn-Tolosa-Marín scale for essential tremor

**Introduction.** The Fahn-Tolosa-Marín (FTM) tremor rating scale has been widely used in clinics for the estimation of tremor severity. However, a Spanish language version of this scale has still not been formally validated.

**Aim.** To provide support to the validity and reliability of this version of FTM Scale.

**Patients and methods.** A cross-sectional study was conducted on essential tremor patients. Severity was rated using the FTM scale. Upper limb disability was evaluated by terms of Disabilities of the Arm, Shoulder and Hand Scale (DASH), and to health-related quality of life using the Short Form-36 Health Survey (SF-36). Statistical analysis included sample description, reliability (Cronbach's alpha), convergent validity, and discrimination capacity tests (ROC curves).

**Results.** Forty patients with essential tremor (22 women, 18 men) were included, with a mean age of  $65.8 \pm 14.7$  years (range: 21-90 years). Internal consistency of the FTM was high: Cronbach's alpha: 0,90 (subscale A: 0.85; subscale B: 0.91; subscale C: 0.77), and the floor and ceiling effects were negligible. The FTM (subscale C) showed high correlations with DASH, and acceptable diagnostic capacity, with an area under the curve of 0.86 (95% CI: 0.67-1.00), sensitivity 78% and specificity 75% for a cut-off score  $> 5.5$ .

**Conclusion.** The Spanish version of FTM the rating scale is a reliable and valid tool to evaluate disability in patients with essential tremor, and a suitable instrument for use in medical research, as well as in clinical practice.

**Key words.** Clinimetric analysis. Disability. Essential tremor. Fahn-Tolosa-Marín scale. Outcomes. Rating scales. Tremor.