

A. Serra-Mayoral  
J. Peña-Casanova

## Fiabilidad test-retest e interevaluador del Test Barcelona

Sección de Neurología de la Conducta y Demencias  
Servicio de Neurología  
Hospital del Mar (IMAS)  
Barcelona

**Introducción.** El programa integrado de exploración neuropsicológica (Test Barcelona) en su versión abreviada (TBA) es un instrumento de evaluación clínica que contempla un amplio abanico de capacidades neuropsicológicas. Elaborado en el año 1977, ha sido a menudo objeto de estudio tanto de aspectos internos como de validación con otros instrumentos. Los resultados obtenidos avalan el TBA como un eficaz instrumento de detección de deterioro cognitivo.

**Método.** Se estudia la fiabilidad test-retest e interevaluador del TBA en sujetos normales. Muestra: consta de 31 sujetos normales (14 hombres y 17 mujeres) de edades comprendidas entre los 20 y los 65 años. Instrumentos: Test Barcelona Abreviado (TBA), *Mini-Mental Status Examination* (MMSE), *Syndrom Kurztest* (SKT) e *Interview for Deterioration of Daily living in Dementia* (IDDD). Análisis estadístico: estadística descriptiva y coeficientes de correlación intraclass.

**Resultados.** El análisis de la fiabilidad test-retest del Test Barcelona en su versión abreviada obtiene buenas correlaciones tanto para las puntuaciones globales directas ICC = 0,92 (IC 95 %: 0,85-0,96) como para las tipificadas ICC = 0,79 (IC 95 %: 0,61-0,89). Sin embargo, cuando se contempla cada uno de los subtests las correlaciones ponen de manifiesto las características de las pruebas: subtest con nula o escasa variabilidad de puntuaciones y subtest con distribución de puntuaciones (como en el caso de la memoria y las funciones ejecutivas). Respecto a la fiabilidad interevaluador, el índice intraclass obtenido es de ICC = 0,99 (IC 95 %: 0,99-0,99) con un alfa de 0,99, que indica una excelente fiabilidad. En conjunto, el TBA presenta excelentes características psicométricas test-retest y fiabilidad interevaluadores.

**Palabras clave:**  
Test Barcelona Abreviado. Fiabilidad test-retest e interevaluador. Exploración neuropsicológica.

*Neurología* 2006;21(6):277-281

Correspondencia:  
Jordi Peña-Casanova  
Sección de Neurología de la Conducta y Demencias  
Servicio de Neurología  
Hospital del Mar (IMAS)  
Passeig Marítim, 25-29  
08003 Barcelona  
Correo electrónico: jpcasanova@imas.imim.es

Recibido el 18-5-05  
Aceptado el 2-2-06

### Test-retest and interrater reliability of Barcelona Test

**Introduction.** The short version of the Barcelona Test is a clinical tool that covers a wide spectrum of neuropsychological skills. Since its creation in 1977, it has often been subject of study either for its internal aspects or in validation studies with other cognitive tests. The results indicate that the Barcelona Test is a useful instrument in detecting cognitive impairment.

**Method.** The aim of this study is to measure the test-retest reliability and the interrater reliability of the Abbreviated Barcelona Test in normal population. Sample: the sample consists of 31 normal subjects (14 men and 17 women) from 20 to 65 years old. Instruments: Abbreviated Barcelona Test (TBA), Mini-Mental Status Examination (MMSE), Syndrome-Kurztest (SKT) and Interview for Deterioration of Daily living in Dementia (IDDD). Statistical analysis: descriptive statistics and intraclass correlation coefficient.

**Results.** Test-retest reliability shows good intraclass correlation coefficients with both direct ICC = 0.92 (IC 95 %: 0.85-0.96) or typified scores ICC = 0.79 (IC 95 %: 0.61-0.89). Nevertheless, when each of the subtests is considered, correlations clearly show two characteristics: subtest with no or minimal range of scores and subtest with distribution of scores (as is the case of memory and executive function). The intraclass coefficient found for the interrater reliability is ICC = 0.99 (IC 95 %: 0.99-0.99) and an alpha of 0.99 which indicate excellent reliability. As a whole, the abbreviated version of the Barcelona Test has excellent test-retest and interrater reliability psychometric characteristics.

**Key words:**  
Abbreviated Barcelona Test. Test-retest and interrater reliability. Neuropsychological assessment.

### INTRODUCCIÓN

A partir del Test Barcelona original<sup>1</sup> se desarrolló una versión abreviada (TBA)<sup>2,3</sup> con el objetivo de establecer un

perfil neuropsicológico básico con un tiempo ajustado a la práctica clínica. El TBA incluye variables que evalúan funciones como la atención, orientación, lenguaje, gnosias, praxias, memoria y funciones ejecutivas. En el momento de su diseño se estableció la validez de criterio<sup>4</sup>.

En la elaboración de un instrumento psicométrico es importante establecer su fiabilidad<sup>5</sup>. La idea global de la fiabilidad está unida a la precisión de las medidas y a la generalización de las mismas en situaciones similares<sup>5</sup>. El objetivo del presente trabajo es analizar la fiabilidad test-retest e interevaluadores de las puntuaciones globales del TBA en sujetos normales.

El test-retest permite establecer la estabilidad de las medidas obtenidas a partir de los mismos sujetos con el mismo instrumento en dos momentos diferentes. La razón de su selección radica en la simplicidad y bajo coste que supone el test-retest. El mayor inconveniente lo supone la determinación del intervalo de tiempo entre ambas aplicaciones, ya que un intervalo corto puede verse influenciado por los efectos de aprendizaje, mientras que con un intervalo amplio pueden surgir otros inconvenientes, como, por ejemplo, un cambio en el estado cognitivo de los sujetos. El intervalo más común para investigar los efectos de la práctica en un test oscila entre 1 semana y 2 meses, siendo 1 mes el punto medio razonable. Intervalos de esta magnitud son los típicos utilizados en investigaciones test-retest. Normalmente se aconseja un intervalo entre 2-4 semanas<sup>6</sup>; para este trabajo fijamos, de forma arbitraria, un intervalo de 1 mes<sup>8,9</sup>.

La fiabilidad interevaluador consiste en una medida de homogeneidad; indica en qué grado diferentes evaluadores puntuarán de idéntica manera la ejecución de una tarea. En este caso las personas encargadas de hacer las evaluaciones, juicios expertos u observaciones son tratadas como un instrumento de medida, ya que son ellas las que proporcionan las puntuaciones<sup>10</sup>.

## MÉTODOS

### Sujetos

La muestra estuvo formada por 31 sujetos: 14 hombres (45,2%) y 17 mujeres (54,8%) de edades comprendidas entre los 20 y los 65 años, con una media de edad de 41,67 y una desviación típica de 16,33. El 35,5% de la muestra había finalizado el bachillerato elemental y sólo un 3,2% sabía sólo leer y escribir. La mayoría eran diestros y catalanohablantes. Todos ellos fueron voluntarios que fueron reclutados por selección incidente y que dieron verbalmente su consentimiento informado. Se consideraron los siguientes criterios de exclusión: evidencia de trastorno neurológico, antecedentes o condiciones médicas que pudiesen afectar el rendimiento neuropsicológico, trastornos psiquiátricos graves y defectos de visión y/o audición.

### Instrumentos

A todos los sujetos se les administraron conjuntamente el *Mini-Mental State Examination* (MMSE) (Folstein et al., 1975)<sup>11</sup> en la versión NORMACODEM (Blesa et al., 2001)<sup>12</sup>, el *Syndrom Kurztest* (SKT) (Erzigkeit, 1992)<sup>13</sup> y el *Interview for Deterioration of Daily living in Dementia* (IDDD) (Teunisse et al., 1991)<sup>14</sup> en su adaptación castellana (Böhm et al., 1998)<sup>15</sup> con el fin de descartar alteraciones cognitivas y funcionales.

#### Mini-Mental State Examination (MMSE) (Folstein et al., 1975)<sup>11</sup>

Test de cribado cognitivo breve que evalúa diferentes funciones cognitivas: orientación, memoria inmediata y diferida, atención, diferentes aspectos del lenguaje y praxia constructiva. La puntuación máxima es de 30 y establece un punto de corte de 24 para determinar la presencia de deterioro. Constituye el test breve de detección de alteraciones cognitivas más difundido y ha sido objeto de diversos estudios acerca de sus diferentes aplicaciones y su valor diagnóstico<sup>16,17</sup>.

#### Syndrom Kurztest (SKT) (Erzigkeit, 1992)<sup>13</sup>

Test cognitivo breve que consta de nueve subtests que, de forma rápida, avalúan memoria y atención/concentración. De su aplicación resultan dos medidas de puntuación directa (por tiempo y número de omisiones) que se convierten en normalizadas siguiendo unas tablas de conversión según edad y nivel premórbido de inteligencia. En este trabajo se han utilizado las tablas originales extraídas de una población alemana. Consta de cinco formas paralelas, utilizándose para este trabajo la forma E.

#### Interview for Deterioration of Daily living in Dementia (IDDD) (Teunisse et al., 1991)<sup>14</sup>

Escala que evalúa el funcionamiento de una persona en las actividades de la vida diaria a partir de la información aportada por un familiar y/o cuidador. Consta de 33 ítems, 16 de los cuales avalúan el grado de cuidado personal y los 17 restantes las habilidades complejas. Cada ítem se puntúa según la frecuencia con que requiere asistencia para llevar a cabo las actividades evaluadas, contemplándose como marco temporal el mes previo en comparación con la que necesitaba antes de la aparición de los síntomas. Se considera un punto de corte de 36, ya que permite discriminar entre la normalidad y el deterioro cognitivo leve<sup>15</sup>.

#### Test Barcelona Abreviado (TBA) (Peña-Casanova et al., 1997)<sup>2,3</sup>

Batería que consta de 55 ítems que avalúan las funciones neuropsicológicas fundamentales (lenguaje expresivo,

orientación, atención, control mental, comprensión, lectura, escritura, praxis constructiva, memoria, abstracción verbal y abstracción cognitiva). En muchos de los subtests se contempla la variable tiempo, con lo cual se obtiene una puntuación directa y una puntuación por tiempo. Cada puntuación se convierte en percentil, de manera que nos permite obtener un perfil cognitivo normalizado para cinco grupos de individuos dependiendo de su edad y de su grado de escolaridad. Así pues, se obtienen dos puntuaciones: una puntuación directa bruta (rango: 0-110) y otra tipificada (rango: 35-130) según grupo de edad y escolaridad. La puntuación global de la escala tiene una media de 100 y una desviación típica de 15.

## Procedimiento

Para evaluar la fiabilidad test-retest se administró la batería de instrumentos a cada uno de los sujetos en una sola sesión por la misma evaluadora y bajo condiciones ambientales similares. El procedimiento se repitió al cabo de 30 ( $\pm 3$ ) días. Con el fin de analizar la fiabilidad interevaluador se procedió a la grabación de un sujeto que cumplía los criterios DSM-IV y NINCDS-ADRA de enfermedad de Alzheimer en fase CDR-1. El vídeo fue observado por un grupo de 11 evaluadores (psicólogos) entrenados que no conocían el estatus clínico del sujeto.

## Análisis estadístico

Para la estadística descriptiva se tuvieron en cuenta medias y desviaciones estándar. Con el fin de medir tanto la fiabilidad test-retest como interevaluador se calculó el coeficiente intraclase. El estudio estadístico se realizó mediante el programa SPSS-PC, versión 11.0.

## RESULTADOS

Las puntuaciones obtenidas en los tests de selección (MMSE, SKT, IDDD) y las puntuaciones del TBA en sus dos evaluaciones se presentan en la tabla 1. Los resultados de las pruebas de correlación (intraclase) para las puntuaciones globales del TBA se muestran en la tabla 2. Los coeficientes intraclase para cada subtest del TBA se muestran en la tabla 3. En esta tabla se han suprimido los subtests que no muestran variabilidad de puntuaciones (alfa de 0 o de 1). Los subtests suprimidos son los siguientes: conversación, fluencia, orientación personal y espacial, series directas e inversas (respuestas correctas y tiempo), repetición de palabras, denominación de imágenes (respuestas correctas y tiempo), respuesta al denominando (respuestas correctas y tiempo), comprensión lectora de logatomos, mecánica de la escritura, denominación escrita (respuestas correctas y tiempo), gesto simbólico derecho e izquierdo (a la orden e imitación) y las posturas bilaterales. Los índices de correlación para estudiar la fiabilidad interevaluador se muestran en la tabla 4.

Test	N	Media (DE)	Mínimo	Máximo
MMSE	31	28,77 (1,52)	24,00	30,00
SKT	31	1,61 (1,25)	0,00	4,00
IDDD	31	33,00 (—)	33,00	33,00
TBA bruta	31	99,45 (8,17)	68,00	108,00
TBA estandarizada	31	102,77 (9,04)	87,00	124,00
Retest TBA bruta	31	100,19 (7,82)	68,00	109,00
Retest TBA estandarizada	31	104,83 (10,32)	87,00	126,00

DE: desviación estándar; MMSE: *Mini-Mental Status Examination*; SKT: *Syndrom Kurztest*; IDDD: *Interview for Deterioration of Daily living in Dementia*; TBA: Test Barcelona Abreviado.

## DISCUSIÓN

### Características de la muestra

La estadística descriptiva de los tests cognitivos y escalas funcionales (MMSE, SKT, IDDD) aplicadas a los sujetos incluidos en el estudio de fiabilidad test-retest ponen de relieve la normalidad de la población estudiada. La mayoría de los resultados obtenidos por los sujetos del estudio giran alrededor de la media, indicativo de una escasa variabilidad de la muestra (tabla 1). Las puntuaciones medias de los sujetos en el TBA estandarizado son de 102,77 en el test y de 104,83 en el retest, indicando que la población está un poco por encima de la media de 100 de la normalización. Además ningún sujeto presenta puntuaciones por debajo de 85 (una desviación estándar del perfil teórico de normalización). Los pocos puntos de diferencia entre el test y el retest se podrían explicar por factores de aprendizaje de la prueba en población normal. Las puntuaciones brutas siguen, lógicamente, el mismo esquema.

Test-retest	ICC	IC 95%	Significación	Alfa
TBA bruta	0,92	0,85-0,96	<0,001	0,96
TBA estandarizada	0,79	0,61-0,89	<0,001	0,89

IC: intervalo de confianza (límites inferior y superior).

Tabla 3		Índices de correlación intraclase (ICC) para cada subtest del TBA			
Test-retest	Valor ICC	IC 95%	Significación	Alfa	
Orientación en tiempo	0,43	0,11-0,68	0,00	0,64	
Digitos directos	0,76	0,57-0,88	0,00	0,87	
Digitos inversos	0,78	0,60-0,89	0,00	0,88	
Series inversas (tiempo)	0,60	0,33-0,79	0,00	0,75	
Evocación categorial (animales)	0,63	0,36-0,80	0,00	0,80	
Comprensión de órdenes	0,10	-0,24-0,43	0,27	0,35	
Material verbal complejo	0,36	0,02-0,63	0,01	0,54	
Material verbal complejo (tiempo)	0,62	0,35-0,79	0,00	0,76	
Lectura de logatomos	0,47	0,15-0,70	0,00	0,63	
Lectura de logatomos (tiempo)	0,47	0,15-0,70	0,00	0,63	
Lectura de un texto	0,96	0,92-0,98	0,00	0,98	
Comprensión lectora de logatomos	0,87	0,75-0,93	0,00	0,93	
Comprensión de frases y textos	0,62	0,35-0,80	0,00	0,76	
Comprensión frases textos (tiempo)	0,91	0,82-0,95	0,00	0,95	
Dictado de logatomos	0,48	0,17-0,71	0,00	0,67	
Dictado de logatomos (tiempo)	0,48	0,17-0,71	0,00	0,67	
Secuencia de posturas (derecha)	0,42	0,09-0,67	0,00	0,58	
Secuencia de posturas (izquierda)	0,63	0,37-0,80	0,00	0,80	
Praxis constructiva	0,93	0,87-0,97	0,00	0,96	
Praxis constructiva (tiempo)	0,76	0,57-0,88	0,00	0,88	
Imágenes superpuestas	0,21	-0,14-0,52	0,12	0,33	
Imágenes superpuestas (tiempo)	0,70	0,46-0,84	0,00	0,83	
Memoria inmediata (evocación)	0,40	0,06-0,65	0,01	0,70	
Memoria inmediata (preguntas)	0,51	0,21-0,73	0,00	0,76	
Memoria diferida (evocación)	0,53	0,22-0,73	0,00	0,70	
Memoria diferida (preguntas)	0,60	0,33-0,78	0,00	0,83	
Memoria visual	0,82	0,66-0,91	0,00	0,93	
Problemas aritméticos	0,84	0,70-0,92	0,00	0,92	
Problemas aritméticos (tiempo)	0,87	0,75-0,93	0,00	0,93	
Semejanzas	0,85	0,72-0,92	0,00	0,93	
Clave de números	0,89	0,80-0,94	0,00	0,94	
Cubos	0,78	0,60-0,89	0,00	0,88	
Cubos-tiempo	0,88	0,77-0,94	0,00	0,89	

\*Se han suprimido los subtests sin variabilidad (alfa de 0 o de 1). IC: intervalo de confianza; TBA: Test Barcelona Abreviado.

### Fiabilidad de las puntuaciones globales

Cuando se contempla el TBA en su conjunto las correlaciones obtenidas tanto para las puntuaciones brutas como para las normalizadas son altas, obteniéndose un índice de correlación intraclase (ICC) = 0,92 (IC 95 %: 0,85-0,96) para las directas y un ICC = 0,79 (IC 95 %: 0,61-0,89) para las normalizadas. La interpretación del índice ICC más utilizada es la que aparece a continuación: hasta 0,39 (pobre), 0,40-0,59 (notable), 0,60-0,79 (buena) y 0,80-1 (excelente)<sup>17</sup>. Consecuentemente los datos obtenidos en este estudio se sitúan en el rango de excelentes.

Tabla 4		Índice de correlación intraclase (ICC) para la fiabilidad interevaluador			
Medida	ICC	Intervalo de confianza		Valor de F	Significación
		Inferior	Superior		
Evaluador aislado	0,994	0,991	0,996	1893,6925	<0,0001
Promedio de los evaluadores*	0,999	0,999	0,999	1893,6925	<0,0001

Grados de libertad para los F-tests: 54 y 540. Valor test: 0. Coeficientes de fiabilidad: número de casos: 55 (variables); número de ítem: 11 (evaluadores). Alfa: 0,9995.

Cabe comentar que en literatura reciente se considera que el comúnmente utilizado índice de correlación de Pearson no es el más adecuado para medir la fiabilidad test-retest, puesto que la *r* de Pearson mide la fuerza de asociación entre las medidas test-retest, no el acuerdo entre ellas<sup>18,19</sup>. Este coeficiente tiende a sobreestimar la fiabilidad ya que es un coeficiente interclase y lo adecuado es uno intraclase. El ICC evalúa el grado de variación entre un tiempo 1 y tiempo 2; si la diferencia es pequeña el ICC será grande; si hay grandes diferencias entonces el ICC será pequeño<sup>20</sup>. El ICC mide la relación entre variables que comparten la misma métrica y varianza. El ICC se calcula como la proporción de la varianza observada atribuible a las diferencias en la puntuación real entre sujetos<sup>21</sup>. Es un coeficiente muy utilizado en la literatura porque es sensible tanto a la aleatorización como a las variaciones entre sujetos<sup>22</sup>.

### Fiabilidad de las puntuaciones de cada subtest

En muchos de los pares de ítems no es posible calcular el coeficiente de correlación puesto que al menos uno de los elementos es constante. Este hecho se explica porque cuando se aplica el TBA a sujetos normales en una serie de subtests muchos sujetos alcanzan la puntuación máxima («efecto techo») que no varía en las dos evaluaciones, por tanto la diferencia entre puntuaciones es 0. Esta baja variabilidad obligadamente se debe reflejar en índices de correlación bajos. Este fenómeno ya se comenta en los estudios de desarrollo del Test Barcelona<sup>1</sup>. Desde una perspectiva psicométrica este hecho podría ser un síntoma de baja discriminabilidad del subtest, puesto que no permitiría la graduación de las puntuaciones. Sin embargo, y ante la realidad clínica, estos subtests se pueden recodificar, de modo que los «sujetos no patológicos» ofrecerán respuestas íntegras, mientras que los sujetos «patológicos» ofrecerán respuestas incompletas. Esto hecho es típico en el estudio de las afasias<sup>24</sup>.

Las mejores correlaciones se encuentran en aquel grupo de subtests que evalúan funciones más complejas como la memoria y las funciones ejecutivas; en estos ítems más

complejos existe más variabilidad entre los sujetos. De la misma manera en algunos subtests cuando se contempla la variable tiempo ofrecen una buena correlación.

### Fiabilidad interevaluador

Para establecer la fiabilidad interevaluador se calculó el ICC dado que las variables eran de tipo ordinal y continuo y el número de jueces 11. El coeficiente intraclass para la media de evaluadores obtenido fue de 0,99 (IC 95 %: 0,99-0,99) con un alfa de 0,99, lo cual indica una excelente fiabilidad interevaluador.

### Comentarios finales

El presente trabajo permite concluir que la forma abreviada del Test Barcelona presenta una buena fiabilidad test-retest en sujetos normales. En sujetos con lesiones cerebrales la fiabilidad test-retest podría estar influida por factores de evolución (positiva o negativa) del cuadro clínico. La fiabilidad interevaluador se ha estudiado en un paciente afecto de enfermedad de Alzheimer leve (CDR 1). El estudio de esta característica psicométrica en un paciente tiene mayor interés que en un sujeto normal ya que es esperable obtener respuestas anormales, más difíciles de evaluar que las correcciones propias de la normalidad. Finalmente se puede afirmar que este estudio se suma al estudio previo de fiabilidad realizado en la fase de desarrollo del test<sup>1</sup>. En conjunto, el TBA tiene buenas características psicométricas de fiabilidad test-retest e interevaluador.

### BIBLIOGRAFÍA

- Peña-Casanova J. Programa Integrado de Exploración Neuropsicológica Test Barcelona. Normalidad, semiología y patología neuropsicológica. Barcelona: Masson, 1991.
- Peña-Casanova J, Guardia J, Bertran-Serra I, Manero RM, Jarne A. Versión abreviada del Test Barcelona (I): subtests y perfiles normales. *Neurología* 1997;12:99-111.
- Guardia J, Peña-Casanova J, Bertran-Serra I, Manero RM, Meza M, Böhm P, et al. Versión abreviada del Test Barcelona (II): puntuación global normalizada. *Neurología* 1997;12:112-6.
- Peña-Casanova J, Meza M, Bertran-Serra I, Manero RM, Espel G, Martí A, et al. Versión abreviada del Test Barcelona (III): validez de criterio con el ADAS-cog. *Neurología* 1997; 12:117-9.
- Muñiz J. *Psicometría*. Madrid: Universitat, 1996.
- Hulley S, Cummings S. *Diseño de la investigación clínica: un enfoque epidemiológico*. Barcelona: Doyma, 1993.
- Tornatore JB, Hill E, Hammond KW. Automated screening for mild cognitive impairment and Alzheimer's disease. *J Neuropsychiatry Clin Neurosci* 2005;17:98-105.
- Schmidt K, Mattis PJ, Adams J, Nestor P. Test-retest reliability of the Dementia Rating Scale-2: alternate form. *Dement Geriatr Cogn Disord* 2005;20:42-44.
- Viladrich C, Doval E, Prat R, Vall-Llovera M. *Psicometría*. Barcelona: Fundació per la Universitat Oberta de Catalunya, 2002.
- Folstein M, Folstein S, McHugh P. Mini-Mental State: a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;12:189-98.
- Blesa R, Pujol M, Aguilar M, Santacruz P, Bertran-Serra I, Hernández G, et al. Clinical validity of the Mini Mental State for Spanish speaking communities. *Neuropsychologia* 2001;39: 1150-7.
- Erzigkeit H. SKT. Ein Kurztest zur Erfassung von Gedächtnis und Aufmerksamkeitsstörungen. Manual. Beltz, Weinheim, 4. Auflage, 1989.
- Teunisse S, Dreix MM, Crevel H. Assessing the severity of dementia. Patient and caregiver. *Arch Neurol* 1991;48:274-77.
- Böhm P, Peña-Casanova J, Aguilar M, Hernández G, Sol JM, Blesa R. Clinical validity and utility of the interview for deterioration of daily living in dementia for spanish speaking communities. *Int Psychogeriatr* 1998;10:261-70.
- Thal LJ, Grundman M, Golden R. Alzheimer's disease: a correlation analysis of the Blessed Information-Memory-Concentration Test and the Mini-Mental State Exam. *Neurology* 1986;36: 262-4.
- Bermejo F, Gómez T, Morales JM. El Mini-Mental State Examination en la evaluación del deterioro cognitivo y la demencia. En: del Ser T, Peña-Casanova J, editores. *Evaluación neuropsicológica y funcional de la demencia*. Barcelona: Prous, 1994.
- Hager K. Reliability of fatigue measures in an overhead work task: a study of shoulder muscle electromyography and perceived discomfort. Tesis Octubre 31, 2003. Blacksburg VA: Master of Science in Industrial and Systems Engineering.
- Keating J. Unreliable inferences from reliable measurements. *Austral J Physiother* 1998;44:5-10.
- Larsson B, Karlsson S, Eriksson M, Gerdle B. Test-retest reliability of EMG and peak torque during repetitive maximum concentric knee extensions. *J Electromiogr and Kinesiol* 2003;13:281-7.
- Wilk CH, Gold J, Bartko J, Dickerson F, Fenton W, Knable M, et al. Test-retest stability of the repeatable battery for the assessment of neuropsychologist status in schizophrenia. *Am J Psychiatry* 2002;159:838-44.
- Nichols DP. Choosing an intraclass correlation coefficient. *SPSS Keywords* 1998;67.
- Bunkan BH, Moen O, Opjordsmoen S, Ljunggren AE, Friis S. Interrater reliability of the comprehensive body examination. *Physiother Theory Pract* 2002;18:121-9.
- Peña-Casanova J, Böhm P, Villaseñor T, Guardia J, Manero RM. Perfil de afasias del Test Barcelona. En: Peña-Casanova J, editor. *Normalidad, semiología y patología neuropsicológicas*. Barcelona: Masson, 2005; p. 49-58.