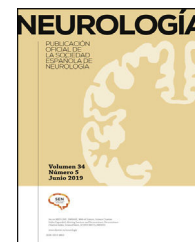




NEUROLOGÍA

www.elsevier.es/neurologia



ORIGINAL

Utilidad diagnóstica de test cognitivos breves en el cribado de deterioro cognitivo

C. Carnero-Pardo^{a,*}, I. Rego-García^b, M. Mené Llorente^c, M. Alonso Ródenas^c
y R. Vélchez Carrillo^b

^a FIDYAN Neurocenter, Granada, España

^b Servicio de Neurología, Hospital Universitario Virgen de las Nieves, Granada, España

^c Centro de Salud Casería de Montijo, Granada, España

Recibido el 1 de febrero de 2019; aceptado el 6 de mayo de 2019

PALABRAS CLAVE

Cribado;
Deterioro cognitivo;
Utilidad diagnóstica;
Test cognitivos
breves;
Atención Primaria

Resumen

Introducción y objetivos: Evaluar y comparar la utilidad diagnóstica (UD) para el cribado de deterioro cognitivo (DC) de los test cognitivos breves (TCB) recomendados por la *Guía de práctica clínica sobre la atención integral a las personas con enfermedad de Alzheimer y otras demencias*.

Material y métodos: Estudio de fase III de evaluación de pruebas diagnósticas en el que se ha incluido en Atención Primaria a sujetos con sospecha de DC. A todos se les ha aplicado Mini-Mental State Examination (Mini-Mental), Mini Examen Cognoscitivo (MEC), Short Portable Mental Status Questionnaire (SPMSQ), Memory Impairment Screen (MIS), test del reloj (TdR), Eurotest, Fototest y test de alteración de memoria (T@M). El diagnóstico de DC se ha realizado de forma independiente y cegada con respecto a los resultados de los TCB. La UD se ha evaluado mediante el área bajo la curva ROC (aROC).

Resultados: Se ha incluido a 141 sujetos (86 con DC). El Eurotest y el T@M ($0,91 \pm 0,02$ [aROC \pm EE] y $0,90 \pm 0,02$, respectivamente), los instrumentos que requieren más tiempo ($7,1 \pm 1,8$ [media \pm DE] y $6,8 \pm 2,2$ min, respectivamente) tienen una UD significativamente superior a la del Mini-Mental, MEC, SPMSQ y TdR, pero no a la del MIS y Fototest ($0,87 \pm 0,03$ ambas), requiriendo este último menos de la mitad del tiempo ($2,8 \pm 0,8$ min). T@M y MIS solo evalúan memoria y el último no es aplicable a analfabetos.

Conclusiones: Los instrumentos más recomendables para el cribado de DC en Atención Primaria son Eurotest, T@M y Fototest, siendo el último más eficiente por requerir la mitad de tiempo.

© 2019 Sociedad Española de Neurología. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

* Autor para correspondencia.

Correo electrónico: ccarnero@neurocenter.es (C. Carnero-Pardo).

<https://doi.org/10.1016/j.nrl.2019.05.007>

0213-4853/© 2019 Sociedad Española de Neurología. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Cómo citar este artículo: Carnero-Pardo C, et al. Utilidad diagnóstica de test cognitivos breves en el cribado de deterioro cognitivo. Neurología. 2019. <https://doi.org/10.1016/j.nrl.2019.05.007>

KEYWORDS

Screening;
Cognitive
impairment;
Diagnostic
performance;
Brief cognitive test;
Primary care

Diagnostic performance of brief cognitive tests in cognitive impairment screening

Abstract

Introduction and objectives: This study aims to assess and compare the diagnostic performance of brief cognitive tests for cognitive impairment (CI) screening recommended by the Spanish guidelines for the integral care of people with Alzheimer's disease and other dementias. **Material and methods:** We performed a phase III study into the accuracy of diagnostic tests, including patients with suspected CI in a primary care setting. All patients completed the Mini-Mental State Examination (MMSE), the Mini Examen Cognoscitivo (MEC), the Short Portable Mental Status Questionnaire (SPMSQ), the Memory Impairment Screen (MIS), the Clock Drawing Test (CDT), the Eurotest, the Fototest, and the Memory Alteration Test (M@T). CI was diagnosed independently by researchers blinded to scores on these tests. Diagnostic performance was evaluated by calculating the area under the receiver operating characteristic curve (AUC).

Results: The study included 141 individuals (86 with CI). The Eurotest and M@T ($AUC \pm SE$: 0.91 ± 0.02 and 0.90 ± 0.02 , respectively) took longer to administer (mean [SD]: 7.1 [1.8] and 6.8 [2.2] min, respectively) and have significantly better diagnostic performance compared to the MMSE, MEC, SPMSQ, and CDT, but not compared to MIS or Fototest (both with an AUC of 0.87 ± 0.03), with the latter taking less than half as long to administer (2.8 [0.8] min). The M@T and MIS only evaluate memory, and the latter cannot be administered to illiterate people. **Conclusion:** The most advisable tests for CI screening in primary care are the Eurotest, M@T, and Fototest, with the latter being the most efficient as it takes half as long to administer.

© 2019 Sociedad Española de Neurología. Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introducción

El deterioro cognitivo (DC) en general y la enfermedad de Alzheimer en particular son uno de los más importantes desafíos sociosanitarios a los que se enfrenta la sociedad, tanto por su frecuencia como por su relevancia y coste¹. En nuestro país, la prevalencia de DC en la población se sitúa entre el 14,5 y el 19% en mayores de 65 años²⁻⁴, aunque muchos de estos pacientes no están identificados (más de la mitad en el reciente estudio Gómez de Caso⁴); a pesar de ello, existe un amplio consenso en considerar que en la actualidad no se dan las circunstancias mínimas requeridas para recomendar la realización de un despistaje (*screening*) de DC en sujetos asintomáticos⁵⁻⁹. No obstante, la mayoría de las guías de práctica clínica coinciden en señalar la conveniencia de mantener en Atención Primaria una adecuada actitud de alerta y realizar un cribado con test cognitivos breves (TCB) en los sujetos en los que se plantee la sospecha de DC^{7,8}, sea por la existencia de quejas subjetivas, quejas por parte del cuidador o allegados o por sospecha generada en los propios profesionales.

En España, la *Guía de práctica clínica sobre la atención integral a las personas con enfermedad de Alzheimer y otras demencias* del Sistema Nacional de Salud (G-SNS)¹⁰ promovida por el Ministerio de Sanidad, Política Social e Igualdad, se hace eco de estas recomendaciones y aconseja el uso de TCB en sujetos con sospecha de DC. En concreto, recomendando el uso de la versión española normalizada del Mini-Mental State Examination (Mini-Mental)¹¹ o del Mini Examen Cognoscitivo (MEC)¹² (recomendación A); en segundo término (recomendación B), considera el uso de

otros instrumentos como el test de Pfeiffer (Short Portable Mental State Questionnaire [SPMSQ])¹³, Memory Impairment Screen (MIS)¹⁴, test de los 7 min (T7M)¹⁵, test del reloj (TdR)¹⁶, Eurotest¹⁷, Fototest¹⁸ o el test de alteración de memoria (T@M)¹⁹. La G-SNS no establece recomendaciones concretas sobre puntos de corte o correcciones a realizar y recomienda que la elección entre instrumentos debe realizarse en función del tiempo disponible, de la experiencia clínica, de la disponibilidad de datos normativos y de estudios de validación en la lengua y en el medio en el que se aplicarán (criterio de buena práctica clínica). Existen también otras características a las que no hace alusión la G-SNS, que también pueden ser relevantes a la hora de la selección, como puede ser el que se pueda aplicar a analfabetos, el que evalúe uno o más dominios cognitivos o el coste que pueda tener asociado su uso (tabla 1)²⁰.

Pocos estudios han evaluado y comparado entre sí estos instrumentos en nuestro país²¹⁻²⁴ y por el momento no existe ningún estudio que haya comparado la utilidad diagnóstica (UD) y la efectividad de los distintos TCB recomendados por la G-SNS en condiciones de práctica clínica habitual en la misma muestra de sujetos. Nuestro objetivo es evaluar y comparar la UD y efectividad de las distintas estrategias de cribado recomendadas por la G-SNS utilizando para ello una cohorte prospectiva de sujetos atendidos en Atención Primaria con verificación diagnóstica completa e independiente.

Material y métodos

Diseño. Estudio de fase III de evaluación de pruebas diagnósticas²⁵, con un diseño pareado (a todos los sujetos

Tabla 1 Características de los test cognitivos breves recomendados en la *Guía de práctica clínica sobre la atención integral a las personas con enfermedad de Alzheimer y otras demencias* del Sistema Nacional de Salud

| Tiempo | Test | Aplicables analfabetos | Multidominio | Estudios en AP | Validación DC | Datos normativos | Coste asociado |
|----------|----------|------------------------|--------------|----------------|---------------|------------------|----------------|
| < 5 min | TdR | No | Sí | Sí | Sí | No | No |
| | SPMSQ | Sí | Sí | Sí | No | No | No |
| | Fototest | Sí | Sí | Sí | Sí | Sí | No |
| | MIS | No | No | Sí | Sí | No | No |
| > 5 min | MMSE/MEC | No | Sí | Sí | Sí | Sí | Sí |
| | T@M | Sí | No | Sí | Sí | No | No |
| | Eurotest | Sí | Sí | Sí | Sí | Sí | No |
| > 10 min | T7M | No | Sí | No | No | Sí | Sí |

MEC: Mini Examen Cognoscitivo; MIS: Memory Impairment Screen; MMSE: Mini-Mental State Examination; SPMSQ: Short Portable Mental State Questionnaire; TdR: test del reloj; T7M: test de los 7 min; T@M: test de alteración de memoria;

se les aplican todos los TCB) y con verificación completa (todos los sujetos, independientemente de los resultados en los TCB, son tratados mediante el procedimiento diagnóstico completo descrito en Procedimientos).

Ámbito. La selección se realizó en 4 Centros de Salud del Distrito Metropolitano del Área Granada Norte (Almanjáyar, Casería de Montijo, Cartuja y Salvador Caballero). El estudio diagnóstico se llevó a cabo en la Unidad de Neurología Cognitivo-Conductual del Servicio de Neurología del Hospital Universitario Virgen de las Nieves.

Población de estudio. Sujetos atendidos en los Centros de Salud en los que se plantea la sospecha de DC.

Criterios de inclusión y exclusión. Se ha incluido de forma consecutiva y sistemática durante el periodo de un año a todos los sujetos atendidos en los que se planteó la sospecha de DC por parte de su médico de Atención Primaria; esta sospecha podía estar basada en la presencia de quejas subjetivas de pérdida de memoria o alteración cognitiva, quejas similares planteadas por un familiar o informador, o bien apreciación, por parte del profesional, de signos o síntomas que sugerían esta sospecha. Se excluyó del estudio a aquellos sujetos que no prestaron consentimiento, que habían participado previamente en el estudio o tenían un DC previo; ni la edad, ni la existencia de déficits sensoriales, motores o cualquier otra condición previa, incluidas las enfermedades neurológicas, han sido motivo de exclusión.

Procedimiento. A todos los sujetos seleccionados se les aplicó de forma balanceada todos los TCB recomendados por la G-SNS excepto el T7M (Mini-Mental, MEC, SPMSQ, MIS, TdR, Eurotest, Fototest y T@M). La orientación temporal y espacial fue evaluada tan solo una vez y se aplicaron los resultados correspondientes a los distintos instrumentos que la contienen (Mini-Mental, MEC, SPMSQ, T@M). La puntuación del MEC fue derivada; para ello, tras la aplicación del Mini-Mental, se aplicaron los ítems específicos del MEC y se recalcularon los resultados correspondientes a las formas con 30 y 35 ítems (MEC-30 y MEC-35, respectivamente). Para la corrección del TdR se utilizó el sistema utilizado en el T7M¹⁵. Se controló el tiempo de aplicación de Eurotest, Fototest, TdR, T@M y Mini-Mental, en los 2 últimos sin incluir los ítems de orientación.

Independientemente de los resultados en los TCB, todos los sujetos fueron remitidos a la consulta de Neurología, donde se llevó a cabo una evaluación conductual (adaptación española del NPI²⁶), funcional (índice Barthel²⁷, escala

de Lawton-Brody²⁸ y FAQ de Pfeffer²⁹) y neuropsicológica detallada (orientación, atención/función ejecutiva (*span* dígitos, semejanzas, fluidez verbal semántica³⁰), memoria verbal (lista de palabras del CERAD³¹), lenguaje (test de denominación de Boston abreviado³², fluidez verbal semántica³⁰ y comprensión órdenes), praxias motoras³³ y visuoespaciales (copia de dibujos del CERAD³¹); esta evaluación se realizó por profesionales independientes de los que habían aplicado los TCB y que desconocían los resultados de los mismos; ninguno de los TCB aplicados formó parte de la evaluación neuropsicológica formal. El tiempo máximo transcurrido entre la aplicación de los TCB y la evaluación en la consulta de Neurología fue de 2 semanas.

Diagnóstico de referencia. El diagnóstico *gold standard* lo realizaron de forma independiente 2 neurólogos experimentados en Neurología cognitivo-conductual que, con base en una evaluación clínica detallada y en los resultados de la evaluación cognitiva, conductual y funcional, clasificaron a los sujetos como: no DC (NoDC), DC ligero (DCL) y demencia (DEM). Se usaron los criterios del Grupo de Estudio de Neurología y Conducta y Demencias de la Sociedad Española de Neurología³⁴ para DCL y DEM, incluidos también en la G-SNS. Estos profesionales estaban cegados con respecto a los resultados de los TCB. Las discrepancias se resolvieron por consenso.

Análisis de datos. La UD se ha evaluado mediante el área bajo la curva ROC (aROC) para cada uno de los TCB; la comparación entre aROC se ha realizado mediante el método de Long de comparación de aROC procedentes de la misma muestra³⁵; el nivel de significación se estableció en $p \leq 0,05$. La efectividad se ha calculado para DC (DCL y DEM) vs. NoDC. Se han calculado la sensibilidad (S) y la especificidad (E) para el mejor punto de corte de cada instrumento (aquel que maximiza la proporción de aciertos diagnósticos; en caso de igualdad entre varios puntos, se optó por aquel entre ellos que maximizaba el índice de Jouden³⁶), así como para aquellos otros que maximizan la E teniendo una $S \geq 0,80$. El análisis estadístico se ha llevado a cabo con el programa SPSS (SPSS Inc., Chicago, IL, EE. UU.) y MedCalc 18.9³⁷.

Aspectos formales. El estudio fue aprobado por el Comité de Ética e Investigación del Hospital Universitario Virgen de las Nieves y se obtuvo el consentimiento informado de todos los participantes. El diseño del estudio y la redacción del informe se ajustan a las recomendaciones STARD de estudios de pruebas diagnósticas³⁸ y a las recomendaciones para

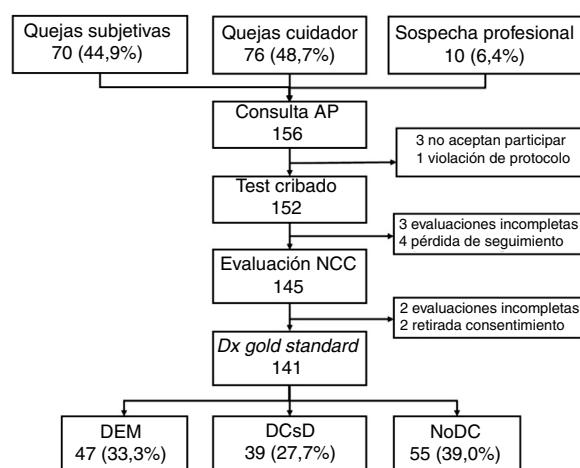


Figura 1 Diagrama de flujo de participantes en el estudio. AP: Atención Primaria; DEM: demencia; DCsD: deterioro cognitivo sin demencia; NCC: Neurología Cognitivo-Conductual; NoDC: no deterioro cognitivo.

describir resultados de estudios diagnósticos de la Food and Drug Administration³⁶.

Resultados

Se seleccionó a un total de 156 sujetos (70 [44,9%] por quejas cognitivas subjetivas, 76 [48,7%] por quejas del cuidador y

10 [6,4%] por sospecha directa del profesional), de los cuales 3 no consintieron participar en el estudio y uno fue excluido por violación de protocolo (tenía un diagnóstico previo de DC); 4 no acudieron a la consulta de Neurología, 2 retiraron el consentimiento a lo largo del estudio, 3 no pudieron ser evaluados por déficits sensoriales y secuelas de ictus previos, y 2 no colaboraron en la evaluación por alteraciones conductuales (fig. 1).

En la tabla 2 se muestran las características sociodemográficas de la muestra final formada por 141 sujetos (55 [39,0%] NoDC, 39 [27,7%] DCL y 47 [33,3%] DEM), con una edad de $72,0 \pm 11,5$ (31-90) años (media \pm DE [rango]), con claro predominio de mujeres (73,0%) y bajo nivel educativo (14,2% analfabetos y solo la mitad [49,6%] con estudios primarios o superiores), así como los resultados de los distintos TCB y el tiempo empleado en su aplicación, estratificados por el diagnóstico cognitivo. Los resultados muestran, en todos los casos y como era de esperar, una diferencia muy significativa entre grupos en el sentido de una peor puntuación y un mayor tiempo en los sujetos con DC.

Todos los pacientes (141) completaron el Mini-Mental y el MEC; 140 completaron el Eurotest y el SPMSQ, y 139 el Fototest y el T@M. El MIS no pudo ser realizado por 22 sujetos (20 de ellos analfabetos); el TdR solo fue completado por 126 sujetos ya que 15 analfabetos se negaron a hacerlo.

En la tabla 3 se resumen los resultados de los distintos instrumentos con indicación del número de sujetos en los que se aplicó, el tiempo que precisó su administración, la UD expresada como aROC, los parámetros diagnósticos (S, E) y la proporción de aciertos diagnósticos correspondiente a los puntos de corte seleccionados.

Tabla 2 Características sociodemográficas y resultados en los test cognitivos breves por grupos diagnósticos

| | Total | DC (DCL + DEM) | NoDC | DCL | DEM |
|---------------------------|-----------------|-----------------|-----------------|----------------|-----------------|
| N.º sujetos | 141 | 86 | 55 | 39 | 47 |
| Sexo (mujeres) | 103 (73,0) | 59 (68,6) | 44 (80,0) | 23 (59,0) | 36 (76,6) |
| Edad (años) | $72,0 \pm 11,5$ | $77,1 \pm 7,5$ | $64,6 \pm 12,2$ | $74,3 \pm 7,3$ | $68,4 \pm 11,6$ |
| Nivel educativo | | | | | |
| Analfabetos | 20 (14,2) | 19 (22,1) | 1 (1,8) | 5 (12,8) | 14 (29,8) |
| Primarios | 51 (36,2) | 35 (40,7) | 16 (29,1) | 16 (41,0) | 19 (40,4) |
| Secundarios | 69 (49,6) | 32 (37,2) | 38 (69,1) | 18 (50,0) | 14 (29,8) |
| MMSE (141) | $19,9 \pm 5,7$ | $17,3 \pm 5,3$ | $24,1 \pm 3,2$ | $21,5 \pm 3,7$ | $13,7 \pm 3,7$ |
| Tiempo (min) ^a | $5,5 \pm 1,6$ | $5,9 \pm 1,5$ | $4,8 \pm 1,4$ | $5,8 \pm 1,7$ | $6,0 \pm 1,4$ |
| MEC (141) | $25,4 \pm 7,2$ | $22,1 \pm 7,2$ | $30,6 \pm 2,9$ | $27,4 \pm 4,8$ | $17,8 \pm 5,9$ |
| Eurotest (140) | $20,2 \pm 10,3$ | $14,7 \pm 9,2$ | $28,6 \pm 4,4$ | $22,3 \pm 6,0$ | $8,3 \pm 6,1$ |
| Tiempo (min) | $7,1 \pm 1,8$ | $7,7 \pm 1,5$ | $6,2 \pm 1,9$ | $7,7 \pm 1,7$ | $7,8 \pm 1,3$ |
| T@M (139) | $29,0 \pm 12,5$ | $22,3 \pm 11,0$ | $39,3 \pm 6,0$ | $30,9 \pm 8,3$ | $15,2 \pm 7,2$ |
| Tiempo (min) ^a | $6,8 \pm 2,2$ | $7,6 \pm 2,0$ | $5,6 \pm 1,9$ | $6,7 \pm 1,9$ | $8,4 \pm 1,9$ |
| Fototest (139) | $29,1 \pm 7,7$ | $25,4 \pm 6,8$ | $34,7 \pm 5,0$ | $30,0 \pm 4,6$ | $21,4 \pm 5,8$ |
| Tiempo (min) | $2,8 \pm 0,8$ | $2,9 \pm 0,8$ | $2,5 \pm 0,7$ | $2,5 \pm 0,5$ | $3,3 \pm 0,9$ |
| TdR (126) | $4,4 \pm 2,6$ | $3,1 \pm 2,6$ | $6,2 \pm 1,4$ | $4,7 \pm 2,2$ | $1,6 \pm 2,0$ |
| Tiempo (min) | $1,3 \pm 0,7$ | $1,5 \pm 0,7$ | $1,0 \pm 0,6$ | $1,3 \pm 0,6$ | $1,6 \pm 0,8$ |
| SPMSQ (140) | $2,6 \pm 2,6$ | $3,8 \pm 2,7$ | $0,6 \pm 0,9$ | $1,8 \pm 1,6$ | $5,6 \pm 2,1$ |
| MIS (119) | $4,2 \pm 2,9$ | $2,6 \pm 2,5$ | $6,3 \pm 1,8$ | $4,0 \pm 2,5$ | $1,0 \pm 1,4$ |

Las cifras son número de sujetos (porcentaje) o media \pm DE.

DC: deterioro cognitivo (DCL y DEM); DCL: deterioro cognitivo ligero; DEM: demencia; MEC: Mini Examen Cognoscitivo; MIS: Memory Impairment Screen; MMSE: Mini-Mental State Examination; NoDC: no deterioro cognitivo; SPMSQ: Short Portable Mental State Questionnaire; TdR: test del reloj; T@M: test de alteración de memoria.

^a En el tiempo no están incluidos los ítems de orientación.

Tabla 3 Utilidad diagnóstica y efectividad de los test cognitivos breves

| Test | N.º | Tiempo (min) | aROC | Mejor punto de corte | | Punto de corte con $S \geq 0,80$ | | | |
|----------|-----|------------------------|-------------|----------------------|------|----------------------------------|-------|------|------|
| | | | | PdC | S | E | PdC | S | PdC |
| MMSE | 141 | 5,5 ± 1,6 ^a | 0,85 ± 0,03 | 20/21 | 0,71 | 0,87 | 22/23 | 0,80 | 0,65 |
| MEC-35 | 141 | > 5,5 ± 1,6 | 0,85 ± 0,03 | 27/28 | 0,70 | 0,87 | 30/31 | 0,85 | 0,62 |
| MEC-30 | 141 | > 5,5 ± 1,6 | 0,85 ± 0,03 | 25/26 | 0,81 | 0,78 | 25/26 | 0,81 | 0,78 |
| Eurotest | 140 | 7,1 ± 1,8 | 0,91 ± 0,02 | 24/25 | 0,82 | 0,89 | 24/25 | 0,82 | 0,89 |
| T@M | 139 | 6,8 ± 2,2 ^a | 0,90 ± 0,02 | 32/33 | 0,81 | 0,89 | 32/33 | 0,81 | 0,89 |
| SPMSQ | 140 | — | 0,86 ± 0,03 | 1/2 | 0,74 | 0,85 | 0/1 | 0,88 | 0,54 |
| Fototest | 139 | 2,8 ± 0,8 | 0,87 ± 0,03 | 30/31 | 0,76 | 0,80 | 32/33 | 0,84 | 0,69 |
| TdR | 126 | 1,3 ± 0,7 | 0,83 ± 0,03 | 4/5 | 0,64 | 0,91 | 6/7 | 0,86 | 0,59 |
| MIS | 119 | — | 0,87 ± 0,03 | 5/6 | 0,82 | 0,81 | 5/6 | 0,82 | 0,81 |

aROC: área bajo la curva ROC; E: especificidad; MEC: Mini Examen Cognoscitivo; MIS: Memory Impairment Screen; MMSE: Mini-Mental State Examination; N.º: número de sujetos; PAD: proporción de aciertos diagnósticos; PdC: punto de corte; S: sensibilidad; SPMSQ: Short Portable Mental State Questionnaire; TdR: test del reloj; T@M: test de alteración de memoria.

^a En el tiempo no están incluidos los ítems de orientación.

La comparación de la UD entre los distintos TCB (tabla 4) se ha llevado a cabo de 2 en 2 en las respectivas submuestras de sujetos que completaron ambos TCB. Los resultados muestran que no hay diferencia significativa entre la UD del Eurotest y el T@M, pero ambos tienen una UD significativamente superior a las del Mini-Mental, MEC-30, MEC-35 y SPMSQ; esta diferencia también es significativa en el caso del TdR y T@M, y está en el límite de la significación cuando se compara TdR y Eurotest ($p = 0,08$). No hay diferencia significativa, en cambio, entre la UD del Eurotest y T@M con la del Fototest y MIS.

En la figura 2 se representan los tiempos necesarios para la aplicación de los TCB utilizados (excepto SPMSQ y MIS que no se han controlado). No hay diferencia significativa entre el tiempo de aplicación del Eurotest y T@M ($7,1 \pm 1,8$ vs. $6,8 \pm 2,2$ min, respectivamente), que son claramente superiores al tiempo exigido por el Fototest ($2,8 \pm 0,8$ min) y el TdR ($1,3 \pm 0,7$ min). El tiempo requerido por el Mini-Mental ($5,5 \pm 1,6$ min) está infraestimado porque en él no se incluyen los ítems correspondientes a la orientación temporal y espacial, a pesar de lo cual es un tiempo significativamente superior al del Fototest y el TdR, y sobrepasa claramente los 5 min. El tiempo de aplicación del MEC, al incluir más ítems que el Mini-Mental, debe ser superior al de este.

Discusión

Los resultados de este estudio de fase III de evaluación de pruebas diagnósticas (prospectivo, pareado y con verificación completa) muestran que el Eurotest ($0,91 \pm 0,02$) y el T@M ($0,90 \pm 0,02$), los TCB que requieren más tiempo de aplicación ($7,1 \pm 1,8$ y $6,8 \pm 2,2$ min, respectivamente), cuentan con la mayor UD sin que exista diferencia significativa entre ellas y ambos clasifican correctamente el 84% de la muestra. Ambos instrumentos han sido desarrollados y validados en nuestro país, son aplicables a analfabetos y de uso libre^{17,19}; el T@M tiene los inconvenientes, en cambio, de que no cuenta con un estudio normativo formal y que solo evalúa memoria, por lo que sería insensible al DC que afecte a otros dominios cognitivos con preservación de la memoria. La UD del T@M es muy similar a la mostrada por este instrumento en un reciente estudio realizado en sujetos con quejas de memoria reclutados en Atención Primaria (aROC 0,91), si bien los resultados globales en nuestra muestra son inferiores probablemente por el menor nivel educativo³⁹.

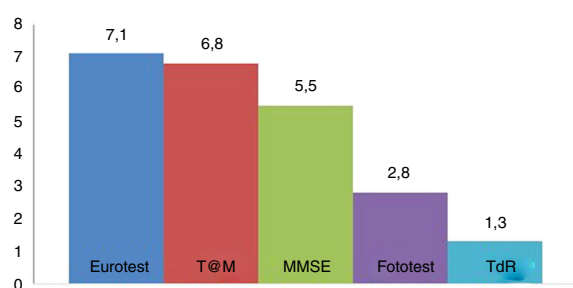
La UD de Mini-Mental, MEC-30 y MEC-35, los instrumentos más utilizados y la recomendación preferente de la G-SNS, que también requieren más de 5 min para ser aplicados (al menos $5,5 \pm 1,6$ min), es significativamente inferior a la de T@M y Eurotest. El bajo rendimiento diagnóstico para DC de estos instrumentos de uso tan universal y generalizado es la norma más que la excepción y viene a sumarse a múltiples estudios de fuera^{40,41} y dentro de nuestro país⁴². Nuestros resultados actuales, en línea con estudios previos de nuestro grupo, también muestran que el punto de corte más efectivo (22/23) es inferior al que habitualmente se utiliza⁴². La baja UD, el tiempo exigido para su aplicación, sumado a otros inconvenientes (influencia nivel educativo, etc.) y, sobre todo, al hecho de que se trata de un TCB cuya aplicación no es gratuita justifica que muchos lo hayan dejado de

Tabla 4 Comparación utilidad diagnóstica de los test cognitivos breves

| | TdR | MEC-35 | MEC-30 | MMSE | SPMSQ | Fototest | MIS | T@M | Eurotest |
|----------|-----|----------|----------|----------|----------|----------|----------|------------|------------|
| TdR | — | ns (126) | ns (126) | ns (126) | ns (126) | ns (126) | ns (117) | 0,03 (125) | 0,08 (126) |
| MEC-35 | | — | ns (141) | ns (141) | ns (140) | ns (139) | ns (119) | 0,02 (139) | 0,02 (140) |
| MEC-30 | | | — | ns (141) | ns (140) | ns (139) | ns (119) | 0,03 (139) | 0,04 (140) |
| MMSE | | | | — | ns (140) | ns (139) | ns (119) | 0,03 (139) | 0,03 (140) |
| SPMSQ | | | | | — | ns (139) | ns (119) | 0,05 (139) | 0,05 (139) |
| Fototest | | | | | | — | ns (119) | ns (138) | ns (138) |
| MIS | | | | | | | — | ns (118) | ns (118) |
| T@M | | | | | | | | — | ns (138) |
| Eurotest | | | | | | | | | — |

MEC: Mini Examen Cognoscitivo; MIS: Memory Impairment Screen; MMSE: Mini-Mental State Examination; ns: no significativo; SPMSQ: Short Portable Mental State Questionnaire; TdR: test del reloj; T@M: test de alteración de memoria.

Los resultados indican el valor de p en el test de deLong y entre paréntesis el número de sujetos en el que se lleva a cabo la comparación.

**Figura 2** Tiempo en minutos requeridos por los test cognitivos breves.

MMSE: Mini-Mental State Examination; T@M: test de alteración de memoria; TdR: test del reloj.

considerar como instrumento de uso preferente⁴³ y otros ni tan siquiera lo incluyen ya entre sus recomendaciones⁴⁴.

También es significativamente inferior la UD del SPMSQ, un instrumento que, sin embargo, requiere mucho menos tiempo de aplicación (no cuantificado en este estudio). El SPMSQ se usa ampliamente en nuestro país para demencia¹³, pero no se había evaluado hasta ahora su UD para DC. Nuestros resultados muestran un marcado fenómeno «suelo», ya que incluso una puntuación de 0 puntos tiene una S 0,88, lo que implica que el 12% de los sujetos con DC tienen una ejecución perfecta (0 errores) y, por tanto, no podrían nunca ser detectados por este instrumento. El SMSPQ tiene además el inconveniente de que la corrección de 3 de sus 10 ítems (número de teléfono, fecha nacimiento y apellido de la madre) exige información previa que muchas veces es difícil de tener o de verificar, en especial en sujetos mayores que pueden acudir solos a consulta.

El TdR, el TCB más breve ($1,3 \pm 0,7$ min), cuenta con la menor UD ($0,83 \pm 0,03$), que también es significativamente inferior, cuenta además con el inconveniente añadido de que exige cierta destreza grafomotora de la que carecen los sujetos analfabetos y con bajo nivel educativo, lo que justifica que 15 sujetos (10,6%), la mayor parte de los analfabetos, rechazara su ejecución. Estos resultados confirman las recomendaciones previas de que no es un TCB adecuado para sujetos que no dispongan de un nivel educativo mínimo⁴⁵.

No hay diferencia significativa, en cambio, entre la UD de T@M y Eurotest y las de Fototest y MIS ($0,87 \pm 0,02$ ambos), 2 instrumentos que precisan mucho menos tiempo para su aplicación ($2,8 \pm 0,8$ min en el caso del Fototest). Estos resultados confirman los de estudios previos^{22,23} y apuntan a que estos instrumentos, al tener similar UD y requerir menos de la mitad de tiempo, son más eficientes. El MIS tiene, en cambio, los inconvenientes de no contar con estudio normativo, evaluar solo memoria y, sobre todo, no ser aplicable a analfabetos, por lo que no podría ser aplicado a una respetable proporción de las personas mayores de nuestro país (14,2% de nuestra muestra).

Hemos elegido la aROC como medida de la UD porque es un parámetro que evalúa la capacidad discriminativa global de una prueba diagnóstica a lo largo de todos sus posibles resultados y permite comparar las UD de los distintos instrumentos en una misma muestra^{46,47}, independientemente de la prevalencia de la condición diagnóstica y de las características sociodemográficas (edad, educación, etc.) de la misma, y por tanto, de su representatividad. También hemos optado por evaluar la UD para DC y no para DEM porque entendemos que es lo más pragmático desde el punto de vista clínico. El interés diagnóstico del profesional en general y del de Atención Primaria en particular ha de ser el DC pues la orientación a DEM impediría detectar casos de DC sin repercusión funcional en los que cabría aplicar medidas preventivas, correctoras o paliativas, que en muchos casos corrigieran esta situación, retrasando o incluso impidiendo su progresión a DEM²⁰.

La principal limitación de este estudio es el hecho de que los TCB se han aplicado de forma seriada, lo que pudiera haber determinado un sesgo debido al cansancio o fatiga del sujeto a pesar de tratarse de instrumentos breves; para contrarrestar en parte este efecto, la aplicación se hizo de forma balanceada según 3 ordenaciones distintas y se ha comprobado que los resultados de los distintos TCB no están asociados a la posición serial de estos (datos no mostrados). Otra limitación es no haber incluido en la evaluación el T7M también recomendado por la G-SNS; no obstante, el T7M⁴⁸ es un instrumento complejo, tanto en su aplicación como corrección, que requiere, a pesar de su nombre, más de 10 min para su aplicación, que no puede ser aplicado a analfabetos y que también tiene asociado un coste; en conjunto

pues, se trata de un instrumento que no creemos indicado para el ámbito de la Atención Primaria²⁰.

La principal fortaleza de este estudio es su diseño. Por un lado, el carácter prospectivo, consecutivo y sistemático del reclutamiento, el hecho de que el periodo de inclusión se haya prolongado un año, la práctica ausencia de criterios de exclusión y el bajo índice de pérdidas en el estudio hacen que nuestra muestra sea un fiel reflejo del problema diagnóstico que se plantea en condiciones de práctica clínica habitual en este entorno asistencial, dotando al estudio de un carácter naturalístico y pragmático. Por otro lado, la evaluación de todos los sujetos con todos los TCB, el hecho de someter a todos los sujetos al protocolo de estudio y diagnóstico *gold standard* independientemente de sus resultados en los TCB, y la evaluación ciega e independiente de los instrumentos evaluados y del *gold standard* aseguran el control de los principales sesgos de los estudios de evaluación de pruebas diagnósticas^{25,49}.

En conclusión, no existe un TCB perfecto e ideal para aplicar en Atención Primaria; los profesionales debieran conocer y manejar varios instrumentos y la elección en un caso determinado debiera realizarse en función de las circunstancias asistenciales (tiempo disponible), las características del sujeto (nivel educativo, déficits sensoriales, etc.), los datos disponibles (datos normativos, validación específica para DC) y las preferencias y experiencia del profesional²⁰. Con base en los resultados de este estudio y a las características deseables resumidas en la [tabla 1](#), nuestra recomendación sería utilizar en Atención Primaria el Eurotest, el T@M o el Fototest, siendo el último más eficiente por requerir menos de la mitad de tiempo.

Financiación

Agencia de Evaluación de Tecnologías Sanitarias, Instituto de Salud Carlos III, España (Expediente. PI06/90034).

Conflicto de intereses

C. Carnero Pardo es el creador del Fototest y Eurotest; ha recibido honorarios por actividades académicas y de asesoría de Nutricia, Schwabe Farma Ibérica, Biogen, Piramal, Janssen Cilag, Pfizer, Eisai, Esteve, Novartis, Lundbeck, Grunenthal y Biogen.

El test de las Fotos o Fototest y el Eurotest están bajo licencia Creative Commons BY-NC-ND; pueden ser usados y difundidos sin fines comerciales, siempre y cuando no sean modificados y se haga reconocimiento expreso de su autoría.

Agradecimientos

Profesionales que han colaborado en el estudio

Hospital Universitario Virgen de las Nieves: C. Sáez Zea, M. Espinosa García, B. Espejo Martínez, L. Montiel Navarro, S. López Alcalde, E. Mora Gavilán.

C.S. Salvador Caballero: F. Padilla Ruiz, P. Concha López, J.A. Henares Civantos, I. Valenzuela López, J.L. Martín Manzano, A. Esteva Rodríguez, M.A. Jurado Duce.

C. S. Almanjáyár: M. Melguizo Jiménez, I. Rodrigo Bravo, B. Martínez Romero, D. Sánchez Mariscal, E. Fernández Román, J.A. Castro Gómez.

C. S. Cartuja: S. Cárdenas Viedma, A.M. de los Ríos Álvarez, J.A. López de Hierro Ruiz.

C. S. Casería de Montijo: F. Romo Serral, M.J. Rodríguez Romero, F. Suárez Pinilla, F. Rodríguez Espinosa, R. Moya Minogorance, V. Molina García, A.M. Zamora Rodrigo, M. Jiménez de la Cruz, C. Romero Molina, A. Cano Carrera, F. Dorador Atienza, J. López Ríos, Y. García Iglesias, C. Pérez Lucena, A. Zambrano Murillo, H. Mahmoud El Awad.

Bibliografía

1. Villarejo Galende A, Eimil Ortiz M, Llamas Velasco S, Llanero Luque M, Lopez de Silanes de Miguel C, Prieto Jurczynska C. Informe de la Fundación del Cerebro. Impacto social de la enfermedad de Alzheimer y otras demencias. Sociedad Española de Neurología. 2017;1–42.
2. Gavrilá D, Antunez C, Tormo MJ, Carles R, Garcia Santos JM, Parrilla G, et al. Prevalence of dementia and cognitive impairment in Southeastern Spain: The Ariadna study. Acta Neurol Scand. 2009;120:300–7.
3. Rodriguez-Sanchez E, Mora-Simon S, Patino-Alonso MC, Garcia-Garcia R, Escribano-Hernandez A, Garcia-Ortiz L, et al. Prevalence of cognitive impairment in individuals aged over 65 in an urban area: DERIVA study. BMC Neurol. 2011;11:147.
4. Vega Alonso T, Miralles Espi M, Mangas Reina JM, Castrillejo Perez D, Rivas Perez AI, Gil Costa M, et al. Prevalence of cognitive impairment in Spain: The Gomez de Caso study in health sentinel networks. Neurología. 2018;33:491–8.
5. Lin JS, O'Connor E, Rossom RC, Perdue LA, Eckstrom E. Screening for cognitive impairment in older adults: A systematic review for the U.S Preventive Services Task Force. Ann Intern Med. 2013;159:601–12.
6. United Kingdom National Screening Committee. The UK NSC recommendation on screening for dementia. London: Public Health England; 2015.
7. Gerontology Society of America. Workgroup on Cognitive Impairment Detection and Earlier Diagnosis. Report and recommendations. Washington DC: The Gerontology Society of America; 2015.
8. Canadian Task Force on Preventive Health Care Pottie K, Rahal R, Jaramillo A, Birtwhistle R, Thombs BD, et al. Recommendations on screening for cognitive impairment in older adults. CMAJ. 2016;188:37–46.
9. Chambers LW, Sivananthan S, Brayne C. Is dementia screening of apparently healthy individuals justified? Adv Prev Med. 2017;2017:9708413.
10. Grupo de trabajo de la Guía de Práctica Clínica sobre la atención integral a las personas con enfermedad de Alzheimer y otras demencias. Guía de práctica clínica sobre la atención integral a las personas con enfermedad de Alzheimer y otras demencias. Plan de calidad para el Sistema Nacional de Salud del Ministerio de Sanidad, Política Social e Igualdad, Agència d'informació, Avaluació i Qualitat en Salut de Catalunya, 2011.
11. Blesa R, Pujol M, Aguilar M, Santacruz P, Bertran-Serra I, Hernandez G, et al. Clinical validity of the "Mini-Mental State" for Spanish speaking communities. Neuropsychologia. 2001;39:1150–7.
12. Lobo A, Saz P, Marcos G, Dia JL, de la Camara C, Ventura T, et al. Revalidación y normalización del Mini-Examen

- Cognoscitivo (primera versión en castellano del Mini-Mental Status Examination) en la población general geriátrica. *Med Clin (Barc)*. 1999;112:767–74.
13. Martínez de la Iglesia J, Duenas Herrero R, Onis Vilches MC, Aguado Taberne C, Albert Colomer C, Luque Luque R. Adaptación y validación al castellano del cuestionario de Pfeiffer (SPMSQ) para detectar la existencia de deterioro cognitivo en personas mayores de 65 años. *Med Clin (Barc)*. 2001;117:129–34.
14. Pérez-Martínez DA, Baztan JJ, González-Becerra M, Socorro A. Evaluación de la utilidad diagnóstica de una adaptación española del Memory Impairment Screen de Buschke para detectar demencia y deterioro cognitivo. *Rev Neurol*. 2005;40:644–8.
15. del Ser Quijano T, Sánchez Sánchez F, García de Yébenes MJ, Otero Puime A, Zúñunegui MV, Muñoz DG. Versión española del Test de los 7 Minutos. Datos normativos de una muestra poblacional de ancianos de más de 70 años. *Neurología*. 2004;19:344–58.
16. Cacho J, García-García R, Arcaya J, Vicente JL, Lantada N. Una propuesta de aplicación y puntuación del test del reloj en la enfermedad de Alzheimer. *Rev Neurol*. 1999;28:648–55.
17. Carnero-Pardo C, Montoro-Ríos MT. Evaluación preliminar de un nuevo test de cribado de demencia (Eurotest). *Rev Neurol*. 2004;38:201–9.
18. Carnero Pardo C, Sáez-Zea C, Montiel Navarro L, del Saz P, Feria Vilar I, Pérez Navarro MJ, et al. Utilidad diagnóstica del Test de las Fotos (FOTOTEST) en deterioro cognitivo y demencia. *Neurología*. 2007;22:860–9.
19. Rami L, Molinuevo JL, Sánchez-Valle R, Bosch B, Villar A. Screening for amnesic mild cognitive impairment and early Alzheimer's disease with M@T (Memory Alteration Test) in the primary care population. *Int J Geriatr Psychiatry*. 2007;22:294–304.
20. Olazarán J, Hoyos-Alonso MC, del Ser T, Garrido Barral A, Condesala JL, Bermejo-Pareja F, et al. Aplicación práctica de los test cognitivos breves. *Neurología*. 2016;31:183–94.
21. Rami L, Bosch B, Valls-Pedret C, Caprile C, Sánchez-Valle Díaz R, Molinuevo JL. Validez discriminativa y asociación del test mini-mental (MMSE) y del test de alteración de memoria (T@M) con una batería neuropsicológica en pacientes con deterioro cognitivo leve amnésico y enfermedad de Alzheimer. *Rev Neurol*. 2009;49:169–74.
22. Carnero-Pardo C, Espejo-Martínez B, López-Alcalde S, Espinosa-García M, Sáez-Zea C, Vilchez-Carrillo R, et al. Effectiveness and costs of Phototest in dementia and cognitive impairment screening. *BMC Neurol*. 2011;11:92.
23. Carnero-Pardo C, Espejo-Martínez B, López-Alcalde S, Espinosa-García M, Sáez-Zea C, Hernández-Torres E, et al. Diagnostic accuracy, effectiveness and cost for cognitive impairment and dementia screening of three short cognitive tests applicable to illiterates. *PLOS ONE*. 2011;6:e27069.
24. Cacho J, Benito-León J, García-García R, Fernández-Calvo B, Vicente-Villardón JL, Mitchell AJ. Does the combination of the MMSE and clock drawing test (mini-clock) improve the detection of mild Alzheimer's disease and mild cognitive impairment? *J Alzheimers Dis*. 2010;22:889–96.
25. Sackett DL, Haynes RB. The architecture of diagnostic research. *BMJ*. 2002;324:539–41.
26. Boada M, Cejudo JC, Tarraga L, López OL, Kaufer D. Neuropsychiatric Inventory Questionnaire (NPI-Q): validación española de una forma abreviada del Neuropsychiatric Inventory (NPI). *Neurología*. 2002;17:317–23.
27. Mahoney FI, Barthel DW. Functional evaluation: The Barthel Index. *Md State Med J*. 1965;14:61–5.
28. Lawton MP, Brody EM. Assessment of older people: Self-maintaining and instrumental activities of daily living. *Gerontologist*. 1969;9:179–86.
29. Pfeffer RI, Kurosaki TT, Harrah CH Jr, Chance JM, Filos S. Measurement of functional activities in older adults in the community. *J Gerontol*. 1982;37:323–9.
30. Carnero Pardo C, Lendínez-González A. Utilidad del test de fluencia verbal semántica en el diagnóstico de demencia. *Rev Neurol*. 1999;29:709–14.
31. Morris JC, Mohs RC, Rogers H, Fillenbaum G, Heyman A. Consortium to establish a registry for Alzheimer's disease (CERAD) clinical and neuropsychological assessment of Alzheimer's disease. *Psychopharmacol Bull*. 1988;24:641–52.
32. Calero MD, Arnedo ML, Navarro E, Ruiz-Pedrosa M, Carnero C. Usefulness of a 15-item version of the Boston Naming Test in neuropsychological assessment of low-educational elders with dementia. *J Gerontol B Psychol Sci Soc Sci*. 2002;57:187–91.
33. Pérez-Marmol JM, López-Alcalde S, Carnero-Pardo C, Canadas-de la Fuente GA, Peralta-Ramírez MI, García-Ríos MC. Creation and design of a test for the Evaluation of Upper Limb Apraxia (EULA) based on a cognitive model: A pilot study. *Rev Neurol*. 2015;60:66–74.
34. Robles A, del Ser T, Alom J, Pena-Casanova J. Propuesta de criterios para el diagnóstico clínico del deterioro cognitivo ligero, la demencia y la enfermedad de Alzheimer. *Neurología*. 2002;17:17–32.
35. DeLong ER, deLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics*. 1988;44:837–45.
36. Statistical guidance on reporting results from studies evaluating diagnostic test: U. S. Department of Health and Human Services, Food and Drug Administration, Center for Devices and Radiological Health, Diagnostic Devices Branch, Division of Biostatistics, Office of Surveillance and Biometrics, 2003.
37. MedCalc Statistical Software versión 18.9.1 [computer program]. Published: MedCalc Software bvba, Ostend, Belgium, 2018.
38. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD initiative. *BMJ*. 2003;326:41–4.
39. Ozer S, Noonan K, Burke M, Young J, Barber S, Forster A, et al. The validity of the Memory Alteration Test and the Test Your Memory test for community-based identification of amnesic mild cognitive impairment. *Alzheimers Dement*. 2016;12:987–95.
40. Pezzotti P, Scalmana S, Mastromattei A, di Lallo D, Progetto Alzheimer Working Group. The accuracy of the MMSE in detecting cognitive impairment when administered by general practitioners: A prospective observational study. *BMC Fam Pract*. 2008;9:29.
41. Mitchell AJ. A meta-analysis of the accuracy of the mini-mental state examination in the detection of dementia and mild cognitive impairment. *J Psychiatr Res*. 2009;43:411–31.
42. Carnero Pardo C, Cruz Orduna I, Espejo Martínez B, Cardenas Viedma S, Torrero García P, Olazarán Rodríguez J. Efectividad del Mini-Mental en la detección del deterioro cognitivo en Atención Primaria. *Aten Primaria*. 2013;45:426–33.
43. Carnero-Pardo C. ¿Es hora de jubilar al Mini-Mental? *Neurología*. 2014;29:473–81.
44. Weintraub S, Besser L, Dodge HH, Teylan M, Ferris S, Goldstein FC, et al. Version 3 of the Alzheimer Disease Centers' Neuropsychological Test Battery in the Uniform Data Set (UDS). *Alzheimer Dis Assoc Disord*. 2018;32:10–7.
45. Carnero-Pardo C, Rego García I, Barrios López J, Blanco Madera S, Calle Calle R, López-Alcalde S, et al. Evaluación de la utilidad diagnóstica y validez discriminativa del Test del Reloj y del Mini-Cog en la detección del deterioro cognitivo. *Neurología*. 2019. <https://doi.org/10.1016/j.nrl.2018.12.002>.

46. Burgueno MJ, Garcia-Bastos JL, Gonzalez-Buitrago JM. Las curvas ROC en la evaluación de las pruebas diagnósticas. *Med Clin (Barc)*. 1995;104:661–70.
47. Aggarwal R, Ranganathan P. Understanding diagnostic tests-Part 3: Receiver operating characteristic curves. *Perspect Clin Res*. 2018;9:145–8.
48. Del Ser T, Sanchez-Sanchez F, Garcia de Yebenes MJ, Otero A, Munoz DG. Validation of the seven-minute screen neurocognitive battery for the diagnosis of dementia in a Spanish population-based sample. *Dement Geriatr Cogn Disord*. 2006;22:454–64.
49. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: A tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res methodol*. 2003;3:25.