

ORIGINAL

Test cognitivos breves como herramienta de decisión en Atención Primaria. Estudio poblacional y de validación



M. Tainta^{a,b,i,*}, A. Iriondo^a, M. Ecay-Torres^a, A. Estanga^a, M. de Arriba^a, M. Barandiaran^{a,c,i}, M. Clerigue^a, M. Garcia-Sebastian^a, J. Villanua^d, A. Izagirre^e, J. Saldias^a, A. Aramburu^f, J. Taboada^{g,1}, J. Múgica^c, A. Barandiaran^{b,2}, A. Arrospide^{h,i}, J. Mar^{h,i,j} y P. Martínez-Lage^a

^a CITA Alzheimer, Donostia-San Sebastián, Gipuzkoa, España

^b Osakidetza, Organización Sanitaria Integrada (OSI) Goierri-Urola Garaia, Zumárraga, Gipuzkoa, España

^c Osakidetza, Organización Sanitaria Integrada (OSI) Donostialdea, Donostia-San Sebastián, Gipuzkoa, España

^d Osatek SA, Hospital Donostia, Donostia-San Sebastián, Gipuzkoa, España

^e Universidad del País Vasco (UPV-EHU), Facultad de Medicina y Enfermería, Donostia-San Sebastián, Gipuzkoa, España

^f Osakidetza, Organización Sanitaria Integrada (OSI) Debabarrena, Mendaro, Gipuzkoa, España

^g Osakidetza, Organización Sanitaria Integrada (OSI) Bilbao Basurto, Bilbao, Bizkaia, España

^h Osakidetza, Organización Sanitaria Integrada (OSI) Debagoiena, Arrasate, Gipuzkoa, España

ⁱ Instituto de Investigación Sanitaria Biodonostia, Donostia-San Sebastián, Gipuzkoa, España

^j Instituto de Investigación en Servicios Sanitarios Kronikgune, Barakaldo, Bizkaia, España

Recibido el 24 de febrero de 2022; aceptado el 24 de mayo de 2022

Accesible en línea el 31 de agosto de 2022

PALABRAS CLAVE

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

Resumen

Introducción y objetivos: Los test cognitivos breves (TCB) se utilizan en Atención Primaria (AP) para la detección de deterioro cognitivo (DC) pero existen pocos datos sobre su utilidad diagnóstica (UD) en el ámbito comunitario.

Este trabajo evalúa la UD de Fototest, T@M, cuestionario AD8 y MMSE en una muestra representativa de la población y aporta nuevos puntos de corte (PdC) que se han validado en un grupo de personas que consultan por quejas cognitivas.

Material y métodos: Ambas muestras, la poblacional y la de validación, se realizó una evaluación en 2 fases; una primera de cribado y administración de los TCB y una segunda de diagnóstico clínico, ciego a los resultados de los TCB, aplicando los criterios NIA-AA actuales.

* Autor para correspondencia.

Correo electrónico: mtainta@cita-alzheimer.org (M. Tainta).

¹ en representación del Grupo de trabajo Proyecto DEBA

² en representación del Grupo de trabajo GOIZ Alzheimer Beasain

◇ Los nombres de los componentes del Grupo de trabajo Proyecto DEBA y del Grupo de trabajo GOIZ Alzheimer Beasain están relacionados en el [anexo 1](#).

La UD de los TCB en la muestra poblacional se evaluó con el área bajo la curva ROC (aROC). Para elegir los PdC óptimos, se evaluaron 2 métodos: el índice de Youden y el PdC con mejor especificidad que asegurase una sensibilidad del 80%. En la muestra de validación se calcularon los parámetros de sensibilidad y especificidad, y los valores predictivos para estos PdC.

Resultados: Se ha incluido a 260 participantes (23,1% con DC) de la muestra poblacional y 177 (42,4% con DC) de la de validación. El Fototest tiene la mejor UD a nivel poblacional (aROC 0,851), que mejora con la combinación de Fototest y AD8 (aROC 0,875; $p < 0,05$). Los PdC propuestos son $AD8 \geq 1$, Fototest ≤ 35 , $T@M \leq 40$ y $MMSE \leq 26$.

Conclusión: Los TCB son útiles en la detección de DC en AP. Este trabajo apoya el uso de PdC más exigentes.

© 2022 Sociedad Española de Neurología. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la CC BY-NC-ND licencia (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

KEYWORDS

Brief cognitive tests;
Primary Care;
Diagnostic utility;
Cognitive impairment

Brief cognitive test as a decision tool in Primary Care settings.

Population and validation study

Abstract

Introduction and objectives: Brief cognitive tests (BCT) are used in primary care (PC) for the detection of cognitive impairment (CI). Still, there are little data on their diagnostic utility (DU) in a community setting. This work evaluates the DU at the population level of Fototest, T@M, AD8 questionnaire and MMSE. It provides new cut-off points (CoP) validated in a CI early detection program.

Material and methods: In the population and validation samples, the evaluation was carried out in 2 phases, a first of screening and administration of BCT and a second of clinical diagnosis, blinded to the results of the BCT, applying the current NIA-AA criteria. The DU of BCT in the population sample was evaluated with the area under the ROC curve (aROC). Youden index and the CoP with the best specificity that ensured a sensitivity of 80% were used to decide on the most appropriate CoP. The sensitivity, specificity, and predictive values for these CoP were calculated in the validation sample.

Results: 260 participants (23.1% with CI) from the population sample and 177 (42.4% with CI) from the validation sample were included. The Fototest has the best UD at the population level (aROC 0.851), which improves with the combination of Fototest and AD8 (aROC 0.875). The proposed CoP are $AD8 \geq 1$, Fototest ≤ 35 , $T@M \leq 40$, and $MMSE \leq 26$.

Conclusion: BCT are helpful in detecting CI in PC. This work supports the use of more demanding PoC.

© 2022 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

A medida que la población está envejeciendo, el potencial aumento de la prevalencia de deterioro cognitivo (DC) supone un importante reto sociosanitario¹. Ya hoy en día, la demencia es la principal causa de dependencia en personas mayores y se estima que dentro de 20 años la población afectada se va a doblar², con las implicaciones sociales y económicas que ello implica.

Las quejas cognitivas son un motivo de consulta frecuente en adultos mayores. La atención correcta y a tiempo del DC es importante para el paciente y para el sistema, permitiendo acceder a asistencia médica, recursos sociales y la toma de decisiones sobre los cuidados a recibir. Sin embargo, el proceso diagnóstico muchas veces se ve dificultado y la mayoría de los pacientes reciben el diagnóstico cuando se encuentran en una fase de demencia³. Desde los servicios de

Atención Primaria (AP) se ha de seguir una estrategia para decidir si un paciente con quejas cognitivas necesita una evaluación más profunda en la atención especializada. Los test cognitivos breves (TCB) son una herramienta primordial de decisión en estos casos^{4,6}.

Sin embargo, los puntos de corte utilizados para determinar cuándo un TCB está alterado proceden en su mayoría de estudios que validan y evalúan la utilidad diagnóstica (UD) de los TCB en ámbitos asistenciales especializados, en los que la prevalencia de DC es mucho mayor que en AP y el estadio clínico en el cual se encuentran los pacientes es generalmente más avanzado^{3,7}. Esto puede afectar al rendimiento esperado, a la validez externa y a la reproducibilidad del TCB al aplicar dichos puntos de corte en AP, donde la prevalencia de DC es más baja, pudiendo resultar en un elevado número de casos falsos negativos. Distintos trabajos^{4,8} proponen la necesidad de actualizar los puntos de corte de los TCB más

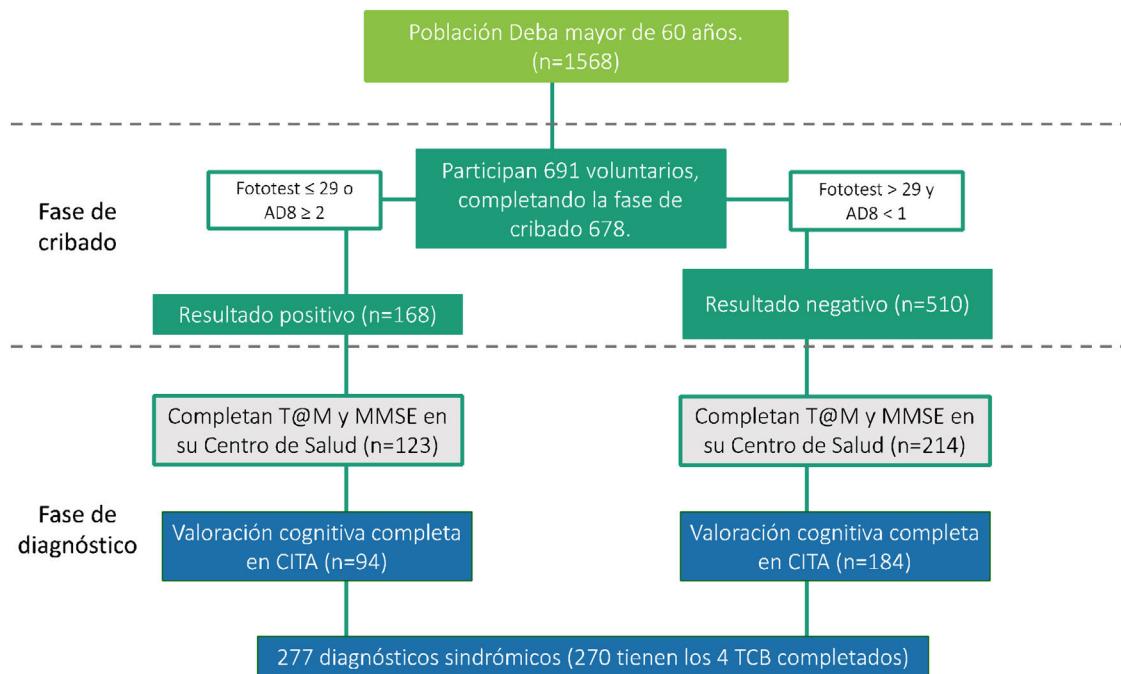


Figura 1 Muestra poblacional. Diseño y flujo de participación en el proyecto STOP Alzheimer 2020-Estudio DEBA.

habituales, como el caso del Mini-Mental State Examination (MMSE) para la detección del DC desde sus fases leves y la necesidad de examinar la precisión diagnóstica de MMSE en el contexto de la vía de diagnóstico experimentada por el paciente⁹. Otros TCB, como el Fototest¹⁰⁻¹², el T@M^{13,14} y el cuestionario AD8¹⁵, han demostrado un buen rendimiento para la detección de casos de DC¹⁶, pero existen pocos datos sobre su uso a nivel poblacional.

En este estudio hemos querido conocer la UD del MMSE y compararla con otros TCB utilizados en nuestro entorno y actualizar los puntos de corte en la poblacional general para, posteriormente, evaluar su aplicación en un contexto clínico no especializado para la detección de casos.

Material y métodos

Diseño

En este estudio se utilizó un diseño de validación de pruebas diagnósticas dividido en 2 partes. La primera se basa en una muestra de voluntarios representativa de la población y se utilizó para evaluar la UD de los TCB y elegir los puntos de corte óptimos. En la segunda, la muestra se compone de personas que consultaban por quejas o preocupación cognitiva y se utilizó para validar la utilidad de los puntos de corte elegidos.

Población de estudio

1. **Muestra poblacional:** la evaluación de la UD de los TCB se basó en la información obtenida del proyecto STOP ALZHEIMER 2020-Proyecto Deba. Se trata de una muestra poblacional obtenida en el estudio epidemiológico

transversal de prevalencia de Deterioro cognitivo leve, de demencia y de enfermedad de Alzheimer en mayores de 60 años llevado a cabo en 2015 en la población de Deba (País Vasco) (fig. 1). Tras realizar una campaña de difusión del proyecto y sensibilización frente al DC, se invitó a participar a todas las personas de la localidad que fueran mayores de 60 años. En la fase de cribado participaron 691 personas; se recogieron datos sociodemográficos, índice de riesgo de demencia CAIDE¹⁷, el cuestionario AD8 y Fototest. Se invitó a participar en una evaluación diagnóstica completa a todos los participantes que dieron un resultado positivo (n = 168) en alguna de las pruebas cognitivas y a una muestra representativa según la edad, sexo y años de educación e índice de riesgo CAIDE de demencia de los voluntarios con un resultado negativo en las pruebas de cribado (tabla suplementaria 1, anexo 2). Iniciaron la fase de diagnóstico en el centro de APde la localidad de Deba 337 participantes (123 con cribado positivo y 214 con cribado negativo), donde se administraron el MMSE y el T@M y se recogieron antecedentes médicos, medicación actual, analítica de sangre. Dicha evaluación la completaron en CITA-Alzheimer 277 voluntarios (94 positivos y 183 negativos).

2. **Muestra de validación:** la validación del estudio se llevó a cabo con los datos del estudio transversal incluido en el desarrollo del programa de atención precoz del DC denominado GOIZ-Alzheimer, llevado a cabo en la localidad de Beasain (País Vasco) en el año 2017. Tras una campaña informativa sobre DC y diagnóstico temprano, a todas las personas mayores de 60 años que tuvieran quejas cognitivas o preocupación por su estado cognitivo se ofreció la posibilidad de realizar una evaluación inicial. Esta evaluación inicial consistió en la recogida de datos sociodemográficos, índice de riesgo de demencia CAIDE y la administración de MMSE, T@M, Fototest y

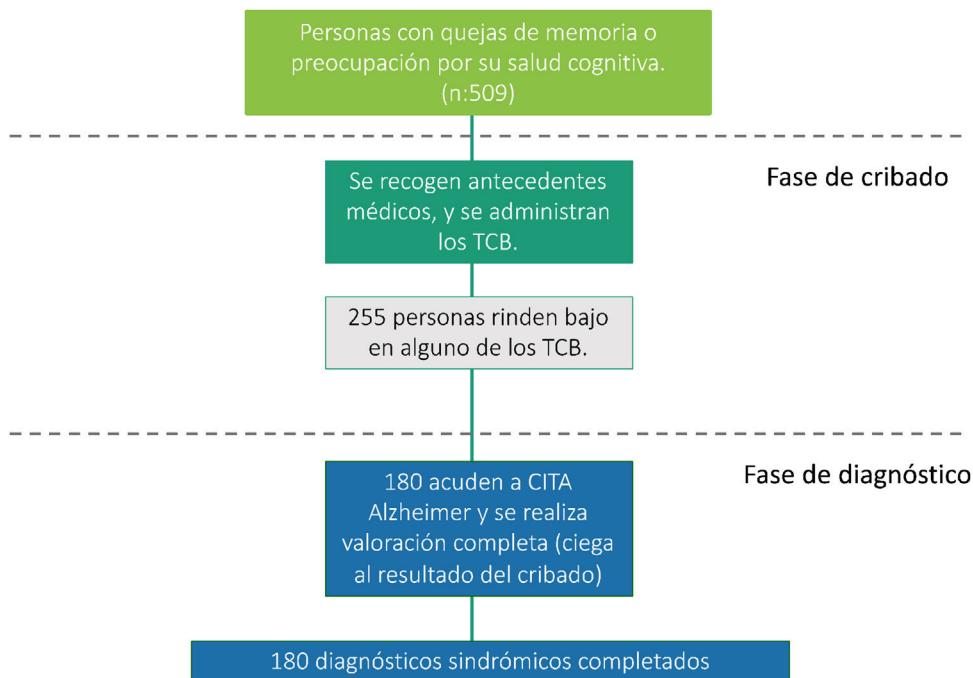


Figura 2 Muestra de validación. Flujo de participación del programa GOIZ Alzheimer.

AD8. A todas las personas con algún resultado alterado en estos TCB se ofreció la posibilidad de realizar una evaluación diagnóstica que incluía una evaluación neurológica y neuropsicológica completa (fig. 2). De las 509 personas con quejas cognitivas que se acercaron al programa, 255 resultaron positivas en el cribado cognitivo y de ellas 180 completaron la evaluación diagnóstica. Entre las personas con cribado negativo no se hizo ninguna evaluación adicional, pero se revisaron sus historias clínicas electrónicas a los 2 años y no se detectaron nuevas consultas por quejas cognitivas ni casos diagnosticados de DC en ese tiempo.

Definición de cribado positivo: presentar una puntuación alterada en alguno de los TCB según los puntos de corte originales.

Variables recogidas en la fase diagnóstica: en la fase diagnóstica de ambas muestras se recogieron las siguientes variables: datos sociodemográficos, evaluación y exploración neurológica completa, exploración física neurológica, exploración del estado mental. La evaluación neuropsicológica estudió los siguientes aspectos: inteligencia premórbida, memoria, lenguaje, capacidad constructiva, capacidad visuoperceptiva, atención/concentración y funciones ejecutivas. Las secuencias obtenidas en la resonancia magnética estructural fueron MPRAGE-T1, T2 axial, Flair axial, difusión axial y SWI axial.

Criterio diagnóstico: los pacientes evaluados en ambos estudios se clasificaron, desde el punto de vista clínico, en 3 grupos; normalidad cognitiva, DC ligero y demencia. El criterio diagnóstico de los casos evaluados se realizó mediante un consenso entre el equipo médico —un médico general, 2 neurólogos y un radiólogo— y de neuropsicología del estudio. En esta valoración el equipo fue ciego al resultado de los TCB. Para el diagnóstico de demencia y DCL se siguieron

los criterios diagnósticos actuales (NIA-AA)¹⁸. Se consideró como caso de DC la presencia de DCL o demencia.

Test cognitivos breves

MMSE: el MMSE es el TCB tradicionalmente más utilizado, sirve como evaluación inicial en muchas consultas de AP y especializada e incluso su puntuación es criterio de inclusión en muchos estudios de investigación. El tiempo de administración es aproximadamente de 10 min; el punto de corte que generalmente se establece para una «cognición normal» es de 24^{5,19}. Existen numerosos trabajos que evalúan su rendimiento poblacional para casos de demencia, pero los datos para la discriminación de DCL son más limitados²⁰.

Test de alteración de la memoria: el T@M es una herramienta de diseño más reciente dirigida a detectar cambios en pacientes con un perfil de DC. El tiempo de administración también es de unos 10 min y su punto de corte es igual o menor de 37. Los datos actuales en nuestro entorno son referentes a unidades especializadas^{13,14}.

Fototest: el tiempo de administración es de 3-4 min y no se ve influido por el nivel cultural. Incluye una prueba de denominación, otra de fluencia verbal y de recuerdo libre y facilitado. Ha demostrado en nuestro medio una efectividad diagnóstica superior a la del MMSE^{21,22}. En este trabajo hemos utilizado el punto de corte de 29 basándonos en revisiones previas¹¹.

Cuestionario AD8: es un cuestionario al que responde un informador fiable con respuestas binarias a 8 preguntas que evalúan los cambios intraindividuo²³, siendo el punto de corte de al menos 2 ítems. El tiempo de administración es de 2-3 min. Ofrece una visión evolutiva del sujeto minimizando los sesgos de edad, sexo, grado educativo o factores socioculturales. El cuestionario AD8 utilizado en el estudio

es su adaptación al castellano, la información recogida que tiene una buena correlación con otras escalas funcionales de amplio uso como la escala Clinical Dementia Rating scale¹⁵.

Combinación Fototest y AD8: el uso combinado de la puntuación de ambos test se ha evaluado de forma conjunta¹⁵, mejorando la capacidad diagnóstica de cada uno por separado. Este uso combinado se puede simplificar sustituyendo la puntuación del AD8 a la obtenida en el Fototest (Fototest-AD8). El punto de corte estimado para considerar la prueba positiva es de 26 y se puede estimar de forma sencilla sustituyendo a la puntuación del Fototest, la obtenida en el AD8.

Protocolo y consentimiento informado: las memorias del Proyecto STOP-ALZHEIMER 2020 Proyecto Deba como del programa GOIZ Alzheimer obtuvieron el informe favorable del Comité de Ética de Euskadi (números PI2015153 y PI2016178, respectivamente). Todos los participantes firmaron un consentimiento informado antes de su inclusión en el estudio.

Análisis estadístico

En el análisis se incluyó a todos los participantes de los que se disponía el valor de los 4 TCB y de la evaluación diagnóstica completa. En primer lugar, se realizó un análisis descriptivo de ambas muestras. En segundo lugar, se calcularon las características diagnósticas de cada TCB para la muestra poblacional: sensibilidad, especificidad, valores predictivos para los puntos de corte previamente recomendados y el área bajo la curva en las curvas ROC (aROC)²⁴. Se realizó el mismo análisis para los grupos por encima y por debajo de 75 años para evaluar cómo afecta al rendimiento diagnóstico de las pruebas el aumento de la prevalencia de DC. Para estimar los mejores puntos de corte se exploraron 2 metodologías; la primera seleccionó el punto de corte con mayor especificidad que asegurase una sensibilidad del 80% y el segundo utilizó el índice Youden ($J = \text{especificidad} + \text{sensibilidad} - 1$)²⁵. El índice de Youden se calcula para cada uno de los puntos de la curva ROC y tiene un valor de -1 a 1 . El valor de 1 significa que no existen falsos positivos ni negativos. Por ello, para la elección del punto de corte óptimo según el índice de Youden, hemos elegido el que hace máximo su valor. Para cada valor del índice de Youden hay una sensibilidad y una especificidad asociada al punto de corte para el que se calcula. La validación de estos puntos de corte se realizó evaluando los valores de sensibilidad, especificidad y valores predictivos de estos puntos de corte en la muestra de validación. El cambio en el aROC obtenidas con los diferentes TCB se analizó mediante la prueba de DeLong²⁶. La prueba De Long se llevó a cabo con el software estadístico para la investigación biomédica MedCalc (Healthcare Technology, Nueva York, EE. UU.).

Resultados

Se incluyó a un total de 437 participantes sumando los de ambos estudios, de los cuales se dispuso de los datos sociodemográficos, los resultados de los 4 TCB analizados y la valoración neurológica completa, incluyendo una

exploración neuropsicológica formal de su estado cognitivo ([tabla 1](#)).

De la *muestra poblacional* se incluyó a 260 personas (94 de las cuales fueron positivas en la fase de cribado), de quienes se tenía el valor de los 4 TCB y se había realizado el diagnóstico clínico de consenso. La edad media fue de 69,9 años, el 51,15% fueron mujeres y la media de años de escolarización fue de 9,85. El 33,46% de la muestra presentó un riesgo elevado (≥ 9 puntos) de demencia medido mediante el índice de riesgo de demencia CAIDE. El porcentaje de participantes con DC en cualquiera de sus grados fue del 23,10%. En el grupo de edad mayor de 75 años casi la mitad de la muestra presentó DC, una cifra mayor a la del grupo menor de 75 años, que fue del 15,42%.

La [tabla 2](#) resume los parámetros de UD (aROC, sensibilidad, especificidad y valores predictivos positivo y negativo) de cada TCB en la muestra poblacional. Dado que por encima de 75 años los casos de DC fueron más prevalentes, se analizó la UD de las pruebas en ambos grupos de edad. Con respecto a la UD, el Fototest fue el TCB con mejor capacidad de discriminación diagnóstica, tanto en toda la muestra como entre los menores de 75 años. Entre los mayores de 75 años, el test con mejor capacidad diagnóstica fue el cuestionario AD8. La combinación de Fototest-AD8 es la prueba con mayor aROC de forma significativa con respecto a su aplicación de forma aislada. En la [figura 3a](#) y b se muestra estos datos de forma gráfica. Los valores de sensibilidad, especificidad y valores predictivos se presentan ([tabla 2](#)) para los puntos de corte originales, los puntos de corte que mejor especificidad tuvieron siempre que se asegurase una sensibilidad superior al 80% y los que considera óptimos el índice de Youden. Los puntos de corte que aseguran una sensibilidad superior al 80% para la muestra global son MMSE ≤ 28 , AD8 ≥ 1 , T@M ≤ 44 , Fototest ≤ 35 y Fototest-AD8 ≤ 33 . Los puntos de corte que considera óptimos el índice de Youden para la muestra global son, al igual que con la metodología anterior, de 35 para el Fototest, de 1 para el cuestionario AD8 y de 33 para Fototest-AD8. Sin embargo, para el MMSE lo sitúa en 26 y en el T@M lo coloca en 40.

La *muestra de validación* fue de 177 participantes. Su edad media fue algo mayor (73,5 años) y el 41,20% fueron mujeres. Los años de educación también se situaron en una cifra algo inferior a la del proyecto DEBA (8,19) y el 58,19% de los participantes presentaron un riesgo elevado de demencia medido con el CAIDE. El porcentaje de personas con DC es del 42,40%. También el grupo de edad mayor de 75 años tiene más personas con deterioro, 63,16%, frente al 26,73% de los menores de 75 años.

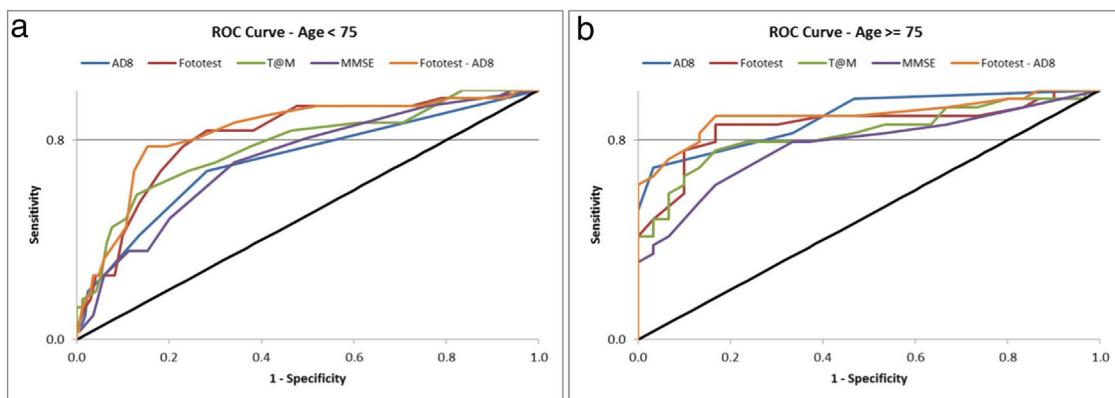
La [tabla 3](#) resume los parámetros de UD (sensibilidad, especificidad, valores predictivos) de cada TCB en la muestra de validación para los puntos de corte obtenidos en la muestra poblacional que aseguran una sensibilidad del 80% y los puntos de corte que hacen máximo el índice de Youden.

Discusión

El presente trabajo ha evaluado la capacidad diagnóstica de 4 pruebas cognitivas breves a partir de la obtención de datos de referencia obtenidos a nivel poblacional y su validación en una muestra clínica. El estudio sugiere unos puntos

Tabla 1 Resultados descriptivos de la muestra de la muestra poblacional y de la muestra de validación

	Muestra poblacional		Muestra de validación		p-valor
	N	%	N	%	
Total (%)	260	100,00%	177	100,00%	
Edad media (DE)	69,9 (7,9)		73,5 (6,9)		< 0,001
Edad					
< 75 años	201	77,31%	101	57,10%	
≥ 75 años	59	22,69%	76	42,90%	< 0,001
Sexo					
Hombres	127	48,85%	104	58,80%	
Mujeres	133	51,15%	73	41,20%	0,050
Años de educación					
Media (DE)	9,85 (3,54)		8,19 (1,85)		< 0,001
1-5 años	14	5,38%	29	16,38%	
6-10 años	145	55,77%	112	63,28%	
11-15 años	75	28,85%	28	15,82%	
Más de 15 años	19	7,31%	8	4,52%	
Desconocido	7	2,69%	0	0,00%	< 0,001
CAIDE					
< 9 puntos	162	62,31%	74	41,81%	
≥ 9 puntos	87	33,46%	103	58,19%	
Desconocido	11	4,23%	0	0,00%	< 0,001
Diagnóstico clínico					
DCL/demencia	60	23,10%	75	42,40%	
Normal	200	76,90%	102	57,60%	< 0,001
< 75 años					
DCL/demencia	31	15,42%	27	26,73%	
Normal	170	84,58%	74	73,27%	0,021
≥ 75 años					
DCL/demencia	29	49,15%	48	63,16%	
Normal	30	50,85%	28	36,84%	0,117

**Figura 3** a) aROC de los TCB en la muestra de la población de DEBA < 75 años. b) aROC de los TCB en la muestra de la población de DEBA ≥ 75 años.

de corte distintos a los utilizados previamente basados en muestras clínicas especializadas. El Fototest fue el TCB que mejor discriminó entre personas con y sin DC. Además, su uso combinado con el cuestionario de síntomas cognitivos AD8 mejoró su rendimiento como se ha sugerido previamente¹⁵. Los resultados de este trabajo pueden contribuir a mejorar la detección del DC en AP. Los datos que se aportan en cuanto a UD y puntos de corte ayudarán a los profesionales

sanitarios a elegir la mejor herramienta de cribado inicial de las personas con quejas cognitivas.

Este trabajo ha examinado la UD de los TCB en la población general a diferencia de los estudios que lo han hecho en atención especializada, por lo que los puntos de corte son más exigentes. La contextualización de este estudio en la población general permite defender que estos resultados y los puntos de corte que se proponen son más adecuados para

Tabla 2 S, E, VPP, VPN y aROC de cada TCB en la muestra poblacional

	aROC	Puntos de corte originales				Punto de corte con $S \geq 80\%$				Punto de corte según índice de Youden							
		c	Sensi- bilidad	Especi- ficidad	VPP	VPN	c ^a	Sensi- bilidad	Especi- ficidad	VPP	VPN	Youden index	c ^a	Sensi- bilidad	Especi- ficidad	VPP	VPN
<i>Total</i>																	
AD8	0,809	≥ 2	61,70%	83,50%	52,90%	87,90%	≥ 1	81,70%	69,00%	44,10%	92,60%	0,507	≥ 1	81,70%	69,00%	44,10%	92,60%
Fototest	0,851 ^c	≤ 29	48,50%	96,00%	80,50%	85,00%	≤ 35	83,30%	74,50%	49,50%	93,70%	0,578	≤ 35	83,30%	74,50%	49,50%	93,70%
Fototest- AD8	0,875	≤ 26	38,3%	98,5%	88,5%	84,2%	≤ 33	83,30%	81,50%	57,50%	94,20%	0,648	≤ 33	83,30%	81,50%	57,50%	94,20%
T@M	0,822	≤ 37	56,70%	92,00%	68,00%	87,60%	≤ 44	85,00%	58,00%	37,80%	92,80%	0,518	≤ 40	68,30%	83,50%	55,40%	89,80%
MMSE	0,763	≤ 23	40,00%	93,00%	63,20%	83,80%	≤ 28	83,30%	48,00%	32,50%	90,60%	0,408	≤ 26	63,30%	77,50%	45,80%	87,60%
<i>Edad < 75 años</i>																	
AD8	0,718	≥ 2	41,90%	86,50%	36,10%	89,10%	≥ 1	67,7% ^b	71,80%	30,40%	92,40%	0,395	≥ 1	67,70%	71,80%	30,40%	92,40%
Fototest	0,813 ^c	≤ 29	16,10%	97,10%	50,00%	86,40%	≤ 36	83,90%	71,80%	35,10%	96,10%	0,557	≤ 36	83,90%	71,80%	35,10%	96,10%
Fototest- AD8	0,833	≤ 26	12,90%	98,80%	66,70%	86,20%	≤ 35	80,60%	74,10%	36,20%	95,50%	0,621	≤ 33	77,40%	84,70%	48,00%	95,40%
T@M	0,777	≤ 37	38,70%	93,50%	52,20%	89,30%	≤ 45	83,90%	53,50%	24,80%	94,80%	0,452	≤ 40	58,10%	87,10%	45,00%	91,90%
MMSE	0,721	≤ 23	25,80%	94,10%	44,40%	87,40%	≤ 28	80,60%	50,60%	22,90%	93,50%	0,369	≤ 27	71,00%	65,90%	27,50%	92,60%
<i>Edad ≥ 75 años</i>																	
AD8	0,891	≥ 2	82,80%	66,70%	70,60%	80,00%	≥ 2	82,80%	66,70%	70,60%	80,00%	0,657	≥ 3	69,00%	96,70%	95,20%	76,30%
Fototest	0,859 ^c	≤ 29	72,40%	90,00%	87,50%	77,10%	≤ 33	86,20%	83,30%	83,30%	86,20%	0,695	≤ 33	86,20%	83,30%	83,30%	86,20%
Fototest- AD8	0,900	≤ 26	65,50%	96,70%	95,00%	74,40%	≤ 31	89,70%	83,30%	83,90%	89,30%	0,730	≤ 31	89,70%	83,30%	83,90%	89,30%
T@M	0,827	≤ 37	75,90%	83,30%	81,50%	78,10%	≤ 41	82,80%	53,30%	63,20%	76,20%	0,592	≤ 37	75,90%	83,30%	81,50%	78,10%
MMSE	0,778	≤ 23	55,20%	86,70%	80,00%	66,70%	≤ 27	82,80%	46,70%	60,00%	73,70%	0,460	≤ 25	79,30%	66,70%	69,70%	76,90%

^a Punto de corte (criterio de positividad).^b No existe punto de corte con sensibilidad mayor de 0,80.^c Diferencia estadísticamente significativa ($p < 0,05$) con respecto al aROC de la variable Fototest-AD8 mediante el test DeLong.

Tabla 3 Sensibilidad, especificidad, valores predictivos positivo y negativo de cada test en la muestra de validación con los puntos de corte óptimos según ambas metodologías

	Punto de corte con Sensibilidad $\geq 80\%$ obtenido en la muestra poblacional					Punto de corte según índice de Youden obtenido en la muestra poblacional				
	c ^a	Sensibilidad	Especificidad	VPP	VPN	c ^a	Sensibilidad	Especificidad	VPP	VPN
Total										
AD8	≥ 1	74,70%	52,90%	53,80%	74,00%	≥ 1	74,70%	52,90%	53,80%	74,00%
Fototest	≤ 35	84,00%	54,90%	57,80%	82,40%	≤ 35	84,00%	54,90%	57,80%	82,40%
Fototest-AD8	≤ 33	82,70%	58,80%	59,60%	82,20%	≤ 33	82,70%	58,80%	59,60%	82,20%
T@M	≤ 44	100,00%	4,90%	43,60%	100,00%	≤ 40	97,30%	17,60%	46,50%	90,00%
MMSE	≤ 28	96,00%	24,50%	48,30%	89,30%	≤ 26	73,30%	71,60%	65,50%	78,50%
<i>Edad < 75 años</i>										
AD8	≥ 1	66,70%	52,70%	34,00%	81,30%	≥ 1	66,70%	52,70%	34,00%	81,30%
Fototest	≤ 36	74,10%	47,30%	33,90%	83,30%	≤ 36	74,10%	47,30%	33,90%	83,30%
Fototest-AD8	≤ 35	74,10%	51,40%	35,70%	84,40%	≤ 33	66,70%	64,90%	40,90%	84,20%
T@M	≤ 45	100,00%	5,40%	27,80%	100,00%	≤ 40	96,30%	21,60%	31,00%	94,10%
MMSE	≤ 28	96,30%	28,40%	32,90%	95,50%	≤ 27	88,90%	52,70%	40,70%	92,90%
<i>Edad >= 75 años</i>										
AD8	≥ 2	58,30%	71,40%	77,80%	50,00%	≥ 3	37,50%	78,60%	75,00%	42,30%
Fototest	≤ 33	85,40%	42,90%	71,90%	63,20%	≤ 33	85,40%	42,90%	71,90%	63,20%
Fototest-AD8	≤ 31	83,30%	42,90%	71,40%	60,00%	≤ 31	83,30%	42,90%	71,40%	60,00%
T@M	≤ 41	100,00%	7,10%	64,90%	100,00%	≤ 37	91,70%	14,30%	64,70%	50,00%
MMSE	≤ 27	85,40%	39,30%	70,70%	61,10%	≤ 25	64,60%	78,60%	83,80%	56,40%

VPN: valor predictivo negativo; VPP: valor predictivo positivo.

^a Punto de corte (criterio de positividad).

su uso en el entorno en la AP y en la evaluación inicial de las quejas cognitivas. La muestra seleccionada es representativa en cuanto a edad y sexo de población general de nuestro entorno geográfico ([figuras suplementarias 1 y 2, anexo 2](#)).

Por otro lado, el diagnóstico de presencia o ausencia de DC se ha realizado sobre la base de un protocolo diagnóstico extenso y una evaluación neuropsicológica completa de todas las áreas cognitivas, que ha permitido la aplicación de los criterios diagnósticos de la NIA-AA. Una fortaleza adicional del estudio es la validación de los puntos de corte a nivel poblacional en una muestra de personas que representa el tipo de paciente que consulta en AP por quejas cognitivas consulta en AP.

Los valores de sensibilidad y especificidad, pero sobre todo los valores predictivos dependen de la prevalencia de la enfermedad, en este caso, de DC, en la población en la que se usa. En el caso de los TCB evaluados, los puntos de corte propuestos y disponibles en la literatura se han establecido sobre muestras seleccionadas en consultas especializadas^{5,11,13} en las que la prevalencia de DC es más alta que en la población general que se atiende en AP. Dado que los TCB tienen su principal uso en la atención inicial a las personas con quejas cognitivas, en nuestra opinión era necesario investigar si los puntos de corte que se obtienen en la población general son iguales a los que están disponibles en la literatura. Los resultados de este estudio demuestran que no. Los puntos de corte que se obtienen en la población general son más altos y, por tanto, más exigentes. Esto tiene importancia si se tiene en cuenta que uno de los principales objetivos de una estrategia de diagnóstico inicial en AP evitar obtener un test negativo en una persona con enfermedad. Tal y como se ha mostrado, la aplicación en la muestra

poblacional de los puntos de corte originales demostró una gran especificidad, pero una baja sensibilidad ([tabla 2](#)) y un mayor número de falsos negativos.

En este trabajo hemos evaluado 2 métodos para estimar los mejores puntos de corte en una estrategia de detección de casos en AP. El primero, elige el punto de corte con mejor especificidad que asegura una sensibilidad de al menos de un 80%; el segundo método utilizado es el índice de Youden, que resulta de la combinación de la sensibilidad y especificidad, dando valor a las clasificaciones correctas de la prueba²⁵. Para el Fototest y el AD8, ambos métodos eligen los mismos puntos de corte. Sin embargo, para el T@M y el MMSE, para poder asegurar una sensibilidad del 80%, los valores de especificidad se resienten mucho. Este hecho se encontró en los datos del Proyecto Deba (muestra poblacional) y fue corroborado en la muestra del programa GOIZ Alzheimer (muestra de validación). De esta forma, con los datos obtenidos, nos inclinaríamos por elegir los puntos de corte que hacen máximo el índice de Youden (AD8 ≥ 1 , Fototest ≤ 35 , T@M ≤ 40 y MMSE ≤ 26) para la detección del DC en AP, que tienen mejores valores de especificidad manteniendo una buena sensibilidad.

La observación de que la prevalencia de DC es distinta según el grupo de edad, justificó la evaluación de la UD de los TCB en ambos grupos de edad. Así, en el en el grupo de edad menor de 75 años estos puntos de corte deberían ser aún más exigentes dado que la prevalencia de deterioro es todavía menor. En el caso de los mayores de 75 años, el cuestionario AD8 fue el TCB que mejor rindió de forma no combinada, pero, al mismo tiempo, fue el que tuvo peores resultados en los menores de 75 años. Este hallazgo puede reforzar la idea de que, en personas mayores, el mero hecho

de que un informador fiable haya notado un cambio en la cognición del paciente es un dato muy importante de cara al diagnóstico final.

La principal limitación de este trabajo fue la dificultad logística de que conseguir que todos los participantes cumplimentasen procedimientos de cribado y también diagnósticas, ocasionando así pues pérdidas. Para conseguir el mayor número de participantes, el equipo investigador se desplazó al municipio de residencia para hacer las evaluaciones de cribado. Así mismo, se consiguió que las características sociodemográficas fueran similares entre las personas que acuden a la evaluación en CITA-Alzheimer y las que no ([tabla suplementaria 1, anexo 2](#)). Por otro lado, en la muestra de validación, para minimizar el hecho de que las personas con rendimientos normales en la fase de cribado no completaron una evaluación diagnóstica, se revisaron sus historias clínicas 2 años después y no se detectaron nuevas consultas por quejas cognitivas ni casos de diagnóstico de DC.

La evaluación de los pacientes con quejas cognitivas es muy variable en función de los recursos, por lo general escasos en tiempo, y de la experiencia de cada profesional. Con el resultado de nuestro estudio intentamos ayudar a establecer una estrategia de actuación sencilla que ayude a los profesionales de la AP a manejar las quejas cognitivas y en la decisión de derivar a un paciente a una consulta especializada. La comparación realizada en este trabajo entre los TCB más utilizados en nuestro entorno no pretende ser una competición para determinar un ganador. Entendemos que no se puede considerar que haya una prueba mejor que otra ni tampoco pretender encontrar una única prueba que resuelva todos los casos de quejas cognitivas. Así mismo, la información que extraemos de cada uno de los TCB puede ser distinta y complementaria. De esta forma, mientras que el MMSE es una prueba que evalúa varios dominios cognitivos puede ser útil en una primera exploración; por otro lado, el T@M es una prueba unidominio que puede ser especialmente útil en quejas puramente mnésicas. El Fototest tiene un especial interés en la primera atención de una persona con quejas porque de una forma eficiente y sin que intervengan condicionantes como el nivel cultural muestra el rendimiento en áreas de lenguaje y memoria. Por último, el cuestionario AD8 es también una herramienta que de una forma estructurada nos aporta la valiosa impresión de un informante. En cualquiera de los casos los resultados de este trabajo orientan a utilizar unos puntos de corte más exigentes para considerar el resultado como alterado. Esta interpretación se encuentra en consonancia con estudios recientes sobre datos normativos en pacientes neurológicos sin DC²⁷, con base en los cuales también se eleva el punto de corte originalmente propuesto.

Según estos resultados, para pacientes que son evaluados en un primer momento en un contexto no especializado, el Fototest, el cuestionario AD8 y en especial la combinación de ambos, son las pruebas con mejor valor predictivo negativo. Atendiendo a la probabilidad de clasificar correctamente a un enfermo y a un sano (aROC) aplicando el mismo punto de corte, las 2 pruebas que mejores resultados mostraron fueron el Fototest y el cuestionario AD8. Ambos son de fácil administración y tienen un tiempo de aplicación de unos 5 min como máximo. Dado que además el cuestionario AD8 se puede llenar por el acompañante al mismo tiempo

que se administra el Fototest al paciente, es una estrategia interesante para una consulta de AP, en la que el tiempo es escaso.

La capacidad diagnóstica del TCB más utilizado habitualmente en AP, el MMSE, es buena pero no tanto como el resto de los analizados. Es por ello por lo que quizás su uso en AP podría ser sustituido por otros TCB²⁸, como el Fototest, al tener mejor rendimiento diagnóstico^{29,30} y coste-efectividad²².

En el caso del T@M, el rendimiento diagnóstico es bueno también, con una sensibilidad muy alta, sin embargo, más de la mitad de las personas que tienen una puntuación alterada, no tienen DC. Si atendemos a las características de la prueba, ya que esta evalúa de forma casi exclusiva la memoria, planteamos que su ámbito de utilización sea una consulta especializada o en una segunda evaluación, entendiendo que en ese contexto la prevalencia más alta de DC mejorará el rendimiento diagnóstico.

Conclusión

El presente estudio evalúa la UD de Fototest, AD8, T@M y MMSE en una muestra poblacional y su posterior aplicación en una muestra de atención no especializada. El Fototest, y su combinación con AD8, muestran la mejor capacidad de discriminación diagnóstica de todos, postulándose como herramientas útiles para su uso en AP. Los puntos de corte que proponemos son más exigentes y mejoran los valores de sensibilidad de las herramientas de evaluación inicial de personas con quejas cognitivas.

Financiación

Para la realización del presente artículo no se ha recibido financiación. El proyecto DEBA fue parcialmente financiado por la Diputación Foral de Gipuzkoa (convocatoria 2016 del Programa de apoyo a la Red guipuzcoana de Ciencia, Tecnología e Innovación) y por el Gobierno Vasco (exp.2016111096, convocatoria 2016-ayudas a proyectos de investigación en salud). La entidad financiadora es Fundación CITA Alzheimer.

Conflicto de intereses

Los autores declaran la ausencia de conflicto de intereses.

Este trabajo no ha sido comunicado previamente en ninguna reunión o congreso.

Agradecimientos

Los autores de este trabajo quisieran manifestar su profundo agradecimiento a los participantes del Proyecto DEBA y del Programa GOIZ Alzheimer Beasain por el tiempo dedicado y por la disposición mostrada. Así mismo, queremos subrayar el compromiso de los municipios gipuzkoanos de Deba y Beasain con esta investigación.

Appendix A. Anexo 1 Componentes del Grupo de trabajo Proyecto DEBA y del Grupo de trabajo GOIZ Alzheimer Beasain

Componentes del Grupo de trabajo Proyecto DEBA: Álvarez, I.; Álvarez, A.I.; Beltrán Igúñio, A.; Bilbao, M.; Garmendia, M.E.; González-Martin, L.; Ibarbia, A.M.; Sanzo, J.M.; Tapia, A.; Villaverde, F.J.

Componentes del Grupo de trabajo GOIZ Alzheimer Beasain: Aquizu, I.; Arriondo, M.A.; Baztarrika, E.; Etxeberria, L.; García-Arrea, E.; García-Domínguez, M.; Imaz, E.; Iparagirre, M.; Iridoy, M.; Larrea, A.; López, M.D.; Martín, F.; Olaskoaga, A.; Pacheco, P.; Pérez-Rodríguez, A.M.; Porres, Y.; Ruibal, M.; San Juan, B.; Tilves, M.J.; Zapirain, E.

Anexo 2. Material adicional

Se puede consultar material adicional a este artículo en su versión electrónica disponible en [doi:10.1016/j.nrl.2022.05.006](https://doi.org/10.1016/j.nrl.2022.05.006).

Bibliografía

1. Soto-Gordoa M, Arrospide A, Moreno-Izco F, Martínez-Lage P, Castilla I, Mar J. Projecting burden of dementia in Spain 2010-2050: Impact of modifying risk factors. *J Alzheimers Dis.* 2015;48:721–30, <http://dx.doi.org/10.3233/JAD-150233>.
2. Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *Lancet.* 2020;396:413–46, [http://dx.doi.org/10.1016/s0140-6736\(20\)30367-6](http://dx.doi.org/10.1016/s0140-6736(20)30367-6) (Dement Geriatr Cogn Dis 37 2014):.
3. Poveda AJ, Baquero M, Guerreiro G-AM. Clinical stages of patients with Alzheimer disease treated in specialist clinics in Spain. The EACE study. *Neurologia.* 2013;28:477–87, <http://dx.doi.org/10.1016/j.nrleng.2013.10.003>.
4. Contador I, Fernández-Calvo B, Ramos F, Tapias-Merino E, Bermejo-Pareja F. [Dementia screening in primary care: critical review]. *Rev Neurologia.* 2010;51:677–86.
5. Olazarán J, Hoyos-Alonso MC, Ser T del, Barral GA, Conde-Sala JL, Bermejo-Pareja F, et al. Aplicación práctica de los test cognitivos breves. *Neurología.* 2016;31:183–94, <http://dx.doi.org/10.1016/j.nrl.2015.07.009>.
6. Villarejo A, Puertas-Martín V. Utilidad de los test breves en el cribado de demencia. *Neurología.* 2011;26:425–33, <http://dx.doi.org/10.1016/j.nrl.2010.12.002>.
7. Petersen R. Mild cognitive impairment as a diagnostic entity. *J Intern Med.* 2004;256:183–94, <http://dx.doi.org/10.1111/j.1365-2796.2004.01388.x>.
8. Kvitting AS, Johansson MM, Marcusson J. Accuracy of the cognitive assessment battery in a primary care population. *Dement Geriatr Cogn Disord.* 2019;9:294–301, <http://dx.doi.org/10.1159/000501365>.
9. Trivedi D. Cochrane Review Summary: Mini-Mental State Examination (MMSE) for the detection of dementia in clinically unevaluated people aged 65 and over in community and primary care populations. *Prim Heal Care Res Amp Dev.* 2017;18:527–8, <http://dx.doi.org/10.1017/s1463423617000202>.
10. Carnero-Pardo C, Lopez-Alcalde S, Allegri RF, Russo MJ. A systematic review and meta-analysis of the diagnostic accuracy of the Phototest for cognitive impairment and dementia. *Dement Neuropsychol.* 2014;8:141–7, <http://dx.doi.org/10.1590/S1980-57642014DN82000009>.
11. Carnero-Pardo C, Sáez-Zea C, de la Vega Cotarelo R, Gurpegui M, FOTOTRANS en nombre del grupo. Estudio FOTOTRANS: estudio multicéntrico sobre la validez del Fototest en condiciones de práctica clínica. *Neurología.* 2012;27:68–75, <http://dx.doi.org/10.1016/j.nrl.2011.06.001>.
12. Carnero-Pardo C, Montoro-Ríos MT. [The photo test]. *Rev Neurol.* 2004;39:801–6.
13. Rami L, Molinuevo J, Sanchez-Valle R, Bosch B, Villar A. Screening for amnestic 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, <http://dx.doi.org/10.1002/gps.1672>.
14. Rami L, Bosch B, Valls-Pedret C, Caprile C, Diaz S-VR, Molinuevo J. [Discriminatory validity and association of the mini-mental test (MMSE) and the memory alteration test (M@T) with a neuropsychological battery in patients with amnestic mild cognitive impairment and Alzheimer's disease]. *Rev Neurol.* 2009;49:169–74.
15. Pardo CC, de la Vega Cotarelo VR, Alcalde LS, Aparicio MC, Carrillo VR, Gavilán ME, et al. Evaluación de la utilidad diagnóstica de la versión española del cuestionario al informador «AD8». *Neurología.* 2013;28:88–94, <http://dx.doi.org/10.1016/j.nrl.2012.03.012>.
16. Carnero-Pardo C, Rego-García I, Llorente MM, Ródenas MA, Carrillo RV. Utilidad diagnóstica de test cognitivos breves en el cribado de deterioro cognitivo. *Neurología (English Edition).* 2022;37:441–9, <http://dx.doi.org/10.1016/j.nrl.2019.05.007>.
17. Kivipelto M, Ngandu T, Laatikainen T, Winblad B, Soininen H, Tuomilehto J. Risk score for the prediction of dementia risk in 20 years among middle aged people: A longitudinal, population-based study. *Lancet Neurol.* 2006;5:735–41, [http://dx.doi.org/10.1016/S1474-4422\(06\)70537-3](http://dx.doi.org/10.1016/S1474-4422(06)70537-3).
18. Jack CR, Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein S, et al. NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimer Dement.* 2018;14:535–62, <http://dx.doi.org/10.1016/j.jalz.2018.02.018>.
19. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state" A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res.* 1975;12:189–98, [http://dx.doi.org/10.1016/0022-3956\(75\)90026-6](http://dx.doi.org/10.1016/0022-3956(75)90026-6).
20. Creavin ST, Wisniewski S, Noel-Storr AH, Trevelyan CM, Hampton T, Rayment D, et al. Mini-Mental State Examination (MMSE) for the detection of dementia in clinically unevaluated people aged 65 and over in community and primary care populations. *Cochrane Database Syst Rev.* 2016;2016(1):CD011145, <http://dx.doi.org/10.1002/14651858.CD011145.pub2>.
21. 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, <http://dx.doi.org/10.1371/journal.pone.0027069>.
22. Carnero-Pardo C, Espejo-Martínez B, Lopez-Alcalde S, Espinosa-García M, Saez-Zea C, Vilchez-Carrillo R, et al. Effectiveness and costs of phototest in dementia and cognitive impairment screening. *BMC Neurology.* 2011;11:92, <http://dx.doi.org/10.1186/1471-2377-11-92>.
23. Galvin JE, Roe CM, Powlishta KK, Coats MA, Muich SJ, Grant E, et al. The AD8 A brief informant interview to detect dementia. *Neurology.* 2005;65:559–64, <http://dx.doi.org/10.1212/01.wnl.0000172958.95282.2a>.
24. Weinstein S, Obuchowski NA, Lieber ML. Clinical evaluation of diagnostic tests. *Am J Roentgenol.* 2005;184:14–9, <http://dx.doi.org/10.2214/ajr.184.1.01840014>.

25. Obuchowski NA. ROC Analysis. Am J Roentgenol. 2005;184:364–72, <http://dx.doi.org/10.2214/ajr.184.2.01840364>.
26. 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.
27. Pardo CC, Muñoz IC, Cueva LT, Alcalde SL, Carrillo RV. Valores normativos del Fototest en pacientes neurológicos sin deterioro cognitivo. Neurología. 2022;37:45–52, <http://dx.doi.org/10.1016/j.nrl.2018.03.001>.
28. Carnero-Pardo C. ¿Es hora de jubilar al Mini-Mental? Neurología. 2014;29:473–81, <http://dx.doi.org/10.1016/j.nrl.2013.07.003>.
29. Borson S, Scanlan JM, Watanabe J, Tu S, Lessig M. Simplifying detection of cognitive impairment: Comparison of the Mini-Cog and Mini-Mental State Examination in a Multiethnic Sample. J Am Geriatr Soc. 2005;53:871–4, <http://dx.doi.org/10.1111/j.1532-5415.2005.53269.x>.
30. Custodio N, Lira D, Herrera-Perez E, Montesinos R, Castro-Suarez S, Cuenca-Alfaro J, et al. Memory alteration test to detect amnestic mild cognitive impairment and early Alzheimer's dementia in population with low educational level. Front Aging Neurosci. 2017;9:278, <http://dx.doi.org/10.3389/fnagi.2017.00278>.