



ORIGINAL

Utilidad diagnóstica y validez predictiva del uso conjunto de Fototest y Mini-Cog en deterioro cognitivo[☆]

C. Carnero-Pardo^{a,*}, S. López-Alcalde^a, M. Florido-Santiago^b, M. Espinosa-García^a, I. Rego-García^c, R. Calle-Calle^c, I. Carrera-Muñoz^c y R. de la Vega-Cotarelo^d

^a FIDYAN Neurocenter, Granada, España

^b Hospital Vithas Parque San Antonio, Málaga, España

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

^d Servicio de Medicina Interna, Hospital Punta de Europa, Algeciras, Cádiz, España

Recibido el 9 de noviembre de 2020; aceptado el 28 de enero de 2021

Accesible en línea el 23 de abril de 2021

PALABRAS CLAVE

Fototest;
Mini-Cog;
Cribado;
Utilidad diagnóstica;
Validez predictiva;
Deterioro cognitivo

Resumen

Introducción: El Fototest y el Mini-Cog incluyen todos los dominios que debieran formar parte de una evaluación cognitiva. Nuestro objetivo es evaluar la utilidad diagnóstica del uso conjunto de ambos instrumentos para el diagnóstico de deterioro cognitivo (DC).

Métodos: Estudio fase III de evaluación de pruebas diagnósticas con 2 muestras independientes, estudio (448 sujetos), dividida aleatoriamente en 2 dataset (Base 80%, Test 20%), y Externa (61 sujetos). Prueba index: Fototest y Mini-Cog aplicados consecutivamente; prueba de referencia: evaluación cognitiva formal. Se evalúa la UD del uso combinado y escalonado de los modelos simple (Comb-Simple), regresión logística (Comb-RL) y árbol aleatorio (Comb-AA) para identificar DC ($GDS \geq 3$). Se realiza un análisis exploratorio en Base seleccionando los criterios que maximizan la exactitud; la evaluación se realiza en las muestras Test y externa mediante un análisis preespecificado con los criterios seleccionados.

Resultados: La UD de los modelos combinados en Base (Comb-Simple 88,3 [(88,5-91,4) [exactitud, LI95%-LS95%], Comb-RL 91,6 [88,2-94,3] y Comb-AA 95,2 [92,5-97,2]]) es significativamente superior a la de Mini-Cog y Fototest (81,6 [77,1-85,4] y 84,9 [80,8-88,5], respectivamente); estos resultados son replicados en Test (Comb-Simple 88,9 [exactitud], Comb-RL 95,6 y Comb-AA 92,2) y externa (Comb-Simple 91,8, Comb-RL 90,2 y Comb-AA 88,5). La aplicación escalonada mantiene la misma UD pero requiere menos tiempo (197,3 ± 56,7 vs. 233,9 ± 45,2, $p < 0,0001$).

[☆] Un análisis intermedio de este trabajo fue presentado en la LXXI Reunión Anual de la Sociedad Española de Neurología.

* Autor para correspondencia.

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

Conclusiones: El uso conjunto del Fototest y el Mini-Cog requiere menos de 4 min y mejora la UD de ambos instrumentos. El uso escalonado es más eficiente porque manteniendo la misma UD requiere menos tiempo de aplicación.

© 2021 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/>).

KEYWORDS

Fototest;
Mini-Cog;
Screening;
Diagnostic accuracy;
Predictive validity;
Cognitive impairment

Diagnostic accuracy and predictive validity of associated use of Fototest and Mini-Cog in cognitive impairment

Abstract

Introduction: The Fototest and Mini-Cog include all the domains that are necessary in a cognitive assessment. This study aims to evaluate the diagnostic accuracy of the combined use of both instruments for detecting cognitive impairment.

Methods: We performed a phase III diagnostic accuracy study with 2 independent samples: STUDY, which included 448 participants randomly allocated to 2 datasets (BASE [80%] and TEST [20%]); and EXTERNAL, which included 61 participants. The index test was consecutive administration of the Fototest and Mini-Cog, and the reference test was formal cognitive assessment. We evaluated the diagnostic accuracy of two-step vs. consecutive application of the tests and simple (Comb-Simple), logistic regression (Comb-LR), and random decision tree (Comb-RDT) models of their combined use for detecting cognitive impairment (Global Deterioration Scale score ≥ 3). We performed an exploratory analysis of the BASE dataset, selecting criteria that maximise accuracy; a pre-specified analysis was used to evaluate the selected criteria in the TEST and EXTERNAL datasets.

Results: The diagnostic accuracy (95% confidence interval) of the combined models in the BASE dataset (Comb-Simple: 88.3 [88.5-91.4]; Comb-LR: 91.6 [88.2-94.3]; Comb-RDT 95.2 [92.5-97.2]) was significantly higher than the individual values observed for the Mini-Cog and Fototest (81.6 [77.1-85.4] and 84.9 [80.8-88.5], respectively). These results were replicated in the TEST (Comb-Simple: 88.9; Comb-LR: 95.6; Comb-RDT: 92.2) and EXTERNAL datasets (Comb-Simple: 91.8; Comb-LR: 90.2; Comb-RDT: 88.5). Two-step application had the same diagnostic accuracy than consecutive application but required less time (mean [SD] of 197.3 s [56.7] vs. 233.9 s [45.2]; $P < .0001$).

Conclusions: Combined application of the Fototest and Mini-Cog takes less than 4 minutes and improves the diagnostic accuracy of both instruments. Two-step application is more efficient as it requires less time while maintaining the same diagnostic accuracy.

© 2021 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

La evaluación cognitiva tiene como objetivo documentar y dimensionar la existencia de una posible afectación en cualquiera de los distintos dominios cognitivos (atención, memoria, lenguaje, funciones ejecutivas, praxias y habilidades visuoespaciales), integrando estos hallazgos en el proceso diagnóstico y en el abordaje terapéutico. Es un elemento de la exploración que, en algunos casos, como las consultas por quejas cognitivas, cada día más frecuentes¹, se convierte en una parte esencial e imprescindible de la misma².

La evaluación cognitiva puede hacerse informalmente; así, durante la anamnesis se puede constatar la presencia de desorientación, problemas expresivos, flagrantes fallos de memoria³ u observar signos como el «no-sé»⁴ o «la mirada suplicante»⁵, que orienten o directamente documenten, la existencia de déficits cognitivos. No obstante, lo

habitual es realizar la evaluación cognitiva usando instrumentos que pongan en juego y evalúen de forma estructurada y cuantificable las capacidades cognitivas del sujeto. Estos instrumentos o test cognitivos, al igual que otros instrumentos de evaluación, como el oftalmoscopio o el martillo de reflejos, deben reunir determinadas condiciones técnicas y exigen al profesional destreza en su uso y capacidad para interpretar los resultados. La realidad de nuestra práctica clínica, sometida a importantes e ineludibles limitaciones de tiempo, obliga a que estos instrumentos tengan forzosamente que ser breves y fáciles de aplicar; por otro lado, estos instrumentos debieran evaluar todos los dominios cognitivos considerados actualmente en los criterios diagnósticos de los distintos constructos de deterioro cognitivo (DC) (atención, memoria y aprendizaje, lenguaje, funciones ejecutivas, praxias y habilidades visuoespaciales)⁶⁻⁸ y, cómo no, como cualquier otro instrumento diagnóstico, poseer unas cualidades psicométricas

adecuadas⁹. Estos test cognitivos breves (TCB) multidominio debieran además haber sido validados para el diagnóstico de DC y no solo demencia, pues de lo contrario no permitiría la detección de problemas cognitivos en un momento en el que una intervención podría, en algunos casos, revertir esta situación y, en otros, retrasar o detener la progresión a demencia¹⁰. Los TCB deberían, además, disponer de datos normativos aplicables al entorno en el que se usan y ser de uso libre¹¹.

Entre los TCB multidominio validados en nuestro país para DC (anexo, tabla suplementaria 1), solo el MoCA^{12,13} y el ACE-III^{14,15} cubren todos los dominios cognitivos, pero precisan 10-15 min el MoCA y al menos 15-20 min el ACE-III, un tiempo excesivo para una consulta de Neurología General¹⁶ e imposible de asumir en las consultas de Atención Primaria¹⁷. El Mini-Mental¹⁸, el Eurotest¹⁹, el Fototest²⁰, el Mini-Cog²¹ y el Test del Reloj (TDR)²², aunque más breves, no abarcan todos los dominios cognitivos; el Mini-Mental, además, cuenta con otros muchos inconvenientes que explican que cada día sea menos utilizado²³.

El uso combinado de varios TCB es una estrategia conocida para ampliar el número de áreas cognitivas evaluadas y así mejorar la sensibilidad y la capacidad diagnóstica en casos de DC. Los ejemplos más conocidos son el test de los 7 minutos, que combina una prueba de orientación, una prueba de recuerdo facilitado, una prueba de fluidez verbal semántica, y el TDR²⁴ y el propio Mini-Cog, que no deja de ser la combinación de una prueba de recuerdo diferido y el TDR²¹.

El Fototest²⁰ y el Mini-Cog²¹ son 2 instrumentos muy breves (menos de 3 y 2 min, respectivamente) de uso extendido en nuestro país, que han sido específicamente validados para DC y que evalúan lenguaje, funciones ejecutivas y memoria visual episódica con recuerdo libre y facilitado (Fototest)²⁰, y atención, memoria verbal y funciones visuoespaciales y praxias visuoespaciales (Mini-Cog)²¹. Nuestra hipótesis es que el uso conjunto de Fototest y Mini-Cog, que abarcaría todos los dominios cognitivos que debe incluir una evaluación cognitiva, constituiría una evaluación cognitiva abreviada que requeriría, teóricamente, un tiempo muy discreto (< 5 min) más asumible en estos ámbitos asistenciales y mejoraría la utilidad diagnóstica (UD) de ambos.

Nuestro objetivo es evaluar la UD y la validez predictiva del uso conjunto de Fototest y Mini-Cog en la detección de DC en condiciones de práctica clínica habitual.

Métodos

Diseño. Estudio fase III de evaluación de pruebas diagnósticas²⁵, transversal, prospectivo, naturalístico y pragmático^{26,27}, en el que se han incluido pacientes nuevos atendidos entre el 21 de febrero del 2018 y el 25 de septiembre del 2018 (a.i.) en una consulta de Neurología con especial dedicación, aunque no exclusiva, a Neurología Cognitivo-Conductual del Hospital Universitario Virgen de las Nieves de Granada, un hospital público de tercer nivel (muestra Estudio).

A efectos de validación externa, se ha incluido también una muestra independiente de pacientes atendidos por quejas cognitivas a partir de febrero del 2019 en una consulta

específica de Neurología Cognitivo-Conductual en un hospital privado de Málaga (muestra Externa).

En ambos casos, el reclutamiento fue consecutivo y sistemático y solo fueron excluidos los sujetos que no habían completado los estudios primarios, pues estudios previos nuestros han mostrado que el Mini-Cog no es un instrumento recomendable en esta población^{28,29}.

Evaluación cognitiva. A todos los sujetos, independientemente de la muestra de procedencia, la edad y el motivo de consulta, formando parte de la exploración neurológica, el neurólogo (CCP, IRG, RCC o ICM) les realizó una evaluación cognitiva abreviada que incluía la aplicación consecutiva, y por este orden, del Fototest²⁰ y la adaptación española del Mini-Cog²⁹, y que constituyen nuestra prueba *index*. El tiempo de aplicación de ambos TCB se ha medido con un cronómetro digital con una precisión de centésimas de segundos, redondeando los resultados al segundo entero superior.

Todos los sujetos atendidos por quejas cognitivas o alteraciones conductuales, o con una puntuación en el Fototest igual o inferior al percentil 10³⁰, cuentan adicionalmente con una evaluación cognitiva formal como prueba de referencia *standard*, que incluye orientación (temporal, espacial y personal), atención (dígitos directos e inversos, Trail Making test [TMT]-A³¹), memoria (aprendizaje, recuerdo libre y reconocimiento de la lista de palabras del CERAD³² en los pacientes de Granada, o el Free and Cued Selective Reminding Test³³ [FCSRT]) en los de Málaga), lenguaje (versión abreviada del test de Boston³⁴, fluidez verbal semántica³⁵ y comprensión de órdenes), praxias motoras (imitación de gestos de la escala EULA³⁶), funciones ejecutivas (semejanzas WAIS abreviada, cálculo con monedas del Eurotest¹⁹, fluidez verbal semántica³⁵ y TMT-B³¹) y visuoespaciales (copia y dibujos del CERAD³²); se han considerado patológicos resultados iguales o inferiores al percentil 5, -1,5 z o a una puntuación escalar de 6 o menos según las normas disponibles en las distintas pruebas. Esta evaluación fue llevada a cabo por un investigador (SLA, MFS o MEG) cegado con respecto a los resultados de los TCB. Estos sujetos disponen también de una evaluación funcional (escala de Barthel³⁷ y Lawton y Brody³⁸). La evaluación de referencia y los TCB se realizaron en el mismo día en ambas muestras, excepto en 9 casos de la muestra Externa, en los que se realizó con un margen máximo de ± 8 días.

Diagnósticos. Independientemente del motivo de consulta y del diagnóstico final, todos los sujetos fueron clasificados con respecto a su estado cognitivo y funcional con arreglo a una operativización de la escala *Global Deterioration Scale*³⁹ (anexo, tabla suplementaria 2); los estadios 1 y 2 fueron considerados como «no DC» (NoDC) y los estadios 3 y superiores como DC. Este diagnóstico lo realizó un neurólogo experimentado en Neurología Cognitiva y Conductual (CCP) teniendo en consideración los resultados de la evaluación cognitiva formal y no los de los TCB.

Modelos evaluados. Se ha evaluado la UD del Fototest y el Mini-Cog de forma individual e independiente, así como de las siguientes formas conjuntas de usar ambos instrumentos.

Uso combinado:

- *Simple* (Comb-Simple): puntuación única resultado de la suma de las puntuaciones de Fototest y Mini-Cog.

- *Modelo de regresión logística* (Comb-RL), que incluye los resultados en Fototest y Mini-Cog como variables predictoras y considera como covariables el sexo, la edad y el nivel de estudios.
- *Árbol aleatorio de decisión* (Comb-AA): hemos elaborado un modelo de clasificación mediante una técnica de *machine learning* aplicando un procedimiento *classification and regression tree* (CART) de aprendizaje supervisado con optimización bayesiana automática de la Exactitud (porcentaje de predicciones correctas) mediante el procedimiento OptiML⁴⁰ en la plataforma BigML⁴¹, de acceso y uso libre. Para el entrenamiento de los modelos se han incluido un total de 9 variables predictoras, entre las que se incluyen variables sociodemográficas (edad, sexo y estudios), puntuaciones del Fototest total y desagregadas (denominación, recuerdo libre, recuerdo total y fluidez de nombres) y puntuaciones del Mini-Cog total y desagregadas (Mini-Cog recuerdo y Mini-Cog reloj); la variable respuesta ha sido «DC».
- *Uso escalonado*: evaluación en 2 pasos; en un primer paso se consideran 2 puntos de corte del Fototest, uno inferior (Fototest ≤ 25), que maximiza el valor predictivo positivo o *precisión* reduciendo al mínimo o incluso eliminando los falsos positivos, y otro superior (Fototest ≥ 40) que maximiza el valor predictivo negativo minimizando/eliminando los falsos negativos; los sujetos con Fototest ≤ 25 son clasificados directamente como DC y aquellos con Fototest ≥ 40 como NoDC sin tener que valorar el Mini-Cog. El segundo paso tiene en consideración los resultados del Mini-Cog solo en aquellos sujetos con puntuaciones intermedias del Fototest ($26 \leq \text{Fototest} \leq 39$), aplicando a estos los modelos combinados previos (Esc-Simple, Esc-RL, Esc-AA).

Análisis estadístico. Se ha realizado un estudio descriptivo de variables sociodemográficas y resultados; las comparaciones entre grupos se han llevado a cabo con las pruebas t de Student y de la chi cuadrado (χ^2), según sean variables continuas o categóricas.

El análisis se ha llevado a cabo en la muestra Estudio y para evitar el fenómeno de sobreajuste y la consecuente sobreestimación de la UD se ha dividido esta de forma aleatoria en 2 *datasets*; el primero de ellos, Base, con el 80% de los sujetos se ha utilizado para la elaboración de los distintos modelos y la realización de un *análisis exploratorio* que permitiera la selección de los mejores puntos de corte y criterios; y el segundo *dataset*, Test, con el 20% de la muestra, se ha reservado de forma exclusiva para la validación independiente de los distintos modelos y criterios seleccionados en Base.

La UD de los distintos modelos en Base se ha valorado mediante el área bajo la curva ROC (aROC) con sus respectivos errores estándar (ee) para DC vs. NoDC; la comparación entre aROC se ha realizado por el método de Hanley-McNeil⁴². Se ha seleccionado como mejor punto de corte o criterio aquel que maximizara la exactitud de la prueba (porcentaje de clasificaciones correctas). En la elección del Comb-AA, dado que muchos de los 98 modelos elaborados tenían iguales valores de exactitud, se ha optado también por maximizar el coeficiente phi (coeficiente de correlación de Mathews) y por la simplicidad, seleccionando por parsimonia el modelo con menor número de nodos.

La evaluación y la validación interna de los modelos y criterios seleccionados en Base se ha realizado de forma independiente en el *dataset* Test y la validación externa en la muestra Externa realizando un *análisis preespecificado* de la UD mediante los parámetros sensibilidad, precisión, coeficiente phi y exactitud, habitualmente utilizados para la validación de modelos de predicción. Estos datos permiten el cálculo de los parámetros diagnósticos clásicos (especificidad, valores predictivos y cocientes de probabilidad), así como la elaboración de la tabla de contingencia correspondiente.

Los cálculos se han llevado a cabo con SPSS 19.0.0⁴³ y MedCalc 18.9.1⁴⁴.

Aspectos formales. El diseño del estudio y la preparación del manuscrito ha seguido las recomendaciones de las iniciativas STARD 2015⁴⁵ y STARDdem⁴⁶ de estudios de evaluación de pruebas diagnósticas en general y para DC en particular, respectivamente, así como las recomendaciones para el desarrollo de modelos predictivos en Biomedicina⁴⁷.

Este estudio ha atendido los principios de investigación médica combinada con la atención médica contemplados en la última revisión de la Declaración de Helsinki de la Asociación Médica Mundial (Edimburgo, 2000).

Resultados

La muestra Estudio tiene un total de 448 sujetos con una edad media \pm desviación estándar de $60,5 \pm 17,7$ años, con ligero predominio de mujeres y de sujetos con más de estudios primarios (el 53,8% y el 52,2%, respectivamente) y con una prevalencia de sujetos con DC del 46,4%. Los 2 *datasets* generados aleatoriamente, Base y Test, con 358 (80%) y 90 (20%) sujetos, respectivamente, no mostraron diferencias significativas en ninguna de las variables sociodemográficas consideradas ni en la prevalencia de DC ni en los resultados del Fototest y Mini-Cog.

La muestra Externa, en cambio, tiene caracteres distintos, con mucha mayor prevalencia de DC (77,0%), mayor edad ($71,7 \pm 10,0$ años), mayor nivel educativo (60,7% > Primaria) y menor predominio de mujeres (47,5%), aunque en los 2 últimos casos la diferencia no llega a ser significativa; los resultados de Fototest y Mini-Cog también son significativamente menores que en la muestra estudio.

La *tabla 1* resume las características sociodemográficas, la prevalencia de DC y los resultados en las 2 muestras y los 2 *datasets*.

El modelo de AA seleccionado entre los 98 generados en la plataforma tiene una exactitud del 92,2% y un coeficiente phi de 0,84 en la muestra Test, con un total de 7 niveles (*fig. 1*) (*anexo, tabla suplementaria 3*).

La *tabla 2* resume los parámetros diagnósticos de los distintos modelos en las distintas muestras y *datasets*. La UD del Fototest es ligera pero significativamente superior a la del Mini-Cog ($0,93 \pm 0,01$ [aROC \pm ee] vs. $0,89 \pm 0,02$; $t=2,06$, $p=0,04$) y la UD de los modelos Comb-Simple ($0,95 \pm 0,01$) y Comb-RL ($0,98 \pm 0,01$) son significativamente superiores a la de ambos ($t > 4,00$ y $p < 0,0001$ en todas las comparaciones) (*fig. 2*) (*tabla 3*). El modelo Comb-AA carece de aROC pues tan solo incluye una alternativa clasificatoria, pero su exactitud (95,2%) es significativamente superior a

Tabla 1 Características sociodemográficas y resultados por muestras y datasets

	Base (80%)	Test (20%)	Estudio	Externa	a	b
N.º sujetos	358	90	448	61		
DC	166 (46,4)	42 (46,7)	208 (46,4)	47 (77,0)	0,03, 0,96	20,14, 0,0001
Edad (años)	60,7 ± 17,4	60,1 ± 18,7	60,7 ± 17,7	71,7 ± 10,0	0,08, 0,78	23,28, 0,0001
Sexo (mujer)	193 (53,9)	48 (53,3)	241 (53,8)	29 (47,5)	0,01, 0,92	0,84, 0,36
Estudios						
Primarios	171 (47,8)	46 (51,1)	217 (48,4)	24 (39,3)	0,31,	1,78,
> Primarios	187 (52,2)	44 (48,9)	231 (51,6)	37 (60,7)	0,57	0,18
GDS						
1	130 (36,3)	32 (35,6)	162 (36,2)	-	4,30,	37,78,
2	62 (17,3)	16 (17,8)	78 (17,4)	14 (23,0)	0,51	0,0001
3	75 (20,9)	19 (21,1)	94 (21,0)	28 (45,9)		
4	71 (19,8)	16 (17,8)	87 (19,4)	15 (24,6)		
5-6	20 (5,6)	7 (7,8)	27 (6,0)	4 (6,6)		
ECF	228 (63,7)	58 (64,4)	286 (63,8)	61 (100,0)	0,01, 0,98	32,34, 0,0001
Fototest	32,6 ± 7,8	33,0 ± 7,8	32,7 ± 7,8	29,5 ± 6,9	0,22, 0,64	9,28, 0,002
Tiempo (s)	130,2 ± 20,2 (317)	128,7 ± 19,3 (83)	129,9 ± 20,0 (400)	136,7 ± 14,8 (49)	0,38, 0,054	5,36, 0,02
Mini-Cog	2,3 ± 1,8	2,4 ± 1,8	2,4 ± 1,8	1,3 ± 1,4	0,10, 0,75	17,83, 0,0001
Tiempo (s)	103,3 ± 33,7 (310)	100,2 ± 34,9 (81)	102,7 ± 33,9 (391)	109,7 ± 34,8 (48)	0,42, 0,52	1,84, 0,17

Las cifras son n.º sujetos (%), media ± DE. En los resultados referidos a tiempo aparece entre paréntesis el n.º sujetos en los que se ha medido.

^a Comparación Base vs. Test (estadístico χ^2 o t de Student y valor de p).

^b Comparación ESTUDIO vs. EXTERNA (estadístico χ^2 o t de Student y valor de p).

la del uso individualizado del Mini-Cog (81,6%; $\chi^2 = 32,24$, $p < 0,0001$) y Fototest (84,9%; $\chi^2 = 21,16$, $p < 0,001$) y también a la del modelo Comb-Simple (88,3%; $\chi^2 = 11,2$, $p = 0,001$) y Comb-RL (91,6%; $\chi^2 = 3,76$, $p = 0,05$). Los

modelos escalonados replican de forma exacta la UD de los modelos combinados que utilizan, tan solo que el tiempo que tardan en aplicarse es significativamente inferior ($197,3 \pm 56,7$ vs. $233,9 \pm 45,2$; $t = 8,8$, $p < 0,0001$).

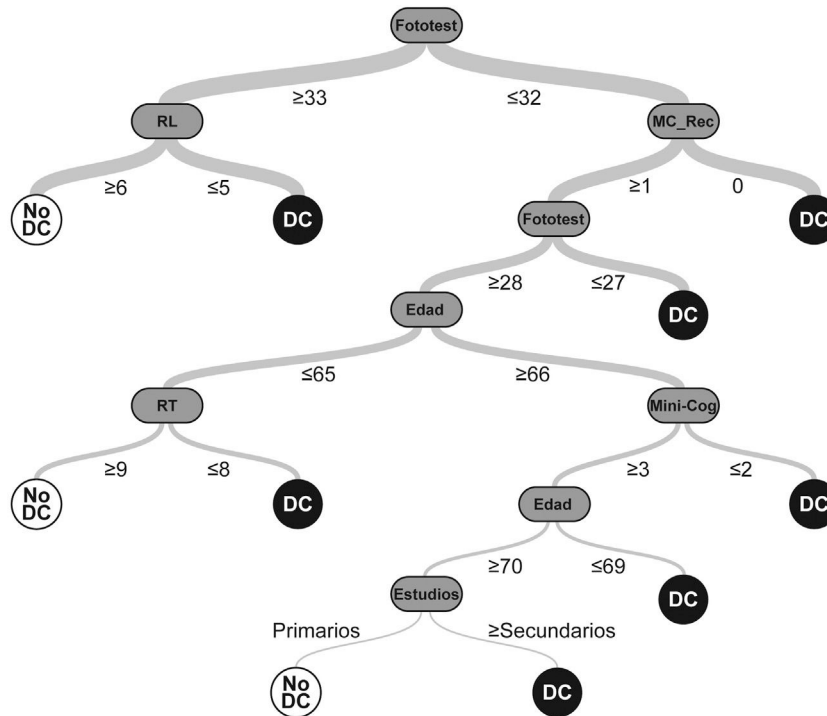


Figura 1 Modelo del árbol aleatorio de decisión

DC: deterioro cognitivo; MC.Rec: puntuación del recuerdo en el Mini-Cog; NoDC: no deterioro cognitivo; RL: puntuación de recuerdo libre en el Fototest; RT: puntuación de recuerdo total en el Fototest.

Tabla 2 Utilidad diagnóstica en el *dataset* Base y evaluación en *dataset* Test y muestra externa

Muestra test	Base (N = 358; 166 DC)				Test (N = 90; 42 DC)				Externa (N = 61; 47 DC)			
	aROC	Criterio	Exactitud	Tiempo	S	Precisión	Phi	Exactitud	S	Precisión	Phi	Exactitud
Fototest	0,93 ± 0,01	≤ 32	84,9 (80,8-88,5)	130,2 ± 20,2	0,93	0,78	0,70	84,4	0,91	0,98	0,79	91,8
Mini-Cog	0,89 ± 0,02	≤ 2	81,6 (77,1-85,4)	103,3 ± 33,7	0,86	0,86	0,73	86,7	0,85	0,87	0,41	78,7
Comb-Simple	0,95 ± 0,01	≤ 34	88,3 (88,5-91,4)	233,9 ± 45,2	0,93	0,85	0,78	88,9	0,91	0,98	0,79	91,8
Comb-RL ^a	0,98 ± 0,01	≥ 0,50	91,6 (88,2-94,3)		0,95	0,95	0,91	95,6	0,91	0,96	0,74	90,2
Comb-AA	—	^b	95,2 (92,5-97,2)		0,90	0,93	0,87	92,2	0,91	0,93	0,68	88,5
Esc-Simple	—	^c	88,3 (88,5-91,4)	197,3 ± 56,7	0,93	0,85	0,78	88,9	0,91	0,98	0,79	91,8
Esc-RL	—	^d	91,6 (88,2-94,3)		0,95	0,94	0,89	94,4	0,94	0,94	0,72	90,2
Esc-AA	—	^e	95,2 (92,5-97,2)		0,90	0,93	0,87	92,2	0,91	0,91	0,63	86,9

Los resultados en Base son media ± DE o porcentaje (intervalo de confianza exacto).

AA: árbol aleatorio; aROC: Área bajo la curva ROC; Comb: modelo combinado; Esc: modelo escalonado DC: deterioro cognitivo. N: n.º de sujetos; Phi: coeficiente de correlación de Mathews; S: sensibilidad; RL: regresión logística.

^a Prob DC = 12,02-0,47 Fototest-1,11 Mini-Cog-2,06 Estudios (> Primarios) + 0,07 Edad (años) - 0,28 sexo (Mujer); si Prob ≥ 0,5 DC, si Prob < 0,5 NoDC.

^b Véase árbol aleatorio de decisión (fig. 2).

^c Paso 1: Fototest ≤ 25 DC & Fototest ≥ 40 NoDC; paso 2: 26 ≥ Fototest ≤ 39: FotoCog ≤ 34 DC, ≥ 35 NoDC.

^d Paso 1: Fototest ≤ 25 DC & Fototest ≥ 40 NoDC; paso 2: 26 ≥ Fototest ≤ 39: aplicar ^a.

^e Paso 1: Fototest ≤ 25 DC & Fototest ≥ 40 NoDC; paso 2: 26 ≥ Fototest ≤ 39: aplicar ^b.

Tabla 3 Comparación utilidad diagnóstica del uso individual y combinado de Fototest y Mini-Cog

	Fototest	Mini-Cog	Comb Simple	Comb RL
Fototest	—	0,036 ± 0,02 (2,06, 0,04)	0,024 ± 0,004 (5,59, < 0,0001)	0,017 ± 0,018 (5,04, < 0,0001)
Mini-Cog		—	0,060 ± 0,015 (4,04, < 0,0001)	0,080 ± 0,014 (5,73, < 0,0001)
Comb-Simple			—	0,020 ± 0,006 (3,56, 0,0004)
Comb-RL				—

Los resultados son: diferencias entre aROC ± EE (valor estadístico z, valor p).

aROC: área bajo la curva ROC; EE: error estándar.

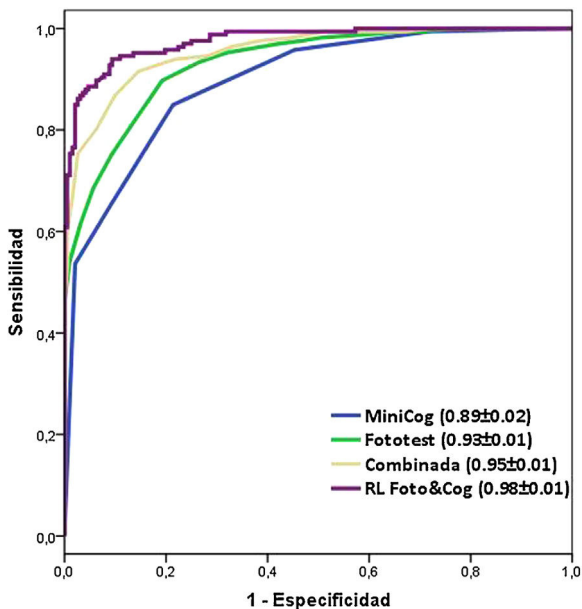


Figura 2 Utilidad diagnóstica del uso individual y combinado de Fototest y Mini-Cog.

Combinada: modelo combinado simple; RL Foto&Cog: modelo combinado de regresión logística.

La validación de los modelos en el *dataset* Test reproduce en gran medida las predicciones de los modelos generados en Base, excepto en los modelos que utilizan el AA en los que la UD el ligeramente inferior, algo esperable por el conocido efecto de sobreajuste. Los modelos también muestran una UD muy satisfactoria en la muestra Externa, con exactitudes entre el 86,9 y 91,8%, con la excepción del Mini-Cog, cuya exactitud es 78,7%, aunque la diferencia con respecto a la exactitud en el *dataset* Base no es significativa ($\chi^2 = 0,29$, $p = 0,59$).

En las [tablas suplementarias 4 a 11 del anexo](#) aparecen las tablas de contingencia y los parámetros diagnósticos clásicos (sensibilidad, especificidad, valores predictivos y cocientes de probabilidad) de cada uno de los modelos y en cada una de las muestras para los puntos de corte y criterios seleccionados.

Discusión

Los resultados de este estudio prospectivo realizado en condiciones de práctica clínica habitual muestran claramente que el uso conjunto de Fototest y Mini-Cog mejora la UD

del uso individual de estos instrumentos. Estos resultados son consistentes ya que han sido replicados en 2 muestras independientes, una, con las mismas características socio-demográficas y clínicas de la muestra que ha servido para elaborar los modelos, y otra, completamente diferente procedente de un entorno distinto desde el punto de vista sociodemográfico y clínico (consulta específica de Neurología Cognitivo-Conductual, mayor edad y prevalencia de DC y distinta ubicación geográfica). La mejor UD de los modelos conjuntos se debe probablemente a que se realiza una evaluación cognitiva más extensa y detallada al incluir más dominios cognitivos en la misma; esta circunstancia también facilita el que se identifiquen perfiles de afectación que mejorarían la orientación diagnóstica, aunque en este estudio no hemos evaluado esta cuestión que sería conveniente abordarla en estudios posteriores.

Nuestros resultados también son consistentes y refrendan los datos previos disponibles en nuestro país con respecto a la UD de los instrumentos utilizados; así, en el único estudio previo del Mini-Cog en nuestro país, el aROC fue $0,88 \pm 0,01$ ²⁹, prácticamente idéntica a $0,89 \pm 0,02$ encontrada en este estudio (existe otro estudio en el que la puntuación del Mini-Cog se reconstruyó a partir del test del reloj y del Mini-Mental, por lo que no lo consideramos comparable²⁸). Del mismo modo, la UD del Fototest en este estudio ($0,93 \pm 0,01$) es muy similar a la encontrada en otros estudios previos con este instrumento en nuestro país ($0,86 \pm 0,02$ ⁴⁸, $0,93 \pm 0,02$ ⁴⁹, $0,95 \pm 0,01$ ²⁰) o incluso en Argentina ($0,93 \pm 0,03$ ⁵⁰).

Los modelos complejos (Comb-RL y Comb-AA) proporcionan una mayor UD que el modelo Comb-Simple, lo que probablemente se deba a la ponderación de las contribuciones diagnósticas de cada test y al hecho de incluir también, estos modelos, datos referidos a variables socio-demográficas (sexo, edad, nivel educativo) que pueden resultar muy relevantes, en especial, en el caso del Mini-Cog²⁹. Estos modelos complejos tienen el inconveniente del proceso de computación que en la práctica clínica puede resultar difícil de llevar a cabo durante la consulta, por lo que resulta más simple y operativo el uso del modelo simple cuyo resultado es la suma de las puntuaciones en ambos instrumentos. No obstante, se ha desarrollado un programa (RVC) que facilita la aplicación en línea de estos instrumentos y que incorpora una calculadora que realiza estas computaciones de forma automática, ofreciendo un informe que incluye todos los resultados ([anexo, figura suplementaria 1](#)); esta aplicación bautizada como Predi-Cog⁵¹ (<https://www.hipocampo.org/Predi-Cog.asp>), es accesible libremente en la página web La Circunvalación del Hipocampo⁵² (www.hipocampo.org).

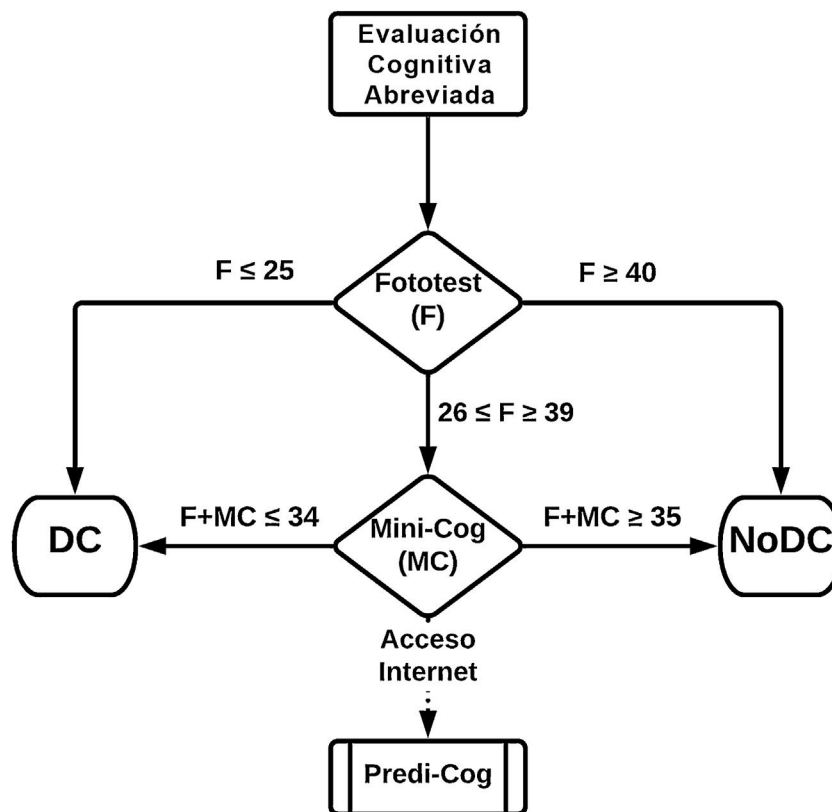


Figura 3 Estrategia de uso conjunto recomendada
DC: deterioro cognitivo; NoDC: no deterioro cognitivo.

Nuestros resultados recomiendan hacer una aplicación escalonada (fig. 3) que permite la clasificación directa, sin menoscabo alguno de la UD, de un buen porcentaje de sujetos tan solo con la aplicación del Fototest (Base 36,6%, Test 37,8% y Externa 42,6%) con el consecuente ahorro de tiempo ($197 \pm 56,7$ vs. $233,9 \pm 45,2$ s).

Nuestro estudio cuenta con algunas debilidades, como es el haber excluido a los sujetos sin al menos estudios primarios lo que hace que nuestros hallazgos no sean extensivos a la población con menos educación; pero nuestros datos previos ya recomendaban evitar el uso de Mini-Cog en esta población^{28,29}. Otra debilidad es el hecho de que aunque el 100% de la muestra Externa cuenta con evaluación cognitiva formal, en la muestra Estudio tan solo disponen de ella el 63,9% de los sujetos, aunque tan solo se ha excluido de ellas a los sujetos sin quejas cognitivas y con puntuaciones en el Fototest superiores al percentil 10, lo que hace poco probable que estos sujetos pudieran tener DC y, por tanto, minimiza el posible sesgo de verificación parcial⁵³. Finalmente, y aunque las pruebas *index* y referencia han sido aplicadas por distintos profesionales ciegos con respecto a los otros resultados y el diagnóstico se ha realizado en base a los resultados de las pruebas de referencia, el profesional que llevó a cabo el diagnóstico no estaba cegado con respecto a los resultados de los TCB. Frente a estas debilidades, nuestro estudio tiene importantes fortalezas como el carácter naturalístico y pragmático del mismo, el gran tamaño muestral y, sobre todo, la doble validación de los resultados en 2 muestras independientes y, en un caso, con caracteres muy diferentes de la muestra en la que se ha llevado a cabo el estudio exploratorio.

En conclusión, el uso conjunto del Fototest y el Mini-Cog, 2 instrumentos muy breves y fáciles de aplicar que cubren todo el abanico de funciones cognitivas imprescindibles para llevar a cabo una evaluación cognitiva abreviada, mejora la UD, ya de por sí buena, de ambos instrumentos. El uso escalonado es más eficiente que el uso combinado, porque manteniendo la eficacia, disminuye significativamente el tiempo medio de aplicación a menos de 200 s.

Datos compartidos

Los datos en los que se basa este estudio están disponibles para cualquier investigador interesado previa petición razonada y razonable.

Conflicto de intereses

C. Carnero Pardo es el creador del Fototest que está bajo licencia y puede ser usado y difundido sin fines comerciales, siempre y cuando no sea modificado y se haga reconocimiento expreso de su autoría.

Agradecimientos

Marc Torres Ciuró ha realizado las figuras.

Anexo. Material adicional

Se puede consultar material adicional a este artículo en su versión electrónica disponible en [doi:10.1016/j.nrl.2021.01.017](https://doi.org/10.1016/j.nrl.2021.01.017).

Bibliografía

- Lopez-Pousa S, Monserrat-Vila S, Turro-Garriga O, Aguilar-Barbera M, Caja-Lopez C, Vilalta-Franch J, et al. Análisis de la demanda asistencial neurológica generada por la atención primaria en un área geográfica de las comarcas de Girona. *Rev Neurol*. 2009;49:288–94.
- Henderson WE. Cognitive assessment in neurology. En: Aminoff MJ, Boller F, Swaab DF, editores. *Handbook of clinical neurology (3rd series): History of Neurology*. Elsevier B. V.; 2010. p. 235–56. Cap. 17.
- Cook C, Fay S, Rockwood K. Verbal repetition in people with mild-to-moderate Alzheimer Disease: A descriptive analysis from the VISTA clinical trial. *Alzheimer Dis Assoc Disord*. 2009;23:146–51, <http://dx.doi.org/10.1097/WAD.0b013e318193cbef>.
- Rego García I, Medina Gámez J, Valderrama Martón C, Guillén Martínez V, Vilchez Carrillo R CCP. "Don't know" sign: Description and evaluation of its diagnostic accuracy for cognitive impairment, compared to other observation based signs. 6th Congress of the European Academy of Neurology. 2020:EPO201.
- Larner AJ. Head turning sign: Pragmatic utility in clinical diagnosis of cognitive impairment. *J Neurol Neurosurg Psychiatry*. 2012;83:852–3, <http://dx.doi.org/10.1136/jnnp-2011-301804>.
- Robles A, del Ser T, Alom J, Peña-Casanova J, Grupo Asesor del Grupo de Neurología de la Conducta y Demencias de la Sociedad Española de Neurología. Propuesta de criterios para el diagnóstico clínico del deterioro cognitivo ligero, la demencia y la enfermedad de Alzheimer. *Neurología*. 2002;17:17–32.
- Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011;7:270–9, <http://dx.doi.org/10.1016/j.jalz.2011.03.008>.
- American Psychiatric Association (APA). *Manual diagnóstico y estadístico de los trastornos mentales (DSM-5)*. 5.ª ed. Madrid: Editorial Médica Panamericana; 2013.
- Peña Casanova J, Sánchez Benavides G. Test cognitivos breves: una confusión, una necesidad y una historia interminable. En: Carnero Pardo C, editor. *Test cognitivos breves*. Madrid: Ediciones SEN; 2015. p. 9–18. Cap. 1.
- Olazarán J, Hoyos-Alonso MC, del Ser T, Garrido Barral A, 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>.
- Carnero-Pardo C, Rego-García I, Mene Llorente M, Alonso Rodeñas M, Vilchez Carrillo R. Utilidad diagnóstica de test cognitivos breves en el cribado de deterioro cognitivo. *Neurología*. 2019, <http://dx.doi.org/10.1016/j.nrl.2019.05.007> (online ahead of print).
- Nasreddine ZS, Phillips NA, Bedirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment MoCA: A brief screening tool for mild cognitive impairment. *J Am Geriatr Soc*. 2005;53:695–9, <http://dx.doi.org/10.1111/j.1532-5415.2005.53221.x>.
- Lozano-Gallego M, Hernández Ferrándiz M, Turró Garriga O, Pericot Nierra I, López-Pousa S, Vilalta Franch J. Validación del Montreal Cognitive Assessment (MoCA): test de cribado para el deterioro cognitivo leve. *Datos preliminares*. *Alzheimer Real Invest Demenc*. 2009;43:4–11.
- Mathuranath PS, Nestor PJ, Berríos GE, Rakowicz W, Hodges JR. A brief cognitive test battery to differentiate Alzheimer's disease and frontotemporal dementia. *Neurology*. 2000;55:1613–20.
- Matias-Guiu JA, Cortes-Martinez A, Valles-Salgado M, Rognoni T, Fernandez-Matarrubia M, Moreno-Ramos T, et al. Addenbrooke's cognitive examination III: Diagnostic utility for mild cognitive impairment and dementia and correlation with standardized neuropsychological tests. *Int Psychogeriatr*. 2017;29:105–13, <http://dx.doi.org/10.1017/S1041610216001496>.
- Morera-Guitart J, Escudero J, Aguilar M, Aguilera JM, Carnero C, Martín R, et al. Conferencia de consenso sobre tiempos de visita en neurología: recomendaciones sobre tiempos de visita para la asistencia neurológica ambulatoria en España. *Neurología*. 2001;16:399–407.
- Irving G, Neves AL, Dambha-Miller H, Oishi A, Tagashira H, Verho A, et al. International variations in primary care physician consultation time: A systematic review of 67 countries. *BMJ Open*. 2017;7:e017902, <http://dx.doi.org/10.1136/bmjopen-2017-017902>.
- Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": A practical method for grading the clinician. *J Psychiatr Res*. 1975;12:129–132, [0022-3956\(75\)90026-6 \[pii\]](https://doi.org/10.1016/0022-3956(75)90026-6).
- Carnero-Pardo C, Montoro-Ríos MT. Evaluación preliminar de un nuevo test de cribado de demencia (Eurotest). *Rev Neurol*. 2004;38:201–9.
- 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.
- Borson S, Scanlan J, Brush M, Vitaliano P, Dokmak A. The mini-cog: A cognitive "vital signs" measure for dementia screening in multi-lingual elderly. *Int J Geriatr Psychiatry*. 2000;15:1021–7.
- Carnero-Pardo C, Rego-García I, Barrios-Lopez JM, Blanco-Madera S, Calle-Calle R, Lopez-Alcalde S, et al. Assessment of the diagnostic accuracy and discriminative validity of the Clock Drawing and Mini-Cog tests in detecting cognitive impairment. *Neurología*. 2019, <http://dx.doi.org/10.1016/j.nrl.2018.12.002> (online ahead of print).
- 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>.
- Solomon PR, Hirschhoff A, Kelly B, Relin M, Brush M, DeVeaux RD, et al. A 7 minute neurocognitive screening battery highly sensitive to Alzheimer's disease. *Arch Neurol*. 1998;55:349–55.
- Sackett DL, Haynes RB. The architecture of diagnostic research. *BMJ*. 2002;324:539–41.
- Patsopoulos NA. A pragmatic view on pragmatic trials. *Dialogues Clin Neurosci*. 2011;13:217–24.
- Larner A. *Dementia in clinical practice: A neurological perspective: Pragmatic studies in the cognitive function clinic*. 3rd ed. Springer; 2018.
- Carnero-Pardo C, Cruz-Orduna I, Espejo-Martinez B, Martos-Aparicio C, Lopez-Alcalde S, Olazarán J. Utility of the Mini-Cog for detection of cognitive impairment in primary care: Data from two Spanish studies. *Int J Alzheimer Dis*. 2013;2013:285462, <http://dx.doi.org/10.1155/2013/285462>.
- Carnero-Pardo C, Rego-García I, Barrios-López JM, 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, <http://dx.doi.org/10.1016/j.nrl.2018.12.002> (online ahead of print).
- Carnero Pardo C, Carrera Muñoz I, Triguero Cueva L, Lopez Alcalde S, Vilchez Carrillo R. Valores normativos del Fototest

- en pacientes neurologicos sin deterioro cognitivo. *Neurologia*. 2018, <http://dx.doi.org/10.1016/j.nrl.2018.03.001> (online ahead of print).
31. Reitan R. Validity of the Trail Making test as an indicador of organic brain damage. *Percept Mot Skills*. 1958;8:271–6.
 32. 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.
 33. Buschke H. Cued recall in amnesia. *J Clin Neuropsychol*. 1984;6:433–40, <http://dx.doi.org/10.1080/01688638408401233>.
 34. 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:P187–91.
 35. Herrera-García JD, Rego-García I, Guillen-Martínez V, Carrasco-García M, Valderrama-Martin C, Vilchez-Carrillo R, et al. Discriminative validity of an abbreviated Semantic Verbal Fluency Test. *Dement Neuropsychol*. 2019;13:203–9, <http://dx.doi.org/10.1590/1980-57642018dn13-020009>.
 36. Pérez-Mármol JM, López-Alcalde S, Carnero-Pardo C, Cañadas-De la Fuente GA, Peralta-Ramírez MI, García-Ríos MC. Creación y diseño de un test para la evaluación de la apraxia de los miembros superiores (EULA) basado en un modelo cognitivo: un estudio piloto. *Rev Neurol*. 2015;60:66–74.
 37. Mahoney FI, Barthel DW. Functional evaluation: The Barthel Index. *Md State Med J*. 1965;14:61–5.
 38. Lawton MP, Brody EM. Assessment of older people: Self-maintaining and instrumental activities of daily living. *Gerontologist*. 1969;9:179–86.
 39. Reisberg B, Ferris SH, de Leon MJ, Crook T. The Global Deterioration Scale for assessment of primary degenerative dementia. *Am J Psychiatry*. 1982;139:1136–9.
 40. The BigML Team. OptiML with the BigML dashboard. 2020 [consultado 22 Jun 2020]. Disponible en: https://static.bigml.com/pdf/BigML_OptiML.pdf?ver=c306567.
 41. BigML Inc. Machine Learning Platform [online]. [consultado 22 Jun 2020]. Disponible en: <https://bigml.com/>. Accessed 22/6/2020.
 42. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology*. 1983;148:839–43.
 43. IBM Corp. Released 2010. IBM SPSS Statistics for Windows, Version 19.9. Armonk, NY: IBM Corp.
 44. MedCalc Statistical Software versión 18.9.1 (MedCalc Software Ltd, Ostend, Belgium; <http://www.medcal.org>; 2018.
 45. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015 an updated list of essential items for reporting diagnostic accuracy studies. *BMJ*. 2015;351:h5527, <http://dx.doi.org/10.1136/bmj.h5527>.
 46. Noel-Storr AH, McCleery JM, Richard E, Ritchie CW, Flicker L, Cullum SJ, et al. Reporting standards for studies of diagnostic test accuracy in dementia: The STARDDem Initiative. *Neurology*. 2014;83:364–73, <http://dx.doi.org/10.1212/WNL.0000000000000621>.
 47. Luo W, Phung D, Tran T, Gupta S, Rana S, Karmakar C, et al. Guidelines for developing and reporting machine learning predictive models in biomedical research: A multidisciplinary view. *J Med Internet Res*. 2016;18:e323, <http://dx.doi.org/10.2196/jmir.5870>.
 48. Carnero-Pardo C, Sáez-Zea C, de la Vega Cotarelo R, Gurpegui M, en nombre del grupo FOTOTRANS. 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>.
 49. Carnero Pardo C, de la Vega Cotarelo R, López Alcalde S, Espinosa García M, Mora Gavilán E, Vilchez Carrillo R, et al. Evaluación de la utilidad diagnóstica y validez del cuestionario al informador AD8. *Neurología*. 2011;26:14.
 50. Russo MJ, Iturry M, Sraka MA, Bartoloni L, Carnero Pardo C, Allegri RF. Diagnostic accuracy of the Phototest for cognitive impairment and dementia in Argentina. *Clin Neuropsychol*. 2014;28:1–15, <http://dx.doi.org/10.1080/13854046.2014.928748>.
 51. De la Vega Cotarelo R. Predi-Cog [online] [consultado 27 Oct 2020]. Disponible en: <https://www.hipocampo.org/Predicog.asp>.
 52. De la Vega Cotarelo R. La Circunvalación del Hipocampo [online]. [consultado 27 Oct 2020]. Disponible en: <https://www.hipocampo.org/>.
 53. O'Sullivan JW, Banerjee A, Heneghan C, Pluddemann A. Verification bias. *BMJ Evid Based Med*. 2018;23:54–5, <http://dx.doi.org/10.1136/bmjebm-2018-110919>.