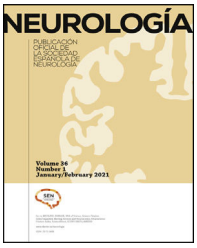
Neurología 38 (2023) 653—662

NEUROLOGÍA

www.elsevier.es/neurologia



ORIGINAL ARTICLE

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

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], R. de la Vega-Cotarelo[d]

[a] *FIDYAN Neurocenter, Granada, Spain*
[b] *Hospital Vithas Parque San Antonio, Málaga, Spain*
[c] *Servicio de Neurología, Hospital Universitario Virgen de las Nieves, Granada, Spain*
[d] *Servicio de Medicina Interna, Hospital Punta de Europa, Algeciras, Cádiz, Spain*

Received 9 November 2020; accepted 28 January 2021
Available online 17 October 2023

KEYWORDS
Fototest;
Mini-Cog;
Screening;
Sensitivity;
Diagnostic accuracy;
Predictive validity;
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 $\geq$ 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.

DOI of refers to article: https://doi.org/10.1016/j.nrl.2021.01.017.
☆ An interim analysis of this study was presented at the 71st Annual Meeting of the Spanish Society of Neurology.
* Corresponding author.
E-mail address: ccarnero@neurocenter.es (C. Carnero-Pardo).

https://doi.org/10.1016/j.nrleng.2023.10.002

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.

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

Utilidad diagnóstica y validez predictiva del uso conjunto de Fototest y Mini-Cog en deteriorocognitivo

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 dos muestras independientes, ESTUDIO (448 sujetos), dividida aleatoriamente en dos 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 $\pm$ 56.7 vs 233.9 $\pm$ 45.2, $p < 0.0001$).
Conclusiones: El uso conjunto del Fototest y el Mini-Cog requiere menos de cuatro minutos y mejora la UD de ambos instrumentos. El uso escalonado es más eficiente porque manteniendo la misma UD requiere menos tiempo de aplicación.

Introduction

The purpose of cognitive assessment is to detect and measure impairment in any cognitive domain (attention, memory, language, executive function, praxis, visuospatial skills), using these findings in the diagnostic and treatment process. In some cases (for instance in consultations due to cognitive complaints, which are increasingly frequent[1]) this assessment is a fundamental part of neurological examination.[2]

Cognitive assessment may be informal: during medical history taking, the physician may detect disorientation, impaired language expression, marked memory problems,[3] the ''don't know'' sign,[4] or the head turning sign,[5] all of which suggest cognitive impairment (CI). However, assessment is usually performed with validated instruments for structured evaluation of the patient's cognitive function. Like such other instruments as the ophthalmoscope or the reflex hammer, these cognitive instruments and tests must meet certain technical requirements, and the physician must be proficient in their use and the interpretation of results. In clinical practice, given the unavoidable limitations on consultation times, these instruments must necessarily be quick and simple to administer. Furthermore, these instruments should evaluate all the cognitive domains currently considered in the diagnostic criteria for different constructs of CI (attention, memory and learning, language, executive function, praxis, and visuospatial skills)[6—8] and, needless to say, present good psychometric properties, like any other diagnostic tool.[9] These multi-domain brief cognitive tests (BCT) should be validated specifically for the diagnosis of CI (and

not only for dementia); otherwise, the tool would be unable to detect cognitive problems at a stage when the appropriate intervention may either revert the problem or delay or halt dementia progression.[10] Normative data applicable to the population of interest should also be available for these BCTs, and these tools should ideally be available free of charge.[11]

Among the multi-domain BCTs validated in Spain for CI (Supplementary material, Table 1), only the MoCA[12,13] and the ACE-III[14,15] evaluate all cognitive domains. However, these tests require 10—15 and 15—20 minutes to administer, respectively; these administration times are excessive for general neurology consultations,[16] and impracticable in primary care.[17] Although the Mini-Mental State Examination,[18] Eurotest,[19] Fototest,[20] Mini-Cog,[21] and Clock Drawing Test[22] present shorter administration times, they do not assess all cognitive domains. Furthermore, the Mini-Mental State Examination presents many other limitations that have led to a decline in its use in recent years.[23]

Physicians frequently administer several BCTs to evaluate more cognitive domains, increasing the sensitivity and diagnostic accuracy of these tests in cases of CI. The most widely known examples are the 7 Minute Screen, which combines an orientation task, a cued recall task, a semantic verbal fluency task, and the Clock Drawing Test[24]; and the Mini-Cog, which includes a delayed recall task and the Clock Drawing Test.[21]

The Fototest[20] and the Mini-Cog[21] are 2 very short tests (administration times of 3 and 2 minutes, respectively) that are widely used in Spain and have been validated specifically for CI. The former evaluates language, executive function, and episodic visual memory with free and cued recall, whereas the latter assesses attention, verbal memory, and visuospatial skills and visuoconstructive praxis. This study explores the hypothesis that the combination of the Fototest and the Mini-Cog, which together evaluate all the cognitive domains that should be included in a cognitive assessment, improves the diagnostic accuracy of either test alone while requiring very little time to administer (< 5 minutes), a particularly interesting feature in our setting.

Our objective was to evaluate the diagnostic accuracy and predictive validity of the combination of Fototest plus Mini-Cog in detecting CI in everyday clinical practice.

Methods

Design

We conducted a phase III study for the evaluation of diagnostic tests,[25] with a cross-sectional, prospective, naturalistic, pragmatic design.[26,27] The study included new patients attended between 21 February 2018 and 25 September 2018 (inclusive) at a neurology consultation specialising in cognitive-behavioural neurology, at Hospital Universitario Virgen de las Nieves, a tertiary-level hospital in Granada, Spain (study sample).

For the purposes of external validation, we also included an independent sample of patients attended due to cognitive complaints at a cognitive-behavioural neurology clinic at a private hospital in Malaga, Spain, from February 2019 (external sample).

In both instances, recruitment was consecutive and systematic, only excluding individuals who had not completed primary education, since previous studies have shown that the Mini-Cog is not recommended for this population.[28,29]

Cognitive assessment

All individuals, regardless of their age, the reason for consultation, or the sample in which they were included, underwent a brief cognitive assessment (index assessment) consisting of the administration of the Fototest[20] and the Spanish-language version of the Mini-Cog.[29] The tests were administered consecutively, and in that order, by a neurologist (CCP, IRG, RCC, or ICM). The administration time for both BCTs was measured with a digital stopwatch, with a precision of one-hundredth of a second. Times were rounded up to the next whole number.

All individuals attended due to cognitive complaints or behavioural alterations, or with Fototest scores equal to or lower than the 10th percentile,[30] also underwent a formal cognitive assessment (reference assessment), including tests of orientation (time, space, and person), attention (forward and backward digit span, Trail Making Test A [TMT-A][31]), memory (learning, free recall, and recognition of the CERAD word list,[32] for patients from Granada, or the Free and Cued Selective Reminding Test [FCSRT][33] for patients from Malaga), language (short version of the Boston Naming Test,[34] semantic verbal fluency,[35] and understanding of commands), motor praxis (gesture imitation from the EULA test[36]), executive function (short-form version of the WAIS similarities subtest, coin test from the Eurotest,[19] semantic verbal fluency,[35] and TMT-B[31]), and visuospatial function (copy and drawing tasks from the CERAD[32]). Scores equal to or below the 5th percentile, with z-score = −1.5, or scaled scores $\leq$ 6, depending on the norms for each test, were considered pathological. This assessment was performed by a researcher (SLA, MFS, or MEG) blinded to the results of the BCTs. These individuals also underwent functional assessment (Barthel Index[37] and Lawton-Brody IADL scale[38]). The reference assessment and the BCTs were administered on the same day in both samples, except in 9 individuals from the external sample, who underwent these tests within a maximum of 8 days.

Diagnosis

Regardless of the reason for consultation and the final diagnosis, all individuals were classified according to their cognitive and functional status using the Global Deterioration Scale[39] (Supplementary material, Table 2); scores 1 and 2 were classified as ''no CI'' and scores $\geq$ 3 as ''CI.'' This diagnosis was established by a neurologist with experience in cognitive and behavioural neurology (CCP), based solely on the results from the formal cognitive assessment.

Models evaluated

We calculated diagnostic accuracy for the Fototest and Mini-Cog independently, and for the following combinations of both instruments.

Combined use.

- *Simple* (Comb-simple): single score resulting from the sum of the Fototest score and the Mini-Cog score.
- *Logistic regression model* (Comb-LR): includes Fototest and Mini-Cog scores as predictors, with sex, age, and level of education as covariates.
- *Random decision tree* (Comb-DT): we developed a classification model using a machine learning technique, applying the algorithm ''classification and regression tree'' (CART) for supervised learning with automatic Bayesian optimisation of accuracy (percentage of correct predictions) using the OptiML feature[40] of the free-access BigML platform.[41] Models were trained with a total of 9 predictors, including sociodemographic variables (age, sex, and level of education), Fototest scores (total score and scores for naming, free recall, total recall, and naming fluency), and Mini-Cog scores (total score and scores for recall and the clock-drawing task); the response variable was ''CI.''
- *Stepped use:* 2-step evaluation. In the first step, we considered 2 cut-off points for the Fototest. The lower cut-off point (Fototest $\leq$ 25) maximises the positive predictive value or precision, minimising or even eliminating false positives. The higher cut-off point (Fototest $\geq$ 40) maximises the negative predictive value, minimising or eliminating false negatives. Individuals

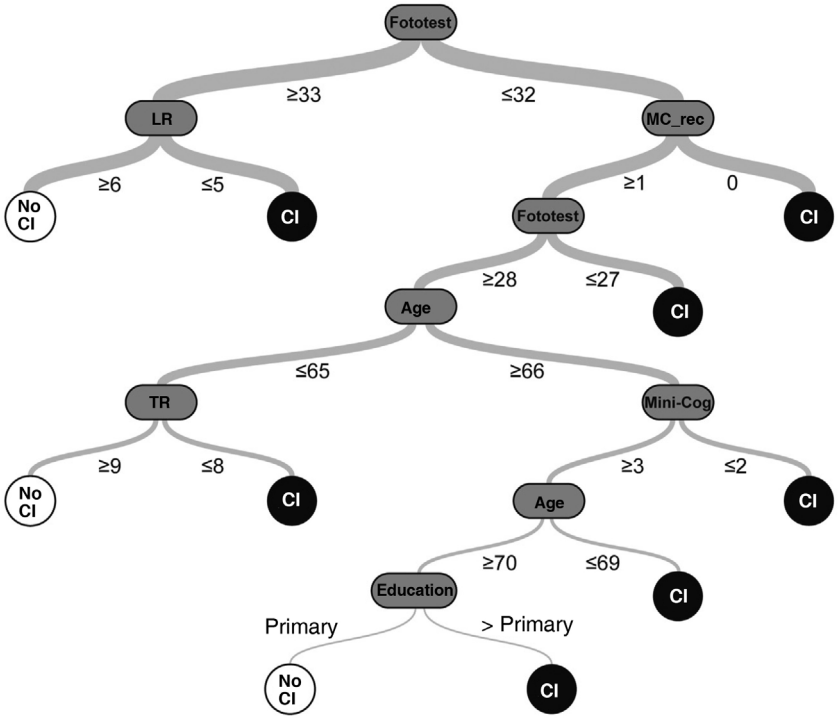


Figure 1 Random decision tree model.
CI: cognitive impairment; FR: Fototest free recall score; LR: logistic regression; MC_rec: Mini-Cog recall score; no CI: no cognitive impairment; TR: Fototest total recall score.

scoring $\leq$ 25 on the Fototest were classified as CI and those scoring $\geq$ 40 as no CI, without the Mini-Cog being administered. The second step considers Mini-Cog results, only in patients with Fototest scores 26—39, applying the previously mentioned combined models (Step-simple, Step-LR, Step-DT).

Statistical analysis

We conducted a descriptive study of sociodemographic variables and results; comparisons were performed using the *t* test and the chi-square test, according to whether variables were continuous or categorical.

The analysis was performed in the study sample; to prevent overadjustment and, consequently, overestimation of diagnostic accuracy, we randomly divided the study sample into 2 datasets. The first (base dataset) included 80% of participants and was used to create the different models and to perform an exploratory analysis enabling selection of the optimal cut-off points and criteria, whereas the second (test dataset), including the remaining 20%, was used for independent validation of the different models and criteria selected in the base dataset.

The diagnostic accuracy of the different models in the base dataset was evaluated with the area under the curve (AUC) and the corresponding standard error (SE) for CI vs no CI; AUCs were compared according to the method proposed by Hanley and McNeil.[42] The optimal cut-off point was that which maximised the accuracy of the test (percentage of correct classifications). To select the Comb-DT model, given that many of the 98 models created presented similar accuracy values, we decided to maximise the phi coefficient (Matthew's correlation coefficient) and, according to the principle of parsimony, to select the model with the smallest number of nodes.

The assessment and internal validation of the models and criteria selected in the base dataset were performed independently on the test dataset, and the external validation was performed on the external sample, using a pre-specified analysis of diagnostic accuracy with the parameters sensitivity, precision, phi coefficient, and accuracy; these parameters are frequently used in the validation of prediction models. These data enabled the calculation of classical measures of the diagnostic accuracy of a test (specificity, predictive values, likelihood ratios), and the creation of the corresponding contingency table.

Statistical analysis was performed with SPSS, version 19.0.0,[43] and MedCalc, version 18.9.1.[44]

Formal considerations

Study design and manuscript drafting followed the recommendations of the STARD 2015 guidelines[45] for reporting diagnostic accuracy studies, the STARDdem statement[46] for reporting the diagnostic accuracy of studies into cognitive disorders, and the guidelines for the development of predictive models in biomedicine.[47]

The study complies with the ethical principles for medical research established in the latest revision of the World Medical Association's Declaration of Helsinki (Edinburgh, 2000).

Results

The study sample included 448 individuals with a mean age (SD) of 60.5 (17.7) years, with a slight predominance of women (53.8%) and individuals with more than primary education (52.2%); the prevalence of CI in this sample was 46.4%. The 2 randomly generated datasets (base and test, including 358 and 90 individuals, respectively) showed no significant differences in any of the sociodemographic variables studied, in the prevalence of CI, or in Fototest and Mini-Cog results.

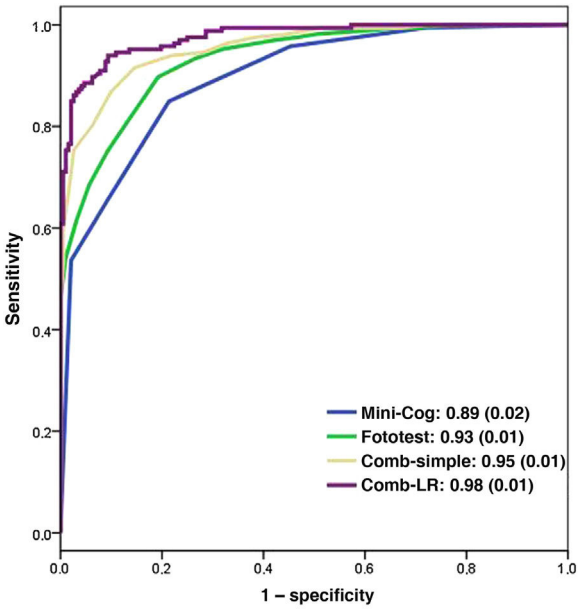


Figure 2 Diagnostic accuracy of the Fototest and Mini-Cog, used alone or in combination.
Comb-LR: logistic regression combined model; Comb-simple: simple combined model.

The external sample, in contrast, presented a much higher prevalence of CI (77.0%), an older mean age (71.7 [10.0] years), a higher education level (60.7% of individuals had at least secondary education), and a lower prevalence of women (47.5%), although differences were not significant in the latter 2 variables. Fototest and Mini-Cog results were significantly lower in the external sample than in the study sample.

Table 1 summarises the sociodemographic characteristics, prevalence of CI, and Fototest and Mini-Cog results of the 2 samples and the 2 datasets.

The random decision tree model selected (among the 98 models generated by the platform) had an accuracy of 92.2% and a phi coefficient of 0.84 in the test sample, with a total of 7 levels (Fig. 1) (Supplementary Material, Table 3).

Table 2 summarises the diagnostic parameters of the different models in the different samples and datasets. The diagnostic accuracy of the Fototest was moderate, but significantly higher than that of the Mini-Cog (0.93 ± 0.01 [AUC ± SE] vs 0.89 ± 0.02; t = 2.06; P = .04), while the diagnostic accuracy of the Comb-simple (0.95 ± 0.01) and Comb-LR models (0.98 ± 0.01) was significantly higher than that of either test alone (t > 4.00 and P < .0001 for all comparisons) (Fig. 2, Table 3). The Comb-DT model does not have an AUC, since it only includes one alternative for classification; however, its accuracy (95.2%) was significantly higher than that of the Mini-Cog (81.6%; χ^2 = 32.24; P < .0001), the Fototest (84.9%; χ^2 = 21.16; P < .001), and the Comb-simple (88.3%; χ^2 = 11.2; P = .001) and Comb-LR models (91.6%; χ^2 = 3.76; P = .05). Stepped models present the exact same diagnostic accuracy as the combined models they include, with the added benefit of significantly shorter administration times (197.3 [56.7] vs 233.9 [45.2]; t = 8.8; P < .0001).

The validation of the models in the test dataset largely replicates the predictions of the models generated in the base dataset, except for the models using random decision trees, which present a slightly lower diagnostic accuracy, as may be expected due to the overadjustment effect. The models also showed a very high diagnostic accuracy in the external sample, with accuracy values ranging from 86.9% to 91.8% (except for the Mini-Cog, at 78.7%), although the difference between the external sample and the base dataset was not statistically significant (81.6%; χ^2 = 0.29; P = .59).

The Supplementary Material (Tables 4—11) presents the contingency tables and classical measures of diagnostic accuracy (sensitivity, specificity, predictive values, likelihood ratios) for each model and in each sample for the selected cut-off points and criteria.

Discussion

The results of this prospective study, which was conducted under conditions of routine clinical practice, clearly show that the combined use of the Fototest and Mini-Cog surpasses the diagnostic accuracy of either instrument alone. These results are consistent, as they were replicated in 2 independent samples: one with the same sociodemographic and clinical characteristics as the sample used to develop the models, and another sample that was completely different in terms of sociodemographic and clinical characteristics (drawn from a cognitive-behavioural neurology consultation, older age, higher prevalence of CI, and different geographical location). The greater diagnostic accuracy of the combined models is probably due to greater comprehensiveness of the cognitive assessment, as these models evaluate more cognitive domains. This, in turn, facilitates the identification of different profiles of CI, which could enhance diagnostic studies. However, our study did not address this issue, which should be explored in future research.

Our results are also consistent with and support the available data from Spain regarding the diagnostic accuracy of the instruments used. The only previous study of Mini-Cog in Spain reported an AUC of 0.88 (SE: 0.01),[29] which is practically identical to that found in our study (0.89 [0.02]); one study reconstructed the Mini-Cog score from the Clock-Drawing Test and the Mini—Mental State Examination, so we do not consider it comparable.[28] Similarly, the diagnostic accuracy of the Fototest in the present study (0.93 [0.01]) is very similar to that reported in previous studies in Spain (0.86 [0.02],[48] 0.93 [0.02],[49] 0.95 [0.01][20]) and even in Argentina (0.93 [0.03][50]).

The complex models (Comb-LR and Comb-DT) provide higher diagnostic accuracy than the Comb-simple model, probably due to the weighting of diagnostic contributions from each test and the inclusion of sociodemographic variables (sex, age, education level) in these models. These variables can be especially relevant, particularly in the case of the Mini-Cog.[29] These complex models have the disadvantage that they involve a computational process that may be difficult to apply during consultations. The simple model, in contrast, is more practical for use in the clinical setting, as it simply requires adding the scores from both instruments. However, one of the authors (RVC) developed a program to facilitate the application of these instruments, incorporating a calculator that performs these computations automatically, providing a report that includes all results (Supplementary Material, Fig. 1). This application, Predi-Cog[51] (https://www.hipocampo.org/Predi-Cog.asp), is available online, free of cost, at www.hipocampo.org.[52]

Based on our results, we recommend applying the tests in a stepped manner (Fig. 3), which allows direct classification of a substantial percentage of individuals, without compromising diagnostic accuracy, using only the Fototest (base dataset, 36.6%; test dataset, 37.8%; external sample, 42.6%), saving time (197 [56.7] vs 233.9 [45.2] seconds).

Our study has some limitations, such as the exclusion of individuals without at least primary education, which limits the possibility of extrapolating our results to the population with lower education levels. In any case, our previous data already recommended avoiding the use of Mini-Cog in this population.[28,29] Another Fototest weakness is that only 63.9% of the individuals in the study sample underwent formal cognitive assessment. However, we only excluded

Table 1 Sociodemographic characteristics and cognitive test results, by sample and dataset.

	Base dataset (80%)	Test dataset (20%)	Study sample	External sample	[a]	[b]
No. subjects	358	90	448	61		
CI	166 (46.4%)	42 (46.7%)	208 (46.4%)	47 (77.0%)	0.03 (0.96)	20.14 (0.0001)
Age (years)	60.7 (17.4)	60.1 (18.7)	60.7 (17.7)	71.7 (10.0)	0.08 (0.78)	23.28 (0.0001)
Sex (women)	193 (53.9%)	48 (53.3%)	241 (53.8%)	29 (47.5%)	0.01 (0.92)	0.84 (0.36)
Education level						
Primary education	171 (47.8%)	46 (51.1%)	217 (48.4%)	24 (39.3%)	0.31 (0.57)	1.78 (0.18)
> Primary education	187 (52.2%)	44 (48.9%)	231 (51.6%)	37 (60.7%)		
GDS						
1	130 (36.3%)	32 (35.6%)	162 (36.2%)	—		
2	62 (17.3%)	16 (17.8%)	78 (17.4%)	14 (23.0%)		
3	75 (20.9%)	19 (21.1%)	94 (21.0%)	28 (45.9%)	4.30 (0.51)	37.78 (0.0001)
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%)		
FCA	228 (63.7%)	58 (64.4%)	286 (63.8%)	61 (100%)	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)
Time (s)	130.2 (20.2) (*n* = 317)	128.7 (19.3) (*n* = 83)	129.9 (20.0) (*n* = 400)	136.7 (14.8) (*n* = 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)
Time (s)	103.3 (33.7) (*n* = 310)	100.2 (34.9) (*n* = 81)	102.7 (33.9) (*n* = 391)	109.7 (34.8) (*n* = 48)	0.42 (0.52)	1.84 (0.17)

Data are presented as either number (%) or mean (standard deviation). CI: cognitive impairment; FCA: formal cognitive assessment; GDS: Global Deterioration Scale. Data are presented as either number (%) or mean (standard deviation).

[a] Base dataset vs test dataset, expressed as χ^2 or t (*P*-value).

[b] Study sample vs external sample, expressed as χ^2 or t (*P*-value).

Table 2 Diagnostic accuracy in the base dataset and assessment in the test dataset and external sample.

Test sample	Base dataset (n = 358; 166 CI)				Test dataset (n = 90; 42 CI)				External sample (n = 61; 47 CI)			
	AUC	Cut-off	Accuracy	Time	Sens.	Precision	Phi	Accuracy	Sens.	Precision	Phi	Accuracy
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)		0.93	0.85	0.78	88.9	0.91	0.98	0.79	91.8
Comb-LR[a]	0.98 (0.01)	≥ 0.50	91.6 (88.2—94.3)	233.9 (45.2)	0.95	0.95	0.91	95.6	0.91	0.96	0.74	90.2
Comb-DT	—	[b]	95.2 (92.5—97.2)		0.90	0.93	0.87	92.2	0.91	0.93	0.68	88.5
Step-simple	—	[c]	88.3 (88.5—91.4)		0.93	0.85	0.78	88.9	0.91	0.98	0.79	91.8
Step-LR	—	[d]	91.6 (88.2—94.3)	197.3 (56.7)	0.95	0.94	0.89	94.4	0.94	0.94	0.72	90.2
Step-DT	—	[e]	95.2 (92.5—97.2)		0.90	0.93	0.87	92.2	0.91	0.91	0.63	86.9

Results for the base dataset are expressed as mean (standard deviation) or percentage (exact confidence interval).
AUC: area under the curve; CI: cognitive impairment; Comb: combined model; DT: decision tree; LR: logistic regression; Phi: Matthew's correlation coefficient; Sens.: sensitivity; Step: stepped model.

[a] Prob CI = 12.02 − 0.47 Fototest − 1.11 Mini-Cog − 2.06 education (> primary) + 0.07 age (years) − 0.28 sex (woman); if Prob ≥ 0.5: CI; if Prob < 0.5: no CI.

[b] See random decision tree (Fig. 2).

[c] Step 1: Fototest ≤ 25 CI & Fototest ≥ 40 no CI; step 2: 26 ≥ Fototest ≤ 39: FotoCog ≤ 34 CI, ≥ 35 no CI.

[d] Step 1: Fototest ≤ 25 CI & Fototest ≥ 40 no CI; step 2: 26 ≥ Fototest ≤ 39: apply [a].

[e] Step 1: Fototest ≤ 25 CI & Fototest ≥ 40 no CI; step 2: 26 ≥ Fototest ≤ 39: apply [b].

Table 3 Diagnostic accuracy of the Fototest and Mini-Cog, used alone or in combination.

	Fototest	Mini-Cog	Comb-simple	Comb-LR
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-LR				—

Results are expressed as the difference between AUCs (SE) (z; *P*).
AUC: area under the curve; LR: logistic regression; SE: standard error.

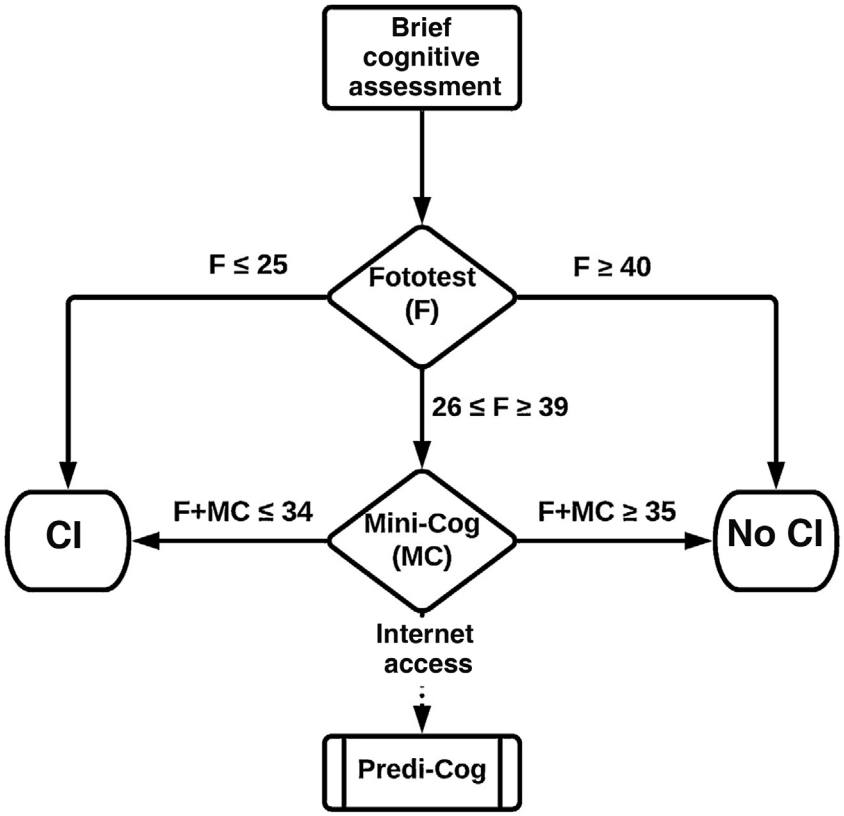


Figure 3 Recommended strategy for combined use of the Fototest and Mini-Cog.
CI: cognitive impairment; No CI: no cognitive impairment.

individuals without cognitive complaints and with Fototest scores above the 10th percentile, which makes it unlikely that these individuals would have CI, thus minimising the risk of partial verification bias.[53] Lastly, while the index and reference assessments were performed by different professionals who were blinded to other results, and diagnosis was based on reference test results, the professional who issued the diagnosis was not blinded to the results of the BCTs. Our study also has several strengths, such as its naturalistic, pragmatic nature, the large sample size, and, most importantly, the double validation of the results in 2 independent samples with very different characteristics.

In conclusion, the combined use of the Fototest and Mini-Cog, 2 very brief tests that are simple to administer and that evaluate all the cognitive functions that should be addressed in a brief cognitive assessment, enhances the already high diagnostic accuracy of both instruments. Stepped use of these tests is more efficient than their combined use, since it achieves a significant decrease in mean administration time (< 200 seconds) while maintaining accuracy.

Data availability statement

The data that support the findings of this study are available upon reasonable request.

Conflicts of interest

C. Carnero Pardo is the creator of the Fototest. Under the terms of its Creative Commons licence, the Fototest may be used and distributed for non-commercial purposes provided that it is not modified and its authorship is explicitly acknowledged.

Acknowledgements

All figures were created by Marc Torres Ciuró.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi: https://doi.org/10.1016/j.nrleng.2023.10.002.

References

1. Lopez-Pousa S, Monserrat-Vila S, Turro-Garriga O, Aguilar-Barbera M, Caja-Lopez C, Vilalta-Franch J, et al. Analisis de la demanda asistencial neurologica generada por la atencion primaria en un area geografica de las comarcas de Girona. Rev Neurol. 2009;49:288—94.
2. Henderson WE. Cognitive assessment in neurology. In: Aminoff MJ, Boller F, Swaab DF, editors. Handbook of Clinical Neurology (3rd series): History of Neurology. Elsevier B. V.; 2010. p. 235—56. Cap: 17.
3. 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.
4. Rego García I, Medina Gámez J, Valderrama Martón C, Guillén Martínez V, Vílchez Carrillo R. C CP. ''Don't know'' sign: Description and evaluation of its diagnostic accuracy for cognitive impairment, comparint to other observation based signs. European Academy of Neurology. 2020:EPO201.
5. 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.
6. 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. Neurologia. 2002;17:17—32.
7. 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.
8. American Psychiatric Association (APA). Manual diagnóstico y estadístico de los trastornos mentales (DSM-5). In: Editorial Médica Panamericana. 5ª ed; 2013.
9. Peña Casanova J, Sánchez Benavides G. Test cognitivos breves: una confusión, una necesidad y una historia interminable. In: Carnero Pardo C, editor. Test cognitivos breves. Madrid: Ediciones SEN; 2015. p. 9—18. Cap: 1.
10. 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.
11. Carnero-Pardo C, Rego-Garcia I, Mene Llorente M, Alonso Rodenas 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).
12. 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.
13. Lozano-Gallego M, Hernández Ferrándiz M, Turró Garriga O, Pericot Nierga 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.
14. Mathuranath PS, Nestor PJ, Berrios GE, Rakowicz W, Hodges JR. A brief cognitive test battery to differentiate Alzheimer's disease and frontotemporal dementia. Neurology. 2000;55:1613—20.
15. 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.
16. Morera-Guitart J, Escudero J, Aguilar M, Aguilera JM, Carnero C, Martin 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.
17. 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.
18. 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, 0022-3956(75)90026-6 [pii].
19. Carnero-Pardo C, Montoro-Ríos MT. Evaluación preliminar de un nuevo test de cribado de demencia (Eurotest). Rev Neurol. 2004;38:201—9.
20. Carnero Pardo C, Sáez-Zea C, Montiel Navarro L, Del Saz P, Feria Vilar I, Pérez Navarro MJ, et al. Utilidad diagnostica del Test de las Fotos (Fototest) en deterioro cognitivo y demencia. Neurologia. 2007;22:860—9.
21. 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.
22. Carnero-Pardo C, Rego-Garcia 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. Neurologia. 2019, http://dx.doi.org/10.1016/j.nrl.2018.12.002 (online ahead of print).
23. Carnero-Pardo C. ¿Es hora de jubilar al Mini-Mental? Neurologia. 2014;29:473—81, http://dx.doi.org/10.1016/j.nrl.2013.07.003.
24. Solomon PR, Hirschoff 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.
25. Sackett DL, Haynes RB. The architecture of diagnostic research. BMJ. 2002;324:539—41.
26. Patsopoulos NA. A pragmatic view on pragmatic trials. Dialogues Clin Neurosci. 2011;13:217—24.
27. Larner A. Dementia in clinical practice: A neurological perspective: Pragmatic studies in the cognitive function clinic. 3rd ed. Springer; 2018.
28. Carnero-Pardo C, Cruz-Orduna I, Espejo-Martinez B, Martos-Aparicio C, Lopez-Alcalde S, Olazaran 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.
29. 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).

30. Carnero Pardo C, Carrera Munoz 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;57:187—91.
35. Herrera-Garcia JD, Rego-Garcia I, Guillen-Martinez V, Carrasco-Garcia 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 [Accessed 22 June 2020]. Available from: OptiML with the BigML dashboard; 2020 https://static.bigml.com/pdf/BigML_OptiML.pdf?ver=c306567.41
41. BigML Inc. Machine Learning Platform [online]. [Accessed 22 June 2020]. Available from: https://bigml.com/.
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 SPSS Statistic [computer program]. Version 19.0.0. USA: SPSS Inc.; 2010 https://www.ibm.com/products/spss-statistics
44. MedCalc Statistical Software versión 18.9.1 [computer program]. Ostend, Belgium: MedCalc Software bvba; 2018 http://www.medcal.org
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 F. 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. Neurologia. 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. 2014:1—15, http://dx.doi.org/10.1080/13854046.2014.928748.
51. De la Vega Cotarelo R. Predi-Cog [online]. [Accessed 27 October 2020]. Available from: https://www.hipocampo.org/Predi-cog.asp.
52. De la Vega Cotarelo R. La Circunvalación del Hipocampo [online]. [Accessed 27 October 2020]. Available at: 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.