ORIGINAL ARTICLE

Diagnostic performance of brief cognitive tests in cognitive impairment screening[☆]

C. Carnero-Pardo [a,*], I. Rego-García [b], M. Mené Llorente [c], M. Alonso Ródenas [c], R. Vílchez Carrillo [b]

[a] FIDYAN Neurocenter, Granada, Spain
[b] Servicio de Neurología, Hospital Universitario Virgen de las Nieves, Granada, Spain
[c] Centro de Salud Casería de Montijo, Granada, Spain

Abstract

Introduction and objectives: This study aims to assess and compare the diagnostic performance of brief cognitive tests for cognitive impairment (CI) screening recommended by the Spanish guidelines for the integral care of people with Alzheimer's disease and other dementias.

Material and methods: We performed a phase iii study into the accuracy of diagnostic tests, including patients with suspected CI in a primary care setting. All patients completed the Mini—Mental State Examination (MMSE), the Mini Examen Cognoscitivo (MEC), the Short Portable Mental Status Questionnaire (SPMSQ), the Memory Impairment Screen (MIS), the Clock Drawing Test (CDT), the Eurotest, the Fototest, and the Memory Alteration Test (M@T). CI was diagnosed independently by researchers blinded to scores on these tests. Diagnostic performance was evaluated by calculating the area under the receiver operating characteristic curve (AUC).

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

Conclusion: The most advisable tests for CI screening in primary care are the Eurotest, M@T, and Fototest, with the latter being the most efficient as it takes half as long to administer.

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

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

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

Resumen

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

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

Resultados: Se han incluido 141 sujetos (86 con DC). El Eurotest y el T@M (0.91 ± 0.02 [aROC $\pm$ ee] y 0.90 ± 0.02 respectivamente), los instrumentos que requieren más tiempo (7.1 ± 1.8 (media $\pm$ de) y 6.8 ± 2.2 minutos respectivamente) tienen una UD significativamente superior a la del Mini-Mental, MEC, SPMSQ y TdR, pero no a la del MIS y Fototest (0.87 ± 0.03 ambas) requiriendo éste último menos de la mitad del tiempo (2.8 ± 0.8 minutos). T@M y MIS sólo evalúan memoria y el último no es aplicable a analfabetos.

Conclusiones: Los instrumentos más recomendables para el cribado de DC en Atención Primaria son Eurotest, T@M y Fototest, siendo el último más eficiente por requerir la mitad de tiempo.
© 2019 Sociedad Española de Neurología. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Cognitive impairment (CI) in general, and Alzheimer disease in particular, pose considerable challenges for healthcare systems and for society due to their high prevalence and associated costs.[1] In Spain, CI affects 14.5% to 19% of individuals older than 65 years,[2-4] although many cases go undetected (over half, according to the recent Gómez de Caso study[4]); despite this, there is broad consensus that the current circumstances do not meet the minimum required standard to recommend screening for CI in asymptomatic individuals.[5-9] However, most clinical practice guidelines agree on the value of maintaining an appropriate level of vigilance in primary care and performing screening with brief cognitive tests (BCT) if CI is suspected,[7,8] whether due to subjective complaints, complaints from the patient's caregiver or family members, or clinical suspicion among healthcare professionals.

In Spain, these recommendations are reflected in the National Health System's (SNS, for its Spanish abbreviation) clinical practice guidelines on the integral care of patients with Alzheimer disease and other dementias,[10] promoted by the Ministry of Health, Social Services, and Equality (hereinafter, SNS guidelines), which recommend the use of BCTs in patients with suspected CI. Specifically, the document recommends the normalised Spanish-language version of the Mini—Mental State Examination (MMSE)[11] or the Mini—Examen Cognoscitivo (MEC)[12] (grade of recommendation A); secondly (grade of recommendation B), it recommends such other instruments as the Short Portable Mental State Questionnaire (SPMSQ),[13] the Memory Impairment Screen (MIS),[14] the Seven-Minute Screen (7MS),[15] the clock-drawing test (CDT),[16] the Eurotest,[17] the Fototest,[18] and the Memory Alteration Test (M@T).[19] The SNS guidelines do not establish specific recommendations on cut-off points or corrections, and recommend that instruments be selected according to the available time, clinical experience, and the availability of normative data and validation studies for the language and setting in which they will be applied (good clinical practice criterion). However, the SNS guidelines do not mention certain other characteristics that may also be relevant when selecting a test; these include whether an instrument can be applied to illiterate individuals, whether it evaluates one or several cognitive domains, and the potential costs associated with its use (Table 1).[20]

Few studies have evaluated and compared these instruments in Spain,[21-24] and no study to date has used a single sample to compare the diagnostic performance and effectiveness in everyday practice of the different BCTs recommended in the SNS guidelines. This study aims to evaluate and compare the diagnostic performance and effectiveness of the different screening strategies recommended by the SNS guidelines in a prospective cohort of individuals attended at primary care centres, with full, independent verification of the diagnosis.

Material and methods

Design

We conducted a phase III study for the evaluation of diagnostic tests,[25] with a paired design (all BCTs were administered to all participants) and full verification (all participants, regardless of BCT results, were studied according to the gold-standard diagnostic procedure described in the Procedure section).

Setting

The sample was drawn from 4 healthcare centres in the Granada Norte metropolitan district (Almanjáyar, Casería de Montijo, Cartuja, and Salvador Caballero). The diagnostic study was performed at the cognitive and behavioural neurology unit at the neurology department of Hospital Universitario Virgen de las Nieves.

Table 1 Characteristics of the brief cognitive tests recommended in the Spanish National Health System's clinical practice guidelines on the integral care of patients with Alzheimer disease and other dementias.

Time	Test	Applicable to illiterate individuals	Multi-domain	Studies in PC	Validated for CI	Normative data	Associated cost
< 5 min	CDT	No	Yes	Yes	Yes	No	No
	SPMSQ	Yes	Yes	Yes	No	No	No
	Fototest	Yes	Yes	Yes	Yes	Yes	No
	MIS	No	No	Yes	Yes	No	No
	MMSE/MEC	No	Yes	Yes	Yes	Yes	Yes
> 5 min	M@T	Yes	No	Yes	Yes	No	No
	Eurotest	Yes	Yes	Yes	Yes	Yes	No
> 10 min	7MS	No	Yes	No	No	Yes	Yes

7MS: Seven-Minute Screen; CDT: clock-drawing test; CI: cognitive impairment; M@T: Memory Alteration Test; MEC: Mini—Examen Cognoscitivo; MIS: Memory Impairment Screen; MMSE: Mini—Mental State Examination; PC: primary care; SPMSQ: Short Portable Mental State Questionnaire.

Study population

We recruited individuals with suspected CI attended at healthcare centres.

Inclusion and exclusion criteria

Over a period of one year, we systematically included consecutive patients attended at healthcare centres with suspected CI reported by their primary care physician. Suspicion of CI may be based on subjective complaints of memory loss or cognitive alterations, similar complaints reported by a family member or informant, or signs or symptoms of CI observed by the patient's primary care physician. We excluded individuals who did not consent to study inclusion, who had previously participated in the study, or who had existing CI. We applied no exclusion criteria based on age, sensory or motor deficits, or any pre-existing condition, including neurological diseases.

Procedure

All selected participants were administered all BCTs recommended by the SNS guidelines, except the 7MS (MMSE, MEC, SPMSQ, MIS, CDT, Eurotest, Fototest, and M@T), in one of 3 different orders. Orientation in time and space were tested once, and the corresponding results were applied to the different instruments collecting this information (MMSE, MEC, SPMSQ, and M@T). The MEC score was derived: items specific to the MEC were administered after application of the MMSE, and the results were recalculated for the 30- and 35-item forms of the MEC (MEC-30 and MEC-35, respectively). The CDT was scored according to the system used for the 7MS.[15] We timed the administration of the Eurotest, Fototest, CDT, M@T, and MMSE (orientation items were not timed for the latter 2 tests).

Independently of BCT results, all participants were referred to the neurology consultation for comprehensive behavioural (Spanish-language adaptation of the Neuropsychological Inventory questionnaire[26]), functional (Barthel index,[27] Lawton-Brody scale,[28] and Pfeffer Functional Activities Questionnaire[29]), and neuropsychological evaluation (orientation and attention/executive function [digit span, similarities, semantic verbal fluency[30]], verbal memory [CERAD word memory list[31]], language [Boston Naming Test short form,[32] semantic verbal fluency,[30] and understanding of instructions], and motor[33] and visuoconstructive praxis [CERAD copy task[31]]). This evaluation was conducted by independent professionals blinded to participants' results in the BCTs; none of the BCTs analysed in this study were included in the formal neuropsychological evaluation. The maximum time between application of BCTs and evaluation at the neurology consultation was 2 weeks.

Reference diagnosis

The gold-standard diagnosis was independently established by 2 neurologists with experience in cognitive-behavioural neurology; based on thorough clinical assessment and on the results of the cognitive, behavioural, and functional evaluation, participants were classified as having no CI, mild CI (MCI), or dementia. Diagnoses of MCI and dementia were based on the criteria of the Spanish Society of Neurology's Study Group for Behaviour and Dementia,[34] which are also included in the SNS guidelines. These neurologists were blinded to participants' results in the BCTs. Disagreements were resolved by consensus.

Data analysis

Diagnostic performance was evaluated according to the area under the receiver operating characteristic curve (AUC) for each BCT and compared using deLong's method for comparing AUCs from the same sample.[35] The threshold for statistical significance was set
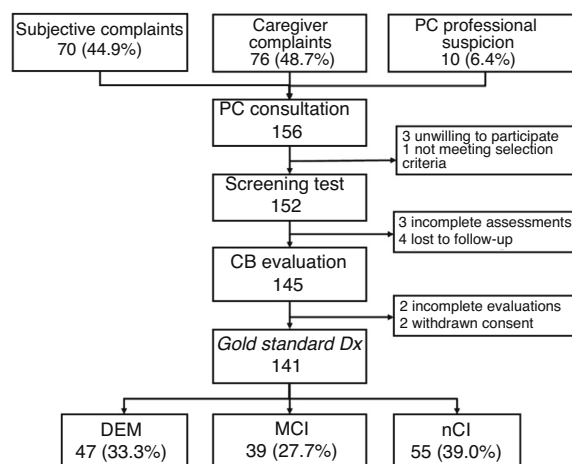

Figure 1 Participant flow chart.
CB: cognitive/behavioural; DEM: dementia; Dx: diagnosis; MCI: mild cognitive impairment; nCI: no cognitive impairment; PC: primary care.

at $P \leq .05$. Effectiveness was calculated for discrimination between CI (MCI and dementia) and no cognitive impairment. Sensitivity and specificity were calculated for the optimal cut-off point for each instrument (the score with the highest proportion of correct diagnoses; if one or more points presented the same proportion, we opted for that with the highest Youden index value[36]) and for cut-off points with the highest specificity and sensitivity ≥ 0.80. Statistical analysis was performed using the SPSS (SPSS Inc.; Chicago, IL, USA) and MedCalc 18.9 software.[37]

Formal considerations

The study was approved by our hospital's research ethics committee; all participants gave written informed consent. Study design and drafting of the report comply with the STARD recommendations for studies evaluating diagnostic tests[38] and with the recommendations of the United States Food and Drug Administration for the reporting of studies evaluating diagnostic tests.[36]

Results

We selected a total of 156 individuals (70 [44.9%] with subjective cognitive complaints, 76 [48.7%] with complaints raised by caregivers, and 10 [6.4%] with suspicion reported by the primary care physician. Three did not consent to be included in the study; one did not meet the selection criteria (previous diagnosis of CI); 4 did not attend the neurology consultation; 2 withdrew consent during the study period; 3 could not be evaluated due to sensory deficits and sequelae of previous strokes; and 2 were uncooperative in the evaluation due to behavioural alterations (Fig. 1).

Table 2 shows the sociodemographic characteristics of the final sample of 141 participants (55 [39.0%] without CI, 39 [27.7%] with MCI, and 47 [33.3%] with dementia); mean age (standard deviation; range) was 72.0 years (11.5; 31-90), with a clear predominance of women (73.0%) and individuals with low levels of schooling (14.2% of participants were illiterate and only half [49.6%] had completed primary education or higher). The table also shows the results and application times for the different BCTs administered, stratified by cognitive diagnosis. As expected, the results show a considerable difference between groups in all cases, with poorer scores and longer application times in individuals with CI.

All participants (n = 141) completed the MMSE and MEC; 140 completed the Eurotest and SPMSQ; and 139 completed the Fototest and M@T. Twenty-two participants (20 of whom were illiterate) were unable to complete the MIS; the CDT was only administered to 126 participants as 15 illiterate individuals were unwilling to complete it.

Table 3 summarises the results for the different instruments, indicating the number of participants who completed each one, application time, diagnostic performance (AUC), sensitivity and specificity, and the proportion of correct diagnoses for the selected cut-off points.

Diagnostic performance was compared for pairs of BCTs in the subsamples of participants who completed both (Table 4). The results show that there was no significant difference between diagnostic performance for the Eurotest and the M@T, but that both performed significantly better than the MMSE, MEC-30, MEC-35, and SPMSQ; the M@T also performed significantly better than the CDT, with the result of the comparison between the Eurotest and the CDT being on the verge of significance (*P* = .08). However, the Eurotest and M@T did not perform significantly better than the Fototest and the MIS.

Fig. 2 shows the administration times for each of the BCTs studied (except the SPMSQ and the MIS, for which administration was not timed). No significant difference was observed in application times for the Eurotest and M@T (7.1 [1.8] vs 6.8 [2.2] minutes); the Fototest (2.8 [0.8] minutes) and CDT (1.3 [0.7] minutes) were clearly quicker to administer. The administration time for the MMSE (5.5 [1.6] minutes) is underestimated, as we did not time the items related to orientation in time and space; despite this, it took significantly longer to administer than the Fototest and CDT, needing longer than 5 minutes. We would expect the MEC, which includes more items than the MMSE, to take even longer.

Discussion

The results from this phase III study evaluating diagnostic tests, following a prospective, paired methodology with full verification, show that the Eurotest and M@T took the longest to apply (7.1 [1.8] and 6.8 [2.2] minutes, respectively) and presented the best diagnostic performance (0.91 [0.02] and 0.90 [0.02], respectively), with no significant difference between the 2; both tests correctly classified 84% of participants. Both instruments were developed and validated in Spain, can be administered to illiterate individuals, and are freely available.[17,19] However, the M@T presents certain limitations: no formal normative study has been conducted, and it only evaluates memory, and therefore would not be sensitive to CI affecting other cognitive domains. The M@T presented very similar diagnostic performance to that reported in a study of primary care patients with memory complaints (AUC of 0.91), although the overall scores in our sample are lower, probably due to the low level of education.[39]

The MMSE, MEC-30, and MEC-35, which are the most widely used instruments and the most strongly recommended in the SNS guidelines, and which also presented application times longer than 5 minutes (at least 5.5 [1.6] minutes), displayed significantly poorer diagnostic performance than the M@T and the Eurotest. The low diagnostic yield for CI of these tests, whose use is so universal and generalised, is more a rule than an exception, and has been reported in numerous Spanish and international studies.[40—42] Our results are consistent with those of previous studies by our research group, showing a lower optimal cut-off point (22/23) than that which is typically used.[42] The poor diagnostic performance, long application times, and other disadvantages (influence of level of education, etc), above all the costs associated with the MMSE, have led many authors to cease to consider it the instrument of choice[43]; other authors even omit the test from their recommendations.[44]

Table 2 Sociodemographic characteristics and scores in the brief cognitive tests by diagnostic group.

	Total	CI (MCI + DEM)	nCI	MCI	DEM
No. participants	141	86	55	39	47
Sex (women)	103 (73.0%)	59 (68.6%)	44 (80.0%)	23 (59.0%)	36 (76.6%)
Age (years)	72.0 (11.5)	77.1 (7.5)	64.6 (12.2)	74.3 (7.3)	68.4 (11.6)
Education level					
Illiterate	20 (14.2%)	19 (22.1%)	1 (1.8%)	5 (12.8%)	14 (29.8%)
Incomplete primary education	51 (36.2%)	35 (40.7%)	16 (29.1%)	16 (41.0%)	19 (40.4%)
Completed primary education	69 (49.6%)	32 (37.2%)	38 (69.1%)	18 (50.0%)	14 (29.8%)
MMSE (n = 141)	19.9 (5.7)	17.3 (5.3)	24.1 (3.2)	21.5 (3.7)	13.7 (3.7)
Time (min)[a]	5.5 (1.6)	5.9 (1.5)	4.8 (1.4)	5.8 (1.7)	6.0 (1.4)
MEC (n = 141)	25.4 (7.2)	22.1 (7.2)	30.6 (2.9)	27.4 (4.8)	17.8 (5.9)
Eurotest (n = 140)	20.2 (10.3)	14.7 (9.2)	28.6 (4.4)	22.3 (6.0)	8.3 (6.1)
Time (min)	7.1 (1.8)	7.7 (1.5)	6.2 (1.9)	7.7 (1.7)	7.8 (1.3)
M@T (n = 139)	29.0 (12.5)	22.3 (11.0)	39.3 (6.0)	30.9 (8.3)	15.2 (7.2)
Time (min)[a]	6.8 (2.2)	7.6 (2.0)	5.6 (1.9)	6.7 (1.9)	8.4 (1.9)
Fototest (n = 139)	29.1 (7.7)	25.4 (6.8)	34.7 (5.0)	30.0 (4.6)	21.4 (5.8)
Time (min)	2.8 (0.8)	2.9 (0.8)	2.5 (0.7)	2.5 (0.5)	3.3 (0.9)
CDT (n = 126)	4.4 (2.6)	3.1 (2.6)	6.2 (1.4)	4.7 (2.2)	1.6 (2.0)
Time (min)	1.3 (0.7)	1.5 (0.7)	1.0 (0.6)	1.3 (0.6)	1.6 (0.8)
SPMSQ (n = 140)	2.6 (2.6)	3.8 (2.7)	0.6 (0.9)	1.8 (1.6)	5.6 (2.1)
MIS (n = 119)	4.2 (2.9)	2.6 (2.5)	6.3 (1.8)	4.0 (2.5)	1.0 (1.4)

Data are expressed either as absolute frequency (%) or as mean (standard deviation).
CDT: clock-drawing test; CI: cognitive impairment; DEM: dementia; M@T: Memory Alteration Test; MCI: mild cognitive impairment; MEC: Mini−Examen Cognoscitivo; MIS: Memory Impairment Screen; MMSE: Mini−Mental State Examination; nCI: no cognitive impairment; SPMSQ: Short Portable Mental State Questionnaire.
[a] Orientation items were not timed.

Table 3 Diagnostic performance and effectiveness of the brief cognitive tests.

Test	n	Time (min)	AUC	Optimal cut-off point				Cut-off point with sensitivity ≥ 0.80			
				COP	Sensitivity	Specificity	PCD	COP	Sensitivity	Specificity	PCD
MMSE	141	5.5 (1.6)[a]	0.85 (0.03)	20/21	0.71	0.87	0.77	22/23	0.80	0.65	0.75
MEC-35	141	>5.5 (1.6)	0.85 (0.03)	27/28	0.70	0.87	0.77	30/31	0.85	0.62	0.76
MEC-30	141	>5.5 (1.6)	0.85 (0.03)	25/26	0.81	0.78	0.79	25/26	0.81	0.78	0.79
Eurotest	140	7.1 (1.8)	0.91 (0.02)	24/25	0.82	0.89	0.84	24/25	0.82	0.89	0.84
M@T	139	6.8 (2.2)[a]	0.90 (0.02)	32/33	0.81	0.89	0.84	32/33	0.81	0.89	0.84
SPMSQ	140	-	0.86 (0.03)	1/2	0.74	0.85	0.78	0/1	0.88	0.54	0.75
Fototest	139	2.8 (0.8)	0.87 (0.03)	30/31	0.76	0.80	0.78	32/33	0.84	0.69	0.77
CDT	126	1.3 (0.7)	0.83 (0.03)	4/5	0.64	0.91	0.76	6/7	0.86	0.59	0.75
MIS	119	-	0.87 (0.03)	5/6	0.82	0.81	0.81	5/6	0.82	0.81	0.81

AUC: area under the receiver operating characteristic curve; CDT: clock-drawing test; COP: cut-off point; M@T: Memory Alteration Test; MEC: Mini−Examen Cognoscitivo; MIS: Memory Impairment Screen; MMSE: Mini−Mental State Examination; n: number of participants; PCD: proportion of correct diagnoses; SPMSQ: Short Portable Mental State Questionnaire.
[a] Orientation items were not timed.

**Table 4** Diagnostic performance of the brief cognitive tests.

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

CDT: clock-drawing test; M@T: Memory Alteration Test; MEC: Mini−Examen Cognoscitivo; MIS: Memory Impairment Screen; MMSE: Mini−Mental State Examination; ns: not significant; SPMSQ: Short Portable Mental State Questionnaire.
For each comparison, we provide the *P*-value in the deLong test and the number of participants included.

The SPMSQ also presents significantly poorer diagnostic performance; however, the application time for this instrument (not measured in this study) is considerably shorter. While the SPMSQ is widely used to screen for dementia in Spain,13 its diagnostic performance in screening for CI has not previously been evaluated. Our results show a clear floor effect, with even scores of 0 having a sensitivity of 0.88; this would imply that 12% of individuals with CI present perfect execution (no errors), and therefore will never be detected by this instrument. Another limitation of the SPMSQ is that for 3 of the 10 items (telephone number, date of birth, and mother's maiden name), scoring requires reference to information that is often difficult to access or verify, particularly for older individuals attending the consultation alone.

The CDT, which presented the shortest application time (1.3 [0.7] minutes), had the poorest diagnostic performance (0.83 [1.03]), a significantly lower value; it also presents the disadvantage that it requires a certain level of graphomotor skill that individuals who are illiterate or have a low level of education may not possess; this would explain why 15 participants (10.6%), most of whom were illiterate, were unwilling to complete the test. These results confirm the previous recommendation that the CDT is not an appropriate BCT for individuals without a minimum level of education.[45] However, the M@T and the Eurotest presented no significant differences in diagnostic performance when compared to the Fototest and the MIS (0.87 [0.02] for both tests), which are far quicker to administer (2.8 [0.8] minutes for the Fototest). This confirms the results of previous studies[22,23] suggesting that these instruments are more efficient on account of their similar diag-
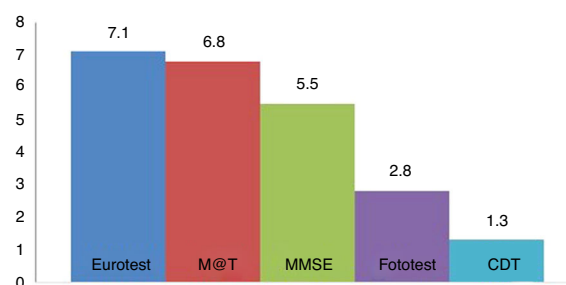
Figure 2 Time taken in minutes to complete each brief cognitive test.
MMSE: Mini−Mental State Examination; M@T: Memory Alteration Test; CDT: clock-drawing test.

nostic performance and shorter administration time. Despite this, the MIS presents several limitations: no normative study has been performed, it only evaluates memory, and it is not applicable to illiterate individuals. Therefore, the test cannot be administered to a considerable percentage of Spanish adults (14.2% in our sample).

The AUC was selected as a measure of diagnostic performance because it enables evaluation of an instrument's global discriminative capacity for all possible scores, and enables comparison of different instruments in a single sample,[46,47] independently of

the prevalence of the diagnosis and the sociodemographic characteristics of the sample (age, schooling, etc) and, therefore, independently of the representativeness of the sample. We also opted to study diagnostic performance for CI rather than for dementia only, as we deem this approach to be more pragmatic from a clinical perspective. For healthcare professionals in general, and primary care physicians in particular, it is beneficial to focus more on diagnosing CI, as targeting dementia exclusively would prevent us from identifying patients with CI with no functional impact, who may be eligible for preventive, corrective, or palliative measures to delay or even prevent progression to dementia.[20]

The main limitation of this study is the fact that the BCTs were applied consecutively, which may have introduced bias related to tiredness or fatigue, despite the short application times of these instruments; to partially mitigate this effect, the scales were administered in 3 different orders and we confirmed that results for the different BCTs showed no association with the serial position in which they were administered (data not shown). Another limitation is that the 7MS, also recommended in the SNS guidelines, was not included in the evaluation. However, this test[48] is a complex instrument, both in terms of administration and scoring; despite its name, it takes longer than 10 minutes to administer; it cannot be administered to illiterate individuals; and it is not free of cost. Therefore, we do not consider this instrument to be appropriate for use at primary care centres.[20]

The main strength of our study is its design. On the one hand, the prospective, consecutive, systematic recruitment of participants; the use of a one-year recruitment period; the near absence of exclusion criteria; and the low drop-out rate mean that our sample accurately reflects the issue of diagnosing CI in clinical practice in this healthcare setting, conferring naturalistic and pragmatic value to the study. On the other, the fact that all participants were assessed with all BCTs, the fact that all participants were administered the gold-standard diagnostic protocol independently of their results in the BCTs, and the independent, blinded evaluation of the instruments studied and the gold-standard assessment ensure that we controlled for the main biases affecting studies evaluating diagnostic tests.[25,49]

In conclusion, there is no perfect, ideal BCT for use in primary care, and these professionals must be familiar with the use of various instruments; the selection of a specific test to use in a particular case should be based on the clinical circumstances (time available), patient characteristics (level of schooling, sensory deficits, etc), available data (normative data, specific validation for diagnosis of CI), and the healthcare professional's preferences and experience.[20] Based on our findings and the desirable characteristics summarised in Table 1, we recommend the Eurotest, T@M, and Fototest for use in primary care, with the latter being the most efficient as it takes less than half as long to administer.

Funding

Conflicts of interest

C. Carnero Pardo created the Fototest and Eurotest and has received professional fees for academic and consulting activities for Nutricia, Schwabe Farma Ibérica, Biogen, Piramal, Janssen Cilag, Pfizer, Eisai, Esteve, Novartis, Lundbeck, and Grunenthal.

Acknowledgements

References

1. Villarejo Galende A, Eimil Ortiz M, Llamas Velasco S, Llanero Luque M, Lopez de Silanes de Miguel C, Prieto Jurczynska C. Informe de la Fundacion del Cerebro. Impacto social de la enfermedad de Alzheimer y otras demencias. Soc Esp Neurol. 2017:1—42.

2. Gavrila D, Antunez C, Tormo MJ, Carles R, Garcia Santos JM, Parrilla G, et al. Prevalence of dementia and cognitive impairment in Southeastern Spain: the Ariadna study. Acta Neurol Scand. 2009;120:300—7.

3. Rodriguez-Sanchez E, Mora-Simon S, Patino-Alonso MC, Garcia-Garcia R, Escribano-Hernandez A, Garcia-Ortiz L, et al. Prevalence of cognitive impairment in individuals aged over 65 in an urban area: DERIVA study. BMC Neurol. 2011;11:147.

4. Vega Alonso T, Miralles Espi M, Mangas Reina JM, Castrillejo Perez D, Rivas Perez AI, Gil Costa M, et al. Prevalence of cognitive impairment in Spain: the Gomez de Caso study in health sentinel networks. Neurologia. 2018;33:491—8.

5. Lin JS, O'Connor E, Rossom RC, Perdue LA, Eckstrom E. Screening for cognitive impairment in older adults: a systematic review for the U.S. Preventive Services Task Force. Ann Intern Med. 2013;159:601—12.

6. United Kingdom National Screening Committee. The UK NSC recommendation on screening for dementia. London: Public Health England; 2015.

7. Gerontology Society of America. Workgroup on Cognitive Impairment Detection and Earlier Diagnosis. Report and Recomendations. Washington DC: The Gerontology Society of America; 2015.

8. Canadian Task Force on Preventive Health Care, Pottie K, Rahal R, Jaramillo A, Birtwhistle R, Thombs BD, et al. Recommendations on screening for cognitive impairment in older adults. CMAJ. 2016;188:37—46.

9. Chambers LW, Sivananthan S, Brayne C. Is dementia screening of apparently healthy individuals justified? Adv Prev Med. 2017;2017:9708413.

10. Grupo de trabajo de la Guía de Práctica Clínica sobre la atención integral a las personas con enfermedad de Alzheimer y otras demencias. Guía de Práctica Clínica sobre la atención integral a las personas con enfermedad de Alzheimer y otras demencias. Plan de Calidad para el Sistema Nacional de Salud del Ministerio de Sanidad, Política Social e Igualdad, Agència dÍnformació, Avaluació i Qualitat en Salut de Cataluña, 2011.

11. Blesa R, Pujol M, Aguilar M, Santacruz P, Bertran-Serra I, Hernandez G, et al. Clinical validity of the' mini-mental state' for Spanish speaking communities. Neuropsychologia. 2001;39:1150—7.

12. Lobo A, Saz P, Marcos G, Dia JL, de la Camara C, Ventura T, et al. Revalidacion y normalizacion del Mini-Examen Cognoscitivo (primera version en castellano del Mini-Mental Status Examination) en la poblacion general geriatrica. Med Clin (Barc). 1999;112:767—74.

13. Martinez de la Iglesia J, Duenas Herrero R, Onis Vilches MC, Aguado Taberne C, Albert Colomer C, Luque Luque R. Adaptación y validación al castellano del cuestionario de Pfeiffer (SPMSQ) para detectar la existencia de deterioro cognitivo en personas mayores de 65 años. Med Clin (Barc). 2001;117:129—34.

14. Perez-Martinez DA, Baztan JJ, Gonzalez-Becerra M, Socorro A. Evaluación de la utilidad diagnóstica de una adaptación española del Memory Impairment Screen de Buschke para detectar demencia y deterioro cognitivo. Rev Neurol. 2005;40:644—8.

15. del Ser Quijano T, Sanchez Sanchez F, Garcia de Yebenes MJ, Otero Puime A, Zunzunegui MV, Munoz DG. Versión española del Test de los 7 Minutos. Datos normativos de una muestra poblacional de ancianos de más de 70 anos. Neurologia. 2004;19:344—58.

16. Cacho J, Garcia-Garcia R, Arcaya J, Vicente JL, Lantada N. Una propuesta de aplicacion y puntuacion del test del reloj en la enfermedad de Alzheimer. Rev Neurol. 1999;28:648—55.

17. Carnero-Pardo C, Montoro-Rios MT. Evaluacion preliminar de un nuevo test de cribado de demencia (Eurotest). Rev Neurol. 2004;38:201—9.

18. Carnero Pardo C, Sáez-Zea C, Montiel Navarro L, del Saz P, Feria Vilar I, Pérez Navarro MJ, et al. Utilidad diagnóstica del Test de las Fotos (FOTOTEST) en deterioro cognitivo y demencia. Neurologia. 2007;22:860—9.

19. Rami L, Molinuevo JL, 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.

20. Olazaran 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.

21. Rami L, Bosch B, Valls-Pedret C, Caprile C, Sanchez-Valle Diaz R, Molinuevo JL. Validez discriminativa y asociacion del test mini-mental (MMSE) y del test de alteracion de memoria (T@M) con una bateria neuropsicologica en pacientes con deterioro cognitivo leve amnesico y enfermedad de Alzheimer. Rev Neurol. 2009;49:169—74.

22. Carnero-Pardo C, Espejo-Martinez B, Lopez-Alcalde S, Espinosa-Garcia M, Saez-Zea C, Vilchez-Carrillo R, et al. Effectiveness and costs of Phototest in dementia and cognitive impairment screening. BMC Neurol. 2011;11:92.

23. Carnero-Pardo C, Espejo-Martinez B, Lopez-Alcalde S, Espinosa-Garcia M, Saez-Zea C, Hernandez-Torres E, et al. Diagnostic accuracy, effectiveness and cost for cognitive impairment and dementia screening of three short cognitive tests applicable to illiterates. PLoS One. 2011;6:e27069.

24. Cacho J, Benito-Leon J, Garcia-Garcia R, Fernandez-Calvo B, Vicente-Villardon JL, Mitchell AJ. Does the combination of the MMSE and clock drawing test (mini-clock) improve the detection of mild Alzheimer's disease and mild cognitive impairment? J Alzheimers Dis. 2010;22:889—96.

25. Sackett DL, Haynes RB. The architecture of diagnostic research. BMJ. 2002;324:539—41.

26. Boada M, Cejudo JC, Tarraga L, Lopez OL, Kaufer D. Neuropsychiatric Inventory Questionnaire (NPI-Q): validacion española de una forma abreviada del Neuropsychiatric Inventory (NPI). Neurologia. 2002;17:317—23.

27. Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. Md State Med J. 1965;14:61—5.

28. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist. 1969;9:179—86.

29. Pfeffer RI, Kurosaki TT, Harrah CH Jr, Chance JM, Filos S. Measurement of functional activities in older adults in the community. J Gerontol. 1982;37:323—9.

30. Carnero Pardo C, Lendínez-González A. Utilidad del test de fluencia verbal semántica en el diagnóstico de demencia. Rev Neurol. 1999;29:709—14.

31. Morris JC, Mohs RC, Rogers H, Fillenbaum G, Heyman A. Consortium to establish a registry for Alzheimer's disease (CERAD) clinical and neuropsychological assessment of Alzheimer's disease. Psychopharmacol Bull. 1988;24:641—52.

32. Calero MD, Arnedo ML, Navarro E, Ruiz-Pedrosa M, Carnero C. Usefulness of a 15-item version of the Boston Naming Test in neuropsychological assessment of low-educational elders with dementia. J Gerontol B Psychol Sci Soc Sci. 2002;57:187—91.

33. Perez-Marmol JM, Lopez-Alcalde S, Carnero-Pardo C, Canadas-De la Fuente GA, Peralta-Ramirez MI, Garcia-Rios MC. Creation and design of a test for the Evaluation of Upper Limb Apraxia (EULA) based on a cognitive model: a pilot study. Rev Neurol. 2015;60:66—74.

34. Robles A, del Ser T, Alom J, Pena-Casanova J. Propuesta de criterios para el diagnostico clinico del deterioro cognitivo ligero, la demencia y la enfermedad de Alzheimer. Neurologia. 2002;17:17—32.

35. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988;44:837—45.

36. Statistical guidance on reporting results from studies evaluating diagnostic test: U. S. Department of Health and Human Services, Food and Drug Administration, Center for Devices and Radiological Health, Diagnostic Devices Branch, Division of Bioestatistic, Office of Surveillance and Biometrics, 2003.

37. MedCalc Statistical Software versión 18.9.1 [computer program]. Published: MedCalc Software bvba, Ostend, Belgium; http://www.medcal.org; 2018.

38. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. BMJ. 2003;326:41—4.

39. Ozer S, Noonan K, Burke M, Young J, Barber S, Forster A, et al. The validity of the Memory Alteration Test and the Test Your Memory test for community-based identification of amnestic mild cognitive impairment. Alzheimers Dement. 2016;12:987—95.

40. Pezzotti P, Scalmana S, Mastromattei A, Di Lallo D, Progetto Alzheimer Working Group. The accuracy of the MMSE in detecting cognitive impairment when administered by general practitioners: a prospective observational study. BMC Fam Pract. 2008;9:29.

41. Mitchell AJ. A meta-analysis of the accuracy of the mini-mental state examination in the detection of dementia and mild cognitive impairment. J Psychiatr Res. 2009;43:411—31.

42. Carnero Pardo C, Cruz Orduna I, Espejo Martinez B, Cardenas Viedma S, Torrero Garcia P, Olazaran Rodriguez J. Efectividad del Mini-Mental en la deteccion del deterioro cognitivo en Atencion Primaria. Aten Primaria. 2013;45:426—33.

43. Carnero-Pardo C. ¿Es hora de jubilar al Mini-Mental? Neurologia. 2014;29:473—81.

44. Weintraub S, Besser L, Dodge HH, Teylan M, Ferris S, Goldstein FC, et al. Version 3 of the Alzheimer Disease Centers' Neuropsychological Test Battery in the Uniform Data Set (UDS). Alzheimer Dis Assoc Disord. 2018;32:10—7.

45. Carnero-Pardo C, Rego García I, Barrios López J, Blanco Madera S, Calle Calle R, López-Alcalde S, et al. Evaluación de la utilidad diagnóstica y validez discriminativa del Test del Reloj y del Mini-Cog en la detección del deterioro cognitivo. Neurologia. 2019, http://dx.doi.org/10.1016/j.nrl.2018.12.002.

46. Burgueno MJ, Garcia-Bastos JL, Gonzalez-Buitrago JM. Las curvas ROC en la evaluacion de las pruebas diagnosticas. Med Clin (Barc). 1995;104:661—70.

47. Aggarwal R, Ranganathan P. Understanding diagnostic tests - Part 3: receiver operating characteristic curves. Perspect Clin Res. 2018;9:145—8.

48. Del Ser T, Sanchez-Sanchez F, Garcia de Yebenes MJ, Otero A, Munoz DG. Validation of the seven-minute screen neurocognitive battery for the diagnosis of dementia in a Spanish population-based sampleValidation of the seven-minute screen neurocognitive battery for the diagnosis of dementia in a Spanish population-based sample. Dement Geriatr Cogn Disord. 2006;22:454—64.

49. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol. 2003;3:25.