Neurological Sciences (2022) 43:821–845
https://doi.org/10.1007/s10072-021-05683-4

REVIEW ARTICLE

Psychometrics and diagnostics of Italian cognitive screening tests: a systematic review

Edoardo Nicolò Aiello[1] · Sara Rimoldi[2] · Nadia Bolognini[2,3] · Ildebrando Appollonio[4] · Giorgio Arcara[5]

Received: 17 August 2021 / Accepted: 17 October 2021 / Published online: 24 November 2021

Abstract

Background Cognitive screening tests (CSTs) are crucial to neuropsychological diagnostics, and thus need to be featured by robust psychometric and diagnostic properties. However, CSTs happen not to meet desirable statistical standards, negatively affecting their level of recommendations and applicability. This study aimed at (a) providing an up-to-date *compendium* of available CSTs in Italy, (b) report their psychometric and diagnostic properties, and (c) address related limitations.

Methods This review was implemented by consulting Preferred Reporting Items for Systematic Reviews and Meta-Analyses and pre-registered on the International Prospective Register of Systematic Reviews. Standardization and usability studies focusing on norms, validity, reliability, or sensitivity/specificity (and derived metrics) in adults were considered for eligibility. Quality assessment was performed by means of an ad hoc checklist collecting information on sampling, psychometrics/diagnostics, norming, and feasibility.

Results Sixty studies were included out of an initial $N = 683$. Identified CSTs ($N = 40$) were classified into general, domain-, and disease-specific ($N = 17$, 7, and 16, respectively), the latter being less statistically robust than remaining categories. Validity and reliability evidence was provided for 29 and 26 CSTs, respectively, sensitivity/specificity for 20 and norms for 33. Prevalence- and post-test-based diagnostic metrics were seldomly represented; factorial structures, ceiling/floor effects, and acceptability rarely investigated; content, face, and ecological validity never assessed.

Discussion Although available Italian CSTs overall met basic psychometric/diagnostic requirements, their statistical profile often proved to be poor on several properties that are desirable for clinical applications, with a few exceptions among general and domain-specific ones.

Keywords Cognitive screening · Psychometrics · Diagnostics · Standardization · Neuropsychology · Normative data

Introduction

Cognitive screening in adults and elders is relevant to both neurological/neuropsychiatric diagnostics and prevention in internal patients with possible brain damage [1], as well as, in turn, to prognosis and interventional management [2]. To screen for cognitive deficits is indeed meant to ease clinicians into determining whether II-level neuropsychological assessment (i.e., an in-depth examination of multiple cognitive/behavioral functions) is needed for a given patients [3].

As aimed at providing practitioners with an optimal compromise between informativity and ease of use within the

✉ Edoardo Nicolò Aiello
e.aiello5@campus.unimib.it

Sara Rimoldi
s.rimoldi11@campus.unimib.it

Nadia Bolognini
nadia.bolognini@unimib.it

Ildebrando Appollonio
ildebrando.appollonio@unimib.it

Giorgio Arcara
giorgio.arcara@ospedalesancamillo.net

[1] PhD Program in Neuroscience, School of Medicine and Surgery, University of Milano-Bicocca, Monza, Italy

[2] Department of Psychology, University of Milano-Bicocca, Milano, Italy

[3] Neuropsychological Laboratory, IRCCS Istituto Auxologico Italiano, Milano, Italy

[4] Neurology Section, School of Medicine and Surgery, University of Milano-Bicocca, Monza, Italy

[5] IRCCS San Camillo Hospital, Venice, Italy

early detection of changes in cognition [1], cognitive screening tests (CSTs) need to come with robust psychometric and diagnostic properties, representative norms, and evidence of clinical feasibility in target conditions (i.e., clinical populations that they are meant to be administered to) [2, 4] (see Table 1).

However, it has been already acknowledged that widespread CSTs happen to fail reaching the aforementioned statistical standards, in turn negatively affecting their level of recommendations [9]. In this respect, cross-cultural adaptations of CSTs have been specifically highlighted as suffering from psychometric/diagnostic weaknesses [10], and this representing a major issue in the light of the relevance of culture-/language-specificity to cognitive assessment [11].

In Italy, much attention has been historically devoted to providing norms within the development and adaptation of CSTs [12]. However, it is debated whether this focus might have led to neglecting other fundamental statistical aspects when standardizing tests, such as validity, reliability, and diagnostic properties [13].

In light of the above premises, this study aimed at systematically reviewing evidence on originally Italian/adapted-to-Italian CSTs in order to (a) provide an up-to-date *compendium* of available CSTs in Italy; (b) report their psychometric and diagnostic properties; and (c) address current issues with regard to their development, adaptation, and standardization.

Methods

Search strategy

Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines were consulted [14]. This review was pre-registered on the International Prospective Register of Systematic Reviews (PROSPERO; CRD42021254561: https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=254561).

The following search terms were entered into Scopus and PubMed databases on May 1, 2021 (no date limit set): neuropsych* OR cogniti* AND screen* OR "screening test" OR "screening tool" OR "screening instrument" AND Italy OR Italian. Fields of search were title, abstract, and key words for Scopus whereas title and abstract for PubMed. Only peer-reviewed, full-text contributions written in English/Italian were considered. Hence, non-peer-reviewed literature was not searched for. Further contributions of possible interest were identified within reference lists of included articles/through manual search.

Contributions focusing either on the standardization of Italian/adapted-to-Italian CSTs (i.e., investigations on psychometric/diagnostic or normative studies) or their feasibility/usability in healthy participants (HPs) and in patients with neurological or neuropsychiatric diseases were considered for eligibility. For a non-normative study to be included, at least one property among validity, reliability, and sensitivity/specificity (or related metrics) had to be assessed. Case reports/case series, reviews/meta-analyses, abstracts, research protocols, qualitative studies, and opinion papers were excluded. Among feasibility/usability studies, those focusing on selected clinical populations that would have not allowed sufficient generalizability were not considered. Investigations on proxy-report tools, questionnaires, CSTs for pediatric populations, or requiring ≥45′ to be administered were also excluded in order to improve external validity of conclusions.

Data collection and quality assessment

Screening and eligibility stages were performed by one of the authors (EN.A.) via Rayyan (https://rayyan.qcri.org/welcome); a second author (G.A.) supervised this stage.

Data extraction was performed by two independent collaborators (S.R. and F.C.), whereas one independent author (E.N.A.) supervised this stage and checked extracted data.

Outcomes of interest were (1) sample size, (2) sample representativeness (geographic coverage, exclusion criteria), (3) participants' demographics, (4) test adaptation procedures, (5) modality of assessment (in-person vs. remote), (6) administration time, (7) validity metrics, (8) reliability metrics (including significant change measures), (9) measures of sensitivity and specificity, (l0) metrics derived from sensitivity and specificity, (11) norming methods, and (12) other psychometric/diagnostic properties (e.g., accuracy, acceptability rate, assessment of ceiling/floor effects).

Formal quality assessment was performed for each CST according to the aforementioned categories by developing an ad hoc checklist (Cognitive Screening Standardization Checklist, CSSC) (see Table 2). The CSSC encompasses two sections, "Sampling" (ranging 0–13) and "Psychometrics, diagnostics, and usability" (ranging 0–29). The first section evaluates the sampling adequacy as for representativeness; the second section focuses on psychometric, diagnostic properties, and feasibility. CSSC total scores range from 0 to 42; a given CST was thus judged as "statistically sound" if scoring ≥21 (i.e., 50% of the maximum) on the CSSC. CSSC items were based on [1, 2] and [7].

Scores were "cumulatively" assigned for each CST by evaluating all available studies on it among those included. Items targeting non-cumulative information which were nonetheless retrievable in multiple studies — e.g., the normative sample size — were scored according to that study providing the highest-quality information — e.g., the highest *N*.

Table 1 Desirable psychometric and diagnostic properties of a cognitive screening test (CST)

Property		Common associated method(s)
Validity-the extent to which the CST measures what it is intended to -		
Content validity	The goodness of the operationalization of the construct that the CST is intended to measure	Qualitative
Construct validity	The degree to which the CST is related to constructs that it is/it is not supposed to be related to (convergent/divergent validity)	Correlation
Criterion validity	The degree to which the CST is predictive of a related construct measured either at the same time of its administration (concurrent validity) or at a different time (predictive validity)	Regression
Ecological validity	The extent to which (a) the CST scores are predictive of the subject's everyday level of functioning and (b) the CST simulates real-life tasks	Correlation/regression
Reliability -the extent to which the scores yielded from the administration of the CST in different conditions to the same subjects are consistent -		
Internal consistency	The degree of relatedness among the scores yielded by each item of the CST	Cronbach's α
Test–retest reliability	The degree of consistency of CST scores across different administration to the same subjects over time	Intra-class correlation/correlation
Inter-rater reliability	The degree of consistency between the scores yielded by the administration of the CST to the same subjects by two independent raters	Intra-class correlation/Cohen's *k*/correlation
Parallel-form reliability	The degree of consistency among the scores yielded by different versions of the same CST administered to the same subjects	Correlation
Significant change	The capability of a CST to detect change over time by taking into account reliability and (possibly) practice effect	Reliable Change Index[†]
Diagnostic properties		
Sensitivity	The extent to which the CST is able to classify as impaired a subject that is actually impaired	TP/(TP + FN)
Specificity	The extent to which the CST is able to classify as unimpaired a subject that is actually unimpaired	TN/(TN + FP)
Positive predictive value	The probability that a subject classified as impaired by the CST is actually impaired	TP/(TP + FP)
Negative predictive value	The probability that a subject classified as unimpaired by the CST is actually unimpaired	TN/(TN + FN)
Positive likelihood ratio	An estimate of how much the probability that a subject is impaired increases when she/he is classified by the CST as being impaired	SE/(1 − SP)
Negative likelihood ratio	An estimate of how much the probability that a subject is unimpaired decreases when she/he is classified by the CST as being unimpaired	(1 − SE)/SP
Norms - reference values derived from a representative healthy population sample -		
Background adjustment	Raw scores on a CST were adjusted for relevant anagraphic-demographic intervening predictors by means of regression-based approaches	Regression-based methods
Inferential error control	Cut-offs were derived by taking into account measures to reduce the risk of classifying a patients as impaired vs. unimpaired	Tolerance limits; [5] method

Table 1 (continued)

Property		Common associated method(s)
Feasibility/usability -the extent to which the CST is usable by practitioners -		
Back-translation	A procedure according to which the original protocol of the CST is translated from the original language to the target one by one independent translator, whereas a second, independent translator back-translates this last protocol to the original language	Qualitative
Cultural adaptation	If needed, items that are likely to be culture-specific should be adapted to the target population	Qualitative
Acceptability	The rate of subjects who do not refuse to take the CST out of a sample	Proportion out of *N*
Ceiling/floor effects	The rate of subjects scoring the minimum/maximum out of the CST	Proportion out of *N*
Sensitivity to change	The capability of the CST to detect trends in cognition over time	Repeated-measure analyses
Sensitivity to severity	The capability of the CST to discriminate between patients with different levels of cognitive impairment	Between-group, cross-sectional analyses
Case–control discrimination	The capability of the CST to discriminate between (a) patients and healthy individual or (b) a given group of patients from a group of different patients	Between-group, cross-sectional analyses

SE sensitivity, *SP* specificity, *TP* true positives, *TN* true negative, *FP* false positive, *FN* false negatives

†See [6] for an alternative method to RCI, used for inferring significant change, and used in an Italian test. Based on [1–3, 7, 8, 75]

Quality assessment was performed by one of the authors (S.R.) and supervised by a second, independent one (E.N.A.).

Results

Study selection process is shown in Fig. 1.

Sixty-one studies were ultimately included. Extracted outcomes are reported in Table 3. A summarization of most relevant psychometric and diagnostic properties for each included CST, along with CSSC scores, are reported in Table 4.

The vast majority of contributions were studies mostly aimed at providing normative data ($N = 32$) — of which, 11 did not report any further relevant statistical property. Twenty-seven studies instead focused on psychometric/diagnostic properties with only marginal/absent attention to normative values (either in the context of clinical usability or not).

Included CSTs fell under the following categories: (a) domain-/disease-nonspecific (in-person: $N = 14$; remote: $N = 3$); (b) domain-specific ($N = 7$), targeting executive functioning, language, memory, and praxis; and (c) disease-specific ($N = 16$), targeting neurodegenerative disorders (Alzheimer's, Parkinson's, and motor neuron diseases), cerebrovascular accidents, neuropsychiatric conditions, infective sub-cortical dementias, delirium, migraine-related subjective cognitive dysfunction, and dementia in the context of intellectual disabilities. Among all the investigations, target clinical populations were present in 33 studies.

Validity was investigated in 37 studies and mostly by convergence ($N = 31$); divergent validity was assessed in 5 studies, whereas criterion validity in 4 (3 of which via concurrent validity, whereas one via predictive validity). Only 5 of the included studies assessed the factorial structure underlying CSTs by means of dimensionality reduction approaches. No overt evidence of content, face, and ecological validity was detected.

Reliability was investigated in 32 studies and mostly as inter-rater ($N = 17$), internal consistency ($N = 14$), and test-retest ($N = 12$). Parallel forms were developed within 4 studies only.

Although sensitivity and specificity measures were often reported ($N = 22$), derived metrics (e.g., positive and negative predictive values and likelihood ratios) were provided in 10 studies only.

With respect to norming, regression-based and inferential-error-controlling methods — e.g., tolerance limits and/or Equivalent Scores [12], were highly represented ($N = 26$). Several studies ($N = 17$) derived point-estimate cut-offs through receiver-operating characteristic (ROC) analyses.

Table 2 Cognitive Screening Standardization Checklist (CSSC)

Sampling (CSSC-S)	
Are sample demographics reported?	
age, education and sex reported	1
≥ 1 feature(s) missing (age, education, or sex)	0
Are sample demographics adequately described?	
Continuous ones (e.g., age and education) reported as $M \pm SD$ and *range*; categorical ones as frequencies/percentages (e.g., sex)	1
≥ 1 statistics not reported	0
Is/are the normative sample(s) size adequate?	
$N \geq 300$	2
$150 \leq N < 300$	1
$N < 150$	0
Is/are the normative sample(s) geographically representative?	
At least 3 different regions with acceptable geographic coverage (e.g., North, Center, and South)	2
At least 2 different regions with acceptable geographic coverage (e.g., North and South)	1
One region only	0
Does/do the normative sample(s) cover a sufficiently wide range of adult age?	
$Range \geq 50$	2
$30 \leq range < 50$	1
$Range < 30$	0
Is/are the normative sample(s) representative of all levels of education?	
$Range \geq 13$	2
$10 \leq range < 13$	1
$Range < 10$	0
Is/are the normative sample(s) stratification adequate (for age, education, and sex)?	
All co-occurrence cells including at least 1 observation	2
The majority of cells including at least 1 observation, except for "critical" ones (e.g., young age with low education)	1
Clearly unbalanced patterns detectable within table stratification	0
Is/are the normative sample(s) well balanced between males and females (40 and 60% or 50–50% circa, respectively)?	1
Are exclusion criteria for normative sample(s) adequately described?	
Quantitative (e.g., cut-off scores on a test) and qualitative (premorbid conditions sufficiently described, e.g., subdivision into medical classifications)	2
Either one of each (quantitative or well-described qualitative)	1
Only qualitative and not sufficiently described	0
CSSC-S score:	__/13
Psychometrics, diagnostics, and usability (CSSC-PDU)	
Is linguistic adaptation adequate?	
Back-translation/not necessary (de novo instrument)	2
Simple translation with adequate controls for subjectivity biases (e.g., independent judges)	1
Simple translation only	0
Is cross-cultural adaptation adequate (if necessary)?	
Critical items adequately addressed/not necessary	1
Critical items not adequately addressed	0
Is there acceptability evidence? (1 if "yes")	1
Has validity been tested? (1 if "yes")	1
If "yes" which of the following validity measures have been considered? (1 if "yes")	
Convergent validity	1
Divergent validity	1
Criterion validity	1
Ecological validity	1
Dimensionality-reduction techniques	1
Other validity measures	1

Table 2 (continued)

If convergent validity has been tested: is/are the correlational measure(s) appropriate (targeting the same construct)? (1 if "yes")	1
Has reliability been tested? (1 if "yes")	1
If "yes," which of the following reliability measures has been considered? (1 if "yes")	
Inter-rater reliability	1
Test–retest reliability	1
Internal consistency	1
Parallel forms/significant change measures	1
Has sensitivity been tested? (1 if "yes")	1
Has specificity been tested? (1 if "yes")	1
Are sensitivity- and specificity-derived metrics present? (e.g., PPV) (1 if "yes")	1
Is accuracy (e.g., AUC) reported? (1 if "yes")	1
If ROC analyses have been carried out (1 if "yes"):	
has the cut-off value been identified with an explicitly described procedure (e.g., Youden index)?	1
has the target condition been adequately identified?	
Quantitative (e.g., cut-off scores on a test)	1
Qualitative (e.g., clinical criteria)	1
Has a normative study been carried out? (1 if "yes")	1
If "yes":	
Has a regression-based approach been adopted?	1
Has a control for inferential error in deriving normative values been adopted?	1
Have ceiling/floor effects been investigated?	1
Does the adopted approach comes with distributional assumptions (e.g., Gaussian distribution)? (0 if "yes")	1
CSSC-PDU score:	_/29
CCSS total score:	_/42

AUC area under the curve, *ROC* receiver-operating characteristics, *PPV* positive predictive value

Unreported information were scored as 0

Acceptability of the CST was overtly examined in 9 studies, while ceiling/floor effects in 11. When applicable, administration time ranged from 2 to 45 min.

Discussion

The present work investigates statistical features of CSTs currently available in Italy, shedding a new light on their clinical and experimental utilization. Information here reported have the potential to promote a more aware and critical usage of CSTs among Italian clinicians, as well as to serve as overall guidelines for researchers either involved in CST development/adaptation/standardization or devoted to addressing open issues on CST psychometrics/diagnostics.

Overall, although psychometrics and diagnostics for a given CST happened not to be assessed within the same study, basic properties and norms were provided within different ones, especially for most widespread CSTs (e.g., Mini-Mental State Examination, MMSE; Montreal Cognitive Assessment, MoCA).

Moreover, although results show a general trend towards focusing on providing only normative data and cut-off values, the majority of included CSTs proved to be supported by sufficient evidence as for basic psychometric/diagnostic requirements are concerned. The present review hints at a relatively high quality of selected global and domain-specific CSTs — e.g., MoCA (CSSC = 34) and Addenbrooke's Cognitive Examination — Revised (ACE-R; CSSC = 31), Screening for Aphasia in Neurodegeneration (SAND; CSSC = 27), and Frontal Assessment Battery (FAB; CSSC = 24). Disease-specific CSTs were shown to be less statistically robust, with a few exceptions — e.g., ALS Cognitive Behavioral Screen (ALS-CBS; CSSC = 26).

Validity

Findings on validity happened to show misinterpretations of the psychometric concepts and incomplete analyses for certain CSTs.

First, it is worth mentioning that convergent and concurrent validity happened to be mistaken for each other — e.g., concurrent validity being tested by means of correlations instead of regressions, or convergent and concurrent validity being addressed as the same construct [54, 55, 61, 69, 73].

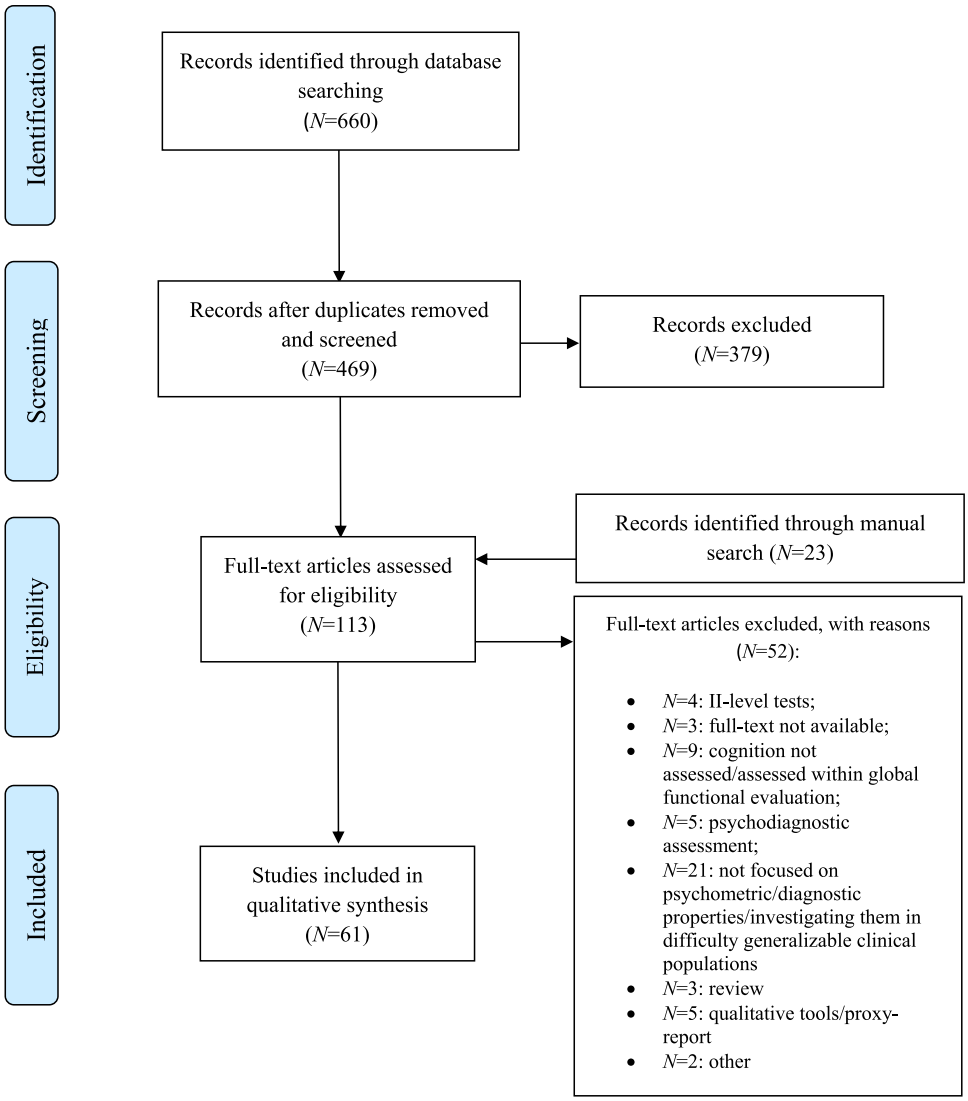

Fig. 1 PRISMA flow chart displaying study selection process. *PRISMA* Preferred Reporting Item for Systematic Reviews and Meta-Analyses, *NPs* neuropsychological. Diagram adapted from [14](www.prisma-statement.org)

In this regard, one should also note that correlational measures happened not to be meant to assess the same construct as that of the target CSTs — e.g., FAB validity being tested against the MMSE [24].

Moreover, predictive validity happened to be almost never assessed [29] — despite the longitudinal dimension being relevant to the monitoring of patients' cognitive profile. This may be due to the high cost of performing a proper longitudinal study to assess predictive validity.

It is also worth mentioning that the vast majority of included CSTs lacked divergent validity evidence. This might be due to the fact that different CSTs are commonly found to correlate despite being meant to assess different functions; this because target constructs often overlap to some extent. Researchers are thus encouraged to test divergence by addressing measures that are supposed to deviate from a given CST as far as either construct or face validity is concerned. This could be done by comparing a CST with either a II-level, domain-specific cognitive test, or with a psychodiagnostic tool.

Furthermore, although the need for cognitive measures that are predictive of daily functioning has been highlighted [12], it has to be noted that the ecological validity of CSTs has never been found to be directly investigated within original standardization studies. This may be due to the lack of a wide consensus on how to investigate ecological validity, as well as to the scarce availability of ad hoc scales designed for assessing the specific impact of cognitive disorders on real-life functioning, going beyond a general evaluation of functional disability.

Finally, researchers should take into consideration to explore content validity and factorial structures of CSTs; this equally applying both to those tests postulated to be mono-factorial (e.g., MMSE) and to domain-specific ones (e.g., SAND), which might nonetheless cover multiple cognitive functions.

Table 3 Summary of extracted outcomes

Authors	Test	*N*	Sampling	Demographics	Validity	Reliability	Sensitivity and specificity	SE- and SP-derived measures	Other properties	Norming method
(1) Mazzoni et al. [15]	MMSE In-person	88 AD;, 57 VaD	Pisa (AD, VaD); EC: neurological	/	/	/	SE = 61–70%	/	/	/
(2) Rocca et al. [16]	HAMTS In-person -	20 dem.; 104 HPs	Macerata; EC: neurological, peripheral, psychiatric	Age: (60–89) M/F: 44/56%	/	/	SE = 90% SP = 89%	/	/	ROC (diagnostic criteria); adjustment
(3) Measso et al. [17]	MMSE In-person -	906 HPs	EC: internal, neurological, psychiatric	Age: (20–79) Education: 8.4 ± 0.1 M/F: 48.7/51.3%	/	/	/	/	Acceptability	TL; adjustment
(4) Brazzelli et al. [71]	MODA In-person 30–45′	312 suspected CI; 217 HPs	Milano; Veruno EC (HPs): neurological, psychiatric	Suspected CI: Age: 64.8 ± 13 Education: 7–3 ± 4–2 M/F: 44.6/55.4% HPs: Age: 60.8 ± 18.5 (20–97) Education: 9.1 ± 4.9 (1–17) M/F: 47.5/52.5%	CV (MMSE): HPs: $r = 0.61$ pts.: $r = 0.84$	TRR (21 d.): $r = 0.83$	/	/	/	TL; adjustment
(5) Fioravanti et al. [18]	ADAS-Cog In-person CT	95 HPs	EC: internal, neurological, psychiatric	Age: (50–79)	FS	/	/	/	/	Adjustment; *SD*-derived norms
(6) Magni et al. [19]	MMSE In-person -	1019 HPs	Ospitaletto; Coccaglio; Tirano; EC: MMSE < 21	Age: 75.4 ± 5.4 (65–89) Education: 5.2 ± 2.5 M/F: 34.3/75.5%	/	/	/	/	Acceptability	TL; adjustment
(7) Ferrucci et al. [20]	TICS In-person CT	41 dem.; 63 HPs	EC: age < 60; education < 3; peripheral, neurological, drugs	Age: 75.3 ± 7.8 (60–103) Education: 5.6 ± 3 (3–17) M/F: 34.6/65.4%	CrV (MMSE) $R^2 = 0.96$	TRR (7 d.) ICC = 0.91	SE > 90% SP > 90%	/	/	/
(8) Cossa et al. [21]	MMSE; MODA In-person 15–45′	829 HPs	Borgomanero EC: age < 60	Age: 72.6 ± 7.5 (60–100) Education: 5.6 ± 2.8 (0–19) M/F: 45/55%	/	IRR (MMSE, MODA): Cohen's $k = 0.4$	MODA: SE = 100% SP = 71.6% MMSE: SE = 85.7% SP = 90%	MODA: PPV = 14.4–11.5% NPV = 100% MMSE: PPV = 29.1–20.8% NPV = 99.2–99.2%	/	/
(9) Grigoletto et al. [4]	MMSE In-person ST ("carne")	908 HPs 912 dem	Arona; Desenzano; Poggio Renatico; Carmignano; Lamezia Terme; Motta Santa Anastasia; San Marino EC (HPs): neurological, peripheral, drugs	HPs: Age: (20–79) Education: (0–10; ≥ 10) M/F: 49.1/50.9% dem Age: (65–79)	/	/	SE = 85% SP = 89%	/	/	5th percentile for sex and education as step functions of age

Table 3 (continued)

Authors	Test	*N*	Sampling	Demographics	Validity	Reliability	Sensitivity and specificity	SE- and SP-derived measures	Other properties	Norming method
(10) Metitieri et al. [22]	Itel-MMSE (0–22) Telephone ST	104 dem. (AD, VaD, FTD)	Brescia; EC: peripheral, neurological	Age: 77.2 ± 8.1 Education: 5.2 ± 2.3 M/F: 24/76%	CV (MMSE): $r = 0.85$ CrV (MMSE) $R^2 = 0.72$	IRR: $r = 0.82$–0.9 TRR: $r = 0.9$–0.95	/	/	/	/
(11) Iavarone et al. [23]	FAB In-person -	236 HPs; 15 AD; 13 FTD	EC (HPs): neurological, psychiatric; CDR = 0	HPs: Age: (20–80; ≥ 80) Education: (3–13; ≥ 13) M/F: 47/53% AD: Age: 73.5 ± 7.2 Education: 7.1 ± 4.9 M/F: 20/80% FTD: Age: 63.2 ± 10.6 Education: 9.7 ± 4 M/F: 23/77%	HPs: CV (AM) $r = 0.29$; CV (TMT-B) $r = -0.62$ CV (DSS) $r = 0.65$ HPs and pts.: DV (MMSE): $r =$ n.s FS	HPs: IRR: Cohen's $k = 0.79$ ($N = 26$ HPs) TRR: Cohen's $k = 0.81$ ($N = 31$) IC: Cronbach's $\alpha = 0.78$	/	/	/	TL; adjustment
(12) Appollonio et al. [24]	FAB In-person CT 10′	364 HPs	Milan; Modena EC: MMSE AS < 24; internal, neurological, psychiatric	Age: 57.4 ± 17.9 (20–94) Education: 10.4 ± 4.3 (1–17) M/F: 40.9/59%	CV (MMSE): $r = 0.41$	IRR: $r = 0.96$ ($N = 56$) TRR (14–28 d.): $r = 0.85$ ($N = 45$)	/	/	/	TL + ES
(13) Dal Forno et al. [25]	I-TICS Telephone ST 10′	45 AD; 64 HPs	EC (AD): neurological EC (HPs): neurological	AD: Age: 73.9 ± 8.8 Education: 7.9 ± 3.9 M/F: 38/62% HPs: Age: 74.4 ± 8.1 Education: 7.5 ± 4.2 M/F: 36/64%	CV (MMSE): $r = 0.9$ CrV (MMSE): $R^2 = 0.82$	IC: Cronbach's $\alpha = 0.91$ IRR: Cohen's k (MMSE) = 0.72 TRR: ICC = 0.73	SE = 84% SP = 86%	/	/	ROC (diagnostic criteria)
(14) Michieletto et al. [26]	Mini-Cog In-person 3–5′ -	2186 HPs	EC: age < 65	/	/	IRR: Cohen's $k = 0.8$–0.9	/	/	Acceptability	/
(15) Vanacore et al. [27]	Itel-MMSE (0–22) Telephone	107 HPs	Rome; EC: neurological	Age: 64 ± 1.6 Education: 10.6 ± 4.3 M/F: 31.8/68.2%	CV (MMSE): $r = 0.26$	IC: Cronbach's $\alpha = 0.37$	SE = 23–75% SP = 61–76%	/	/	ROC (ES = 1/2)
(16) Anselmetti et al. [28]	BACS In-person BT	204 HPs	EC: internal, neurological, psychiatric	Age: 32.99 ± 12.7 (18–69) Education: 13.35 ± 3.7 (5–22) M/F: 42.6/57.4%	/	/	/	/	/	TL + ES

Table 3 (continued)

Authors	Test	*N*	Sampling	Demographics	Validity	Reliability	Sensitivity and specificity	SE- and SP-derived measures	Other properties	Norming method
(17) Pirani et al. [74]	GPCOG-It-Cognitive section In-person ST <5′	68 HPs; 132 dem	EC: age < 55; internal, peripheral, neurological, psychiatric	HPs: Age: 72.6 ± 9.03 Education: 7.1 ± 3.1 M/F: 41/59% Dementia: Age: 77.4 ± 6.7 Education: 5.8 ± 3.1 M/F: 42/58%	CV (MMSE) r = 0.68 CV (CAMCOG) r = 0.84 CV (ADAS-Cog) r = −0.7	/	SE = 98% SP = 58%	NPV = 0.94 PPV = 0.81	AUC = 0.96 MR = 16.5% IOCC = 0.74	ROC (diagnostic criteria)
(18) Lunardelli et al. [29]	BNS In-person ST 5–10′	247 HPs 134 stroke	Trieste; EC (HPs): neurological, psychiatric	HPs: Age: 48.7 ± 20.5 (15–85) Education: 11.5 ± 4.5 M/F: 45.7/54.3% Stroke: Age: 69.7 ± 12.9 (15–94) Education: 8.3 ± 3.4 M/F: 53/47%	18 left stroke: CV (ENPA comprehension, naming, reading) r = 0.85; 0.75; 0.84 CrV (NPs; 1–2 mo.): r: 0.23–0.52	/	/	/	Acceptability Ceiling effect	5th percentile
(19) Caffarra et al. [30]	CDT (Freedman) In-person	248 HPs	EC: internal, neurological, psychiatric	Age: (20–89) M/F: 50/50%	/	/	/	/	/	TL + ES
(20) Pigliautile et al. [73]	ACE-R In-person BT 15′	72 HPs; 46 AD; 18 FTD; 21 LBD	Perugia; Roma; EC (HPs): neurological, psychiatric; MMSE < 26	HPs (< 75 y.): Age: 69.6 ± 2.8 Education: 8.9 ± 4.6 M/F: 43/67% HPs (≥ 75 y.): Age: 80.7 ± 3.6 Education: 80.7 ± 3.6 M/F: 35/65% pts. (< 75 y.): Age: 70.8 ± 3.6 Education: 7.1 ± 3.7 M/F: 35/65% pts. (≥ 75 y.): Age: 80.9 ± 3.6 Education: 7.1 ± 4.8 M/F: 37/63%	CV (MMSE) r = 0.9	IC Cronbach's α = 0.85	SE = 82–90% SP = 82–100%	PPV = 56.8–100% NPV = 81.8–100%	AUC = 0.93	ROC (diagnostic criteria)

Table 3 (continued)

Authors	Test	*N*	Sampling	Demographics	Validity	Reliability	Sensitivity and specificity	SE- and SP-derived measures	Other properties	Norming method
(21) Girtler et al. [31]	SCEB In-person - < 10′	48 HPs; 29 AD; 27 MCI; 27 MDD	EC (HPs): neurological, psychiatric, internal, drugs EC (pts.): neurological, peripheral, psychiatric, internal	HPs: Age: 75.1 ± 6 (65–92) Education: 11.4 ± 3.6 (4–18) M/F: 35/65% AD: Age: 79.1 ± 6.1 (61–88) Education: 7.4 ± 4.3 (3–18) M/F: 24/76% MCI: Age: 76.6 ± 6.5 (66–94) Education: 7.3 ± 4.3 (1–17) M/F: 37/63% MDD: Age: 72 ± 5.5 (61–84) Education: 7.7 ± 2.8 (5–13) M/F: 33/67%	CV (MMSE) $r = 0.73$ AD CV (MMSE) $r = 0.59$	/	SE = 93.1–50.1% SP = 92.6–64.3%	PPV = 91.6–40.4% NPV = 95.7–82.4%	AUC = 0.96–0.52	ROC (clinical diagnosis)
(22) Costa et al. [32]	MMP In-person ST 10–15′	370 HPs	EC: internal, neurological, psychiatric; MMSE < 23.8	Age: 62 ± 13.9 (40–91) Education: 10.6 ± 4.8 (1–23) M/F: 51.8/48.2%	/	/	/	/	/	TL + ES
(23) Isella et al. [33]	SCOPA-Cog In-person BT 20′	121 PD	Monza; Milan; Viareggio; Lucca EC: internal, neurological, psychiatric; GDS > 10	Age: 69 ± 7.8 (50–86) Education: 7.9 ± 4 (3–17) M/F: 55.4/44.6%	CV (DRS; MMP): $r_z = 0.77$ DV: coefficient of variation: 0.34	IC: Cronbach's $\alpha = 0.78$	/	/	Acceptability Ceiling and floor effects	/
(24) Isella et al. [34]	SCOPA-Cog In-person	33 PD-MCI 31 PDD 49 PD	Monza; Milan; Lucca; Viareggio EC: internal, neurological, psychiatric; GDS > 10	PD Age: 64.7 ± 6.4 Education: 9.4 ± 4.1 M/F: 55.1/44.9% PD-MCI Age: 71.5 ± 6.0 Education: 7.4 ± 4.2 M/F: 54.4/45.6% PDD Age: 74.0 ± 7.7 Education: 6.6 ± 3.4 M/F: 54.8/45.2%	/	/	SE = 90–93% SP = 73–97%	PPV = 66–97% NPV = 95–94%	AUC = 0.92–0.99	ROC (clinical diagnosis)

Table 3 (continued)

Authors	Test	*N*	Sampling	Demographics	Validity	Reliability	Sensitivity and specificity	SE- and SP-derived measures	Other properties	Norming method
(25) Timpano et al. [35]	VMMSE (0–28) Video-conference ST 10′	207 suspected CI; 135 HPs	Messina; EC (suspected CI): peripheral, neurological; age < 50; EC (HPs): MMSE < 26; age < 50	Suspected CI: Age: 76.5 ± 8.0 Education: 6.3 ± 3.7 M/F: 34.3/65.7% HPs: Age: 65.7 ± 10.2 Education: 8.7 ± 4.1 M/F: 46.7/53.3%	/	IRR: $r = 0.94$ TRR: ICC = 0.85–0.94	SE = 87% SP = 97%	NPV = 0.83 PPV = 0.97	AUC = 0.96	ROC (MMSE)
(26) Bellelli et al. [36]	4AT In-person ST 2′	207 elderly pts 29 delirium	Monza; Cremona; EC: neurological, peripheral, disease-related (pts.)	HPs: Age: 83.6 ± 5.9 M/F: 37/63% Delirium: Age: 85.5 ± 7.3 M/F: 34/66%	/	IC Cronbach's α = 0.8	SE = 89.7% SP = 84.1%	LR + = 5.6 LR − = 0.1	AUC = 0.93	ROC (diagnostic criteria)
(27) Petrazzuoli et al. [72]	AQT In-person ST 15′	121 HPs	EC: internal, neurological, psychiatric; MMSE < 26	Age: 65.6. ± 12.1 Education: 12.2 ± 3.7 M/F: 57/43%	/	/	/	/	Ceiling effect	*SD* on log-transformed data
(28) Pignatti et al. [37]	PANDA In-person CT	111 PD (29 PCD + ; 82 PCD −); 103 HPs	EC (PD): peripheral, internal, neurological, psychiatric; EC (HPs): internal, neurological, psychiatric; MMSE < 27	PCD + : Age: 73 ± 6 Education: 9 ± 4 M/F: 65.5/34.5% PCD − : Age: 68 ± 9 Education: 9 ± 4 M/F: 57.3/42.7% HPs: Age: 68 ± 7 Education: 11 ± 3 M/F: 33/67%	/	/	SE = 96.6% SP = 82.2%	/	/	ROC (clinical diagnosis)
(29) Santangelo et al. [38]	PD-CRS In-person BT 10′	378 PD	EC: neurological, psychiatric	Age: 65.2 ± 10 M/F: 56.6/43.4%	CV (item 1.1 of MDS-UPDRS $r = -0.306$	IC Cronbach's α = 0.89	/	/	Acceptability	/
(30) Conti et al. [39]	MoCA In-person 15′	225 HPs	Bologna; EC: neurological, peripheral, psychiatric, drugs; MMSE < 23.8; PMT < 6.25	Age: 70.1 ± 5.7 (60–80) Education: 9.9 ± 4.6 (5–23) M/F: 49/51%	CV (MMSE) $r = 0.49$ CV (AS; MMSE) $r = 0.32$	/	/	/	/	TL + ES

Table 3 (continued)

Authors	Test	*N*	Sampling	Demographics	Validity	Reliability	Sensitivity and specificity	SE- and SP-derived measures	Other properties	Norming method
(31) Pigliautile et al. [40]	ACE-R In-person <20′	264 HPs	Perugia; Taranto EC: neurological, psychiatric; MMSE ≥ 24	Age: 72.9 ± 8 (60–93) Education: 9.7 ± 4.8 (1–19) M/F: 44/56%	/	/	/	/	/	TL + ES
(32) Pirrotta et al. [41]	MoCA In-person	154 suspected cognitive impairment 133 HPs	EC: age > 40; HPs MMSE < 26	Suspected cognitive impairment: Age: 76.8 ± 7.7 Education: 6 ± 3.9 M/F: 42.9/57.1%	/	TRR: ICC = 0.87 IRR: ICC = 0.96–0.98	SE = 0.83 SP = 0.97	NPV = 0.83 PPV = 0.98	AUC: 0.96	ROC (MMSE)
(33) Santangelo et al. [42]	MoCA In-person	415 HPs	Naples; Milan; Siena EC: internal, neurological, psychiatric; MMSE < 24.9	Age: 56.8 ± 18.8 (21–95) Education: 11.1 ± 4.8 M/F: 39.3/60.7%	CV (AS; MMSE) $r = 0.43$	/	/	/	/	TL + ES
(34) Tessari et al. [43]	STIMA In-person <5′	111 HPs	EC: neurological, psychiatric; left-handed (Edinburgh Test); age < 30 and > 90	Age: 60.2 ± 15.5 (30–84) Education: 9.8 ± 4.04 (4–20) M/F: 50.5/49.5%	/	/	/	/	Ceiling effect	TL + ES; percentiles
(35) Mancuso et al. [44]	OCS In-person CT <15′	498 HPs	Terranova Bracciolini; Verona; Roma; EC: internal, peripheral, neurological, psychiatric; MMSE < 22	Age: 18–89 M/F: 45.4/54.6%	/	/	/	/	/	Lower bound of the CI (*z*-score-like) on raw/predicted scores
(36) Poletti et al. [45]	ECAS In-person BT 20′	248 HPs; 107 ALS	Milano; EC (HPs): neurological, psychiatric; EC (ALS): internal, neurological, psychiatric	HPs: Age: 57.8 ± 10.6 Education: 14.1 ± 4.6 M/F: 42/58% ALS: Age: 63 ± 12.5 Education: 10.8 ± 4.2 M/F: 65/35%	CV (MoCA, FAB) ALS: $r = 0.7$; $r = 0.63$; HPs: $r = 0.52$	IRR: Cohen's $k = 0.99$ ($N = 30$) IC: Cronbach's α = 0.86 (ALS); 0.73 (HPs)	/	/	Acceptability	−2 *SD* below HPs' mean
(37) Ricci et al. [46]	CDT (three-cluster scoring system) In-person	79/102 AD; 96/104 MCI; 89/104/145 HPs	Perugia; EC: age < 65; neurological, psychiatric	AD: Age: 76.5 ± 5.5 Education: 52.2 ± 3.5 MCI: Age: 75.2 ± 5.8 Education: 7.4 ± 4.7 HPs: Age: 73.3 ± 6.9/71.99 ± 6.5 Education: 7.2 ± 4.2/7.92 ± 4.25	FS	IRR: $r = 0.94$–0.98	SE = 65–91% SP = 72–90%	/	/	ROC (clinical diagnosis); adjustment

Table 3 (continued)

Authors	Test	*N*	Sampling	Demographics	Validity	Reliability	Sensitivity and specificity	SE- and SP-derived measures	Other properties	Norming method
(38) Siciliano et al. [47]	ACE-R In-person 15–20′	528 HPs	EC: internal, neurological, psychiatric; MoCA < 15.5	Age: 52.3 ± 18.1 (20–93) Education: 10.9 ± 5.2 M/F: 46.6/53.4%	CV (AS; MoCA) $r = 0.57$ CV (AS; MMSE) $r = 0.6$	/	/	/	/	TL + ES
(39) Siciliano et al. [48]	CDT (Rouleou) In-person	872 HPs	Campania; Lombardia; Toscana; EC: internal, neurological, psychiatric	Age: 53.16 ± 18.27 (20–94) Education: 11.17 ± 4.98 (1–27) M/F: 44.6/55.4%	CV (AS; MMSE): $r_s = 0.13–0.23$	IRR: ICC = 0.89–0.97 Intra-RR: ICC = 0.91–0.99	/	/	/	TL + ES
(40) Bosco et al. [49]	MoCA In-person -	410 HPs; 40 probable AD	Bari; EC (HPs): neurological, psychiatric	HPs: Age: 72.2 ± 7.3 Education: 7.7 ± 4.6 M/F: 46/54% Probable AD: Age: 76.8 ± 8.94 Education: 5.8 ± 3.8 M/F: 40/60%	/	/	SE = 69.2–87.5 SP = 58.5–92.5	LR + = 1.67–7 LR − = 0.14–0.53	AUC = 0.69–0.95	ROC (MMSE)
(41) Catricalà et al. [50]	SAND In-person CT	134 HPs	EC: neurological, psychiatric; MMSE < 24	Age: 63.3 ± 11.2 (45–85) Education: 11 ± 5 (2–25) M/F: 42/58%	/	/	/	/	/	TL; adjustment
(42) Santangelo et al. [51]	PD-CRS In-person -	268 HPs	Naples; Genova; Verona; EC: age < 30; age > 79; MoCA < 15.5; BDI-II < 13; neurological, psychiatric	Age: 54.5 ± 12.6 Education: 12.5 ± 5 M/F: 46.7/53.3%	CV (AS; MoCA): $r = 0.55$	/	/	/	/	TL + ES
(43) Siciliano et al. [52]	ECAS In-person 15–20′	277 HPs	EC: internal, neurological, psychiatric; MoCA < 15.5	Age: 55.3 ± 13.2 (30–79) Education: 10.9 ± 5 M/F: 45.5/54.5%	CV (AS; MoCA): $r = 0.67$	/	/	/	Ceiling effect	TL + ES

Table 3 (continued)

Authors	Test	*N*	Sampling	Demographics	Validity	Reliability	Sensitivity and specificity	SE- and SP-derived measures	Other properties	Norming method
(44) Battista et al. [53]	SAND In-person <20′	30 PPA 45 MD (PD or PSP) 130 HPs	Milan; Florence; Bari; Salerno; EC (pts): no Italian native speaker; internal, neurological, psychiatric; MMSE < 10; EC (HPs): neurological, psychiatric; MMSE < 24	PPA: Age: 70.9 ± 6.04 Education: 11.67 ± 4.94 M/F: 53.3/46.7 MD: Age: 66.98 ± 8.05 Education: 11.11 ± 4.90 M/F: 60/40% HPs: Age: 63.3 ± 11.3 (45–85) Education: 10.5 ± 4.89 (2–25) M/F: 41.5/58.5%	CV (SAND sub-tests and conventional language tests): r_s = l.45l–l.84l CV (global score with MMSE and DSB): r_s = l.56l; l.34l DV (memory, attention, executive functioning, constructional praxis): r_s = n.s	Cronbach's α = 0.86	SE = 77–93% SP = 67–95%	PPV = 60–80% NPV = 81–98%	AUC = 0.79–0.98 Ceiling effect	ROC (diagnostic criteria)
(45) Crivelli et al. [54]	CASP In-person CT 10–15′	14 stroke; 15 other pts	Cassano Murge; EC: disease-related, peripheral, internal	Post-stroke: Age: 71.1 ± 8.2 Education: 7 ± 2.5 M/F: 57/43% Other pts.: Age: 56.7 ± 14.8 Education: 7.6 ± 2.6 M/F: 53/47	CV (MMSE, MoCA): r = 0.75; 0.8	IRR: ICC (MMSE, MoCA) = 0.7; 0.43	/	/	/	/
(46) Carpinelli Mazzi et al. [70]	MMSE In-person	47 AD; 314 HPs	Campania; EC (HPs): internal, neurological, psychiatric	AD: Age: 76.6 ± 7.7 Education: 9.3 ± 5.4 HPs: Age: 63.4 ± 9 Education: 11.5 ± 4.4 M/F: 51.3/48.7%	/	/	/	/	/	5th cent. + ES
(47) Iavarone et al. [55]	Q*mci*-I In-person BT 5′	307 HPs	Campania; EC: MMSE ≥ 25; neurological, psychiatric, internal, drugs	Age: 63 ± 8.6 Education: 11.7 ± 4.3 M/F: 50.5/49.5	CV (MMSE): r = 0.2	/	/	/	Ceiling effect	TL + ES
(48) Panebianco et al. [56]	ART In-person CT	52 stroke	Catania; EC: impaired consciousness	Age: 73.73 ± 28.99 M/F: 53.8/46.2%	/	IRR: ICC = 0.99; Cohen's weighted k = 0.88	/	/	/	/

Table 3 (continued)

Authors	Test	*N*	Sampling	Demographics	Validity	Reliability	Sensitivity and specificity	SE- and SP-derived measures	Other properties	Norming method
(49) Pigliautile et al. [57]	ACE-III In-person BT	574 HPs	Perugia; Padova; Taranto EC: neurological, psychiatric, drugs	Age: 68.7 ± 9.7 (50–94) Education: 9.2 ± 4.04 (2–24) M/F: 42.5/57.5%	/	/	/	/	/	TL + ES
(50) Siciliano et al. [58]	MoCA In-person ST	302 MoCA 2; 413 MoCA 3	Campania; EC: neurological, psychiatric	MoCA 2: Age: 53.4 ± 16.9 (20–87) Education: 11 ± 4.7 (1–12) M/F: 48/52% MoCA 3: Age: 54.6 ± 18.6 (18–89) Education: 12.1 ± 4.2 (2–18) M/F: 50/50%	CV (AS; MoCA 2—MMSE) $r = 0.69$ CV (AS; MoCA 3—MMSE) $r = 0.61$	PF	/	/	/	TL + ES
(51) Smirni et al. [59]	RMT In-person 10′	100 HPs	EC: 50 ≤ age ≤ 79; education ≥ 5; MMSE < 26; neurological, psychiatric	Age: 64.37 ± 8.43 (50–79) Education: 12.08 ± 4.54 (5–18) M/F: 46/54%	CV(RMT): $r = 0.88$–0.9	PF	/	/	/	TL + ES
(52) Belvederi Murri et al. [60]	SCIP In-person BT 15′	120 HPs	Ferrara; Aosta EC.: WAIS Voc. defective	Age: 24.1 ± 2.7 Education: 16.3 ± 1.8 M/F: 50/50%	FS	IC: Cronbach's α = 0.7 TRR (2 d.): $r = 0.72$ PF	/	/	Acceptability	
(53) Muò et al. [61]	I-AABT In-person 10–40′ CT	116 acute aphasia; 54 post-acute aphasia; 48 right hemisphere damage; 30 HPS	Turin; Milan EC (HPs): neurological	/	CV (AAT; 30 post-acute aphasia) $r_s = 0.61$–0.95	TRR (24 h; 25 acute aphasia): $r_s = (0.84, 1)$ ICC = (0.84, 1) IRR (21 acute aphasia): $r_s = 0.79$–1 ICC = 0.7–1	Comprehension part SE = 72.9% SP = 79.5% Production part: SE = 75.2% SP = 74.4%	/	/	ROC (clinical diagnosis)
(54) Pasotti et al. [62]	MEPS In-person	27 acute stroke 129 stroke 263 HPs	EC (stroke): age < 18; internal, neurological, psychiatric EC (HPs): internal, neurological, psychiatric	27 acute stroke: Age: 70.7 ± 11 Education: 10.14 ± 4 M/F: 59.3/40.7% 129 stoke: Age: 67.8 ± 13.9 Education: 8.7 ± 4.3 M/F: 31/69% HPs: Age: 52.6 ± 15.9 Education: 11.4 ± 4.3 M/F: 43.7/56.3%	Acute stroke CV (MoCA): $r = 0.88$	/	SE = 58.6% SP = 93.8%	/	AUC = 0.8 Ceiling effect	ES
(55) Russo et al. [63]	I-MIG SCOG In-person BT	20 HPs; 153 MwoA	Napoli; EC: drugs	Age: 35.1 ± 11 (18–61) Education: 13 ± 3.7 (5–18)	DV (MoCA): r = n.s	IC: Cronbach's α = 0.81	/	/	Ceiling and floor effects	/

Table 3 (continued)

Authors	Test	*N*	Sampling	Demographics	Validity	Reliability	Sensitivity and specificity	SE- and SP-derived measures	Other properties	Norming method
(56) Tremolizzo et al. [64]	ALS-CBS In-person BT 5–10′	100 ALS; 458 HPs	Milan; Naples; Messina EC (HPs): internal, neurological, psychiatric; MMSE < 23.8; FAB < 13.5 EC (ALS): dementia; tracheostomy	ALS: Age: 64.3 ± 10.2 (42–82) Education: 10.1 ± 4.2 (3–25) M/F: 67/33% HPs: Age: 56.4 ± 16.8 (20–89) Education: 11.4 ± 4.7 (2–27) M/F: 42.1/57.7%	ALS: CV (AS; FAB): *r* = *0.6* CV (AS; WST): *r* = *0.59*	ALS: IRR: Cohen's *k* (FAB) = 0.55 Cohen's *k* (WST) = 0.6	/	/	Acceptability	TL; adjustment
(57) Aiello et al. [65]	FAB In-person	475 HPs	Lombardy; EC: internal, neurological, psychiatric	Age: 61.1 ± 15.1 (21–96) Education: 11.7 ± 4.6 (1–25) M/F: 64.4/35.6%	CV (MoCA): *r* = *0.49*	/	/	/	/	TL + ES
(58) Aiello et al. [66]	MoCA In-person	579 HPs	Lombardy; EC: internal, neurological, psychiatric	Age: 63.44 ± 15.04 (21–96) Education: 11.27 ± 4.6 (1–25)	FS	IC: Cronbach's α = 0.81	-	-	IRT	TL + ES
(59) Barulli et al. [67]	TYM-I In-person CT 10′	94 MCI (*N* = 40 aMCI; *N* = 54 naMCI) 134 HCs	Tricase; Lecce; Puglia: EC: internal, neurological, psychiatric	MCI: Age: 70.5 ± 9.2 (43–87) Education: 7.9 ± 4.7 (0–18) M/F: 44.68/55.32% HCs: Age: 64.2 ± 8.2 (42–83) Education: 9.8 ± 4.1 (3–17) M/F: 45.52/54.48% aMCI: Age: 73.9 ± 7.8 (59–85) Education: 7.6 ± 4.9 (0–18) M/F: 52.5/47.5% naMCI: Age: 68 ± 9.4 (43–87) Education: 8.1 ± 4.6 (2–18) M/F: 38.89/61.11%	CV (MMSE, FAB): *r* = 0.63; 0.62	IC: Cronbach's α = 0.78	SE = 67.02% SP = 88.06%	/	AUC = 0.85 Ceiling and floor effects	ROC (clinical diagnosis); regression coefficients
(60) De Vreese et al. [68]	s-PCFT-I In-person BT 10–15′	165 non-DS ID 46 DS	EC: neurological, peripheral, psychiatric, disease-related	Age: (40–84) M/F: 59/41%	CV (DLD-I SCS): $r_s = -0.66$ DV (DLD-I SCS): $r_s = -0.38$	TRR: ICC = 0.85 IRR: ICC = 0.9 IC: Cronbach's α = 0.85 PF	/	/	Ceiling and floor effects	/

Table 3 (continued)

Authors	Test	*N*	Sampling	Demographics	Validity	Reliability	Sensitivity and specificity	SE- and SP-derived measures	Other properties	Norming method
(61) Montanucci et al. [69]	HDS-IT In-person BT	44 pts (MS, SIVD, NPH, HIV +) 180 HPs	EC (HPs): 40 ≤ age ≤ 85; internal, neurological, psychiatric; MMSE < 24	pts: Age: 64.9 ± 10.6 Education: 11.1 ± 4.8 M/F: 47.7/52.3% HPs: Age: 67.5 ± 8.3 Education: 11.3 ± 4.1 M/F: 42.2/57.8%	scCI CV (MMSE): $r_s = 0.5$	HPs TRR (3–10 mo.): $r_s = 0.7$	SE = 0.7 SP = 0.82	/	AUC = 0.8	ROC (clinical diagnosis)

4AT 4 'A's Test, *ACE-III* Addenbrooke's Cognitive Examination III, *ACE-R* Addenbrooke's Cognitive Examination Revised, *AD* Alzheimer's disease, *ADAS-Cog* Alzheimer's Disease Assessment Scale-Cognitive subscale, *ALS* amyotrophic lateral sclerosis, *ALS-CBS* ALS Cognitive Behavioral Screen, *AM* Attentional Matrices, *aMCI* amnesic MCI, *AQT* A Quick Test of Cognitive Speed, *ART* Aphasia Rapid Test, *AS* adjusted score, *AUC* area under the curve, *BACS* Brief Assessment of Cognition in Schizophrenia, *BNS* Brief Neuropsychological Screening, *BT* back-translation, *CAMCOG* Cambridge Cognitive Assessment, *CASP* Cognitive Assessment for Stroke Patients, *CDR* Clinical Dementia Rating, *CDT* Clock Drawing Test, *CI* cognitive impairment, *CrV* criterion validity, *CSI* cognitive screening instrument, *CT* controlled translation, *CV* convergent validity, *DLD-I-SCS/SOS* Italian Dementia Questionnaire for Persons with Intellectual Disabilities, *DRS* Dementia Rating Scale, *DS* Down syndrome, *DSB* digit span backward, *FS* factorial structure, *DSS* Digit-Symbol Substitution subtest of WAIS-R, *ECAS* Edinburgh Cognitive and Behavioral ALS Screen, *ENPA* Esame Neuropsicologico per l'Afasia, *EC* exclusion criteria, *ES* Equivalent Score, *F* female, *FAB* Frontal Assessment Battery, *FTD* frontotemporal degeneration, *FV* face validity, *GDS* Geriatric Depression Scale, *GPCOG-It* General Practitioner Cognitive Assessment of Cognition Italian version, *HAMTS* Hodkinson Abbreviated Mental Test, *HDS-IT* HIV Dementia Scale Italian Version, *HPs* healthy participants, *I-AABT* Italian Aachener Aphasie Bedside Test, *IC* internal consistency, *ICC* intra-class correlation, *ID* intellectual disabilities, *I-MIG SCOG* Italian version of the MIGraine attacks-Subjective COGnitive impairment scale, *intern.* general medical conditions, *IOCC* Improvement over chance criterion, *IRR* inter-rater reliability, *IRT* Item Response Theory, *Itel-MMSE* Italian telephone Mini-Mental State Examination, *I-TICS* Italian Telephone Interview for Cognitive Status, *LBD* Lewy Body dementia, *LR+/−* positive/negative likelihood ratio, *M* male, *MCI* mild cognitive impairment, *MDD* major depressive disorder, *MEPS* Mental Performance in Acute Stroke, *MMP* Mini-Mental Parkinson, *MMSE* Mini-Mental State Examination, *MoCA* Montreal Cognitive Assessment, *MODA* Milan Overall Dementia Assessment, *MR* misclassification rate, *MwoA* migraine without aura, *naMCI* non-amnesic MCI, *neurol.* neurological comorbidities, *NPV* negative predictive value, *OCS* Oxford Cognitive Screen, *PANDA* Parkinson Neuropsychometric Dementia Assessment, *PD* Parkinson's disease, *PD-CRS* Parkinson's Disease-Cognitive Rating Scale, *PDD* Parkinson's disease dementia, *periph.* sensory-motor deficits, *PPA* primary progressive aphasia, *PPV* positive predictive value, *PSP* progressive supranuclear palsy, *psych.* psychiatric comorbidities, *Qmci-I* Italian version of Quick mild cognitive impairment screen, *RMT* Recognition Memory Test, *ROC* receiver-operating characteristic, *SAND* Screening for Aphasia in NeuroDegeneration, *SCEB* Short Cognitive Evaluation Battery, *SE* sensitivity, *SP* specificity, *s-PCFT-I* Short forms of Prudhoe Cognitive Function Test Italian version, *scCI* subcortical cognitive impairment, *SCIP* Screen for Cognitive Impairment in Psychiatry, *SCOPA-Cog* Scales for Outcomes in Parkinson's disease-Cognition, *ST* simple translation, *STIMA* Short Test for Ideomotor Apraxia; Sum of Cognitive Scores/Sum of Social Scores, *TL* tolerance limit, *TMT-B* Trail Making Test-Part B, *TRR* test-retest reliability, *TYM-I* Test Your Memory, *VaD* vascular dementia, *VMMSE* Videoconference Mini-Mental State Examination, *WAIS-R* Weschler Adult Intelligence Scale-Revised, *WST* Weigl's Sorting Test. Diagnostics (e.g., SE, AUC) ranges are reported regardless of their reference sub-sample(s)/comparison(s)

Table 4 Synopsis on availability psychometric and diagnostic properties of Italian cognitive screening tests

Cognitive screening test		Validity				Reliability					SE, SP, and derived metrics						Norms				Feasibility				
	N	CV	CrV	DV	FS	TRR	IRR	IC	PF	RCI	SE	SP	PPV	NPV	LR+	LR−	AUC	Adj	TL	ES	BT	Acc	C/F	Time	CSSC
General — in-person																									
MMSE	6	□		□			■				■	■	□	□				■	■	■		■			27
MoCA	6	■		□	■	■	■	■	■		■	■	■	■	■	■	■	■	■	■				10′	34
ACE-R	3	■						■			■	■	■	■			■	■	■	■	■			15–20′	31
CDT	3	■			■	■	■				■	■					■	■	■	■					25
MODA	2	■				■	■				■	■	■	■				■	■				■	30–45′	25
ACE-III	1																	■	■	■	■				17
BNS	1	■	■																			■	■	5–10′	14
HAMTS	1										■	■						■							8
GPCOG	1	■									■	■	■	■			■							< 5′	13
Mini-Cog	1						■															■		3–5′	7
Q*mci*-i	1	■																■	■	■	■		■	5′	16
SCEB	1	■									■	■	■	■			■							< 10′	16
TICS	1		■			■					■	■													15
TYM-I	1	■						■			■	■					■	■					■	10′	24
General — remote																									
Itel-MMSE	2	■	■			■	■	■			■	■													13
I-TICS	1	■	■			■	■	■			■	■												10′	14
VMMSE	1					■	■				■	■	■	■			■							10′	14
Domain-specific																									
FAB	3	■		■	■	■	■	■										■	■	■				10′	24
SAND	2	■		■				■			■	■	■	■			■	■	■				■	< 20′	27
AQT	1																						■	15′	6
ART	1						■																		4
I-AABT	1	■				■	■				■	■												15–40′	10
RMT	1	■						■										■	■	■				10′	16
STIMA	1																			■	■		■	< 5′	16
Disease-specific																									
ECAS	2	■					■	■										■	■	■		■	■	15–20′	21
SCOPA-Cog	2	■		■				■			■	■	■	■			■					■	■	20′	18
PD-CRS	2	■						■										■	■	■	■	■		10′	19
ADAS-Cog	1	□			■													■							6
ALS-CBS	1	■					■											■	■		■	■		5–10′	26
BACS	1																		■	■	■				15
CASP	1	■					■																	10–15′	6

Table 4 (continued)

Cognitive screening test		Validity				Reliability					SE, SP, and derived metrics						Norms				Feasibility				
	N	CV	CrV	DV	FS	TRR	IRR	IC	PF	RCI	SE	SP	PPV	NPV	LR+	LR−	AUC	Adj	TL	ES	BT	Acc	C/F	Time	CSSC
HDS-IT	1	■				■					■	■					■				■				19
I-MIG-SCOG	1			■				■													■		■		11
MEPS	1	■									■	■					■			■			■		17
MMP	1	□																■	■	■				10–15′	15
OCS	1																							< 15′	11
PANDA	1										■	■													7
SCIP	1				■			■	■												■	■		15′	10
s-PCFT-I	1	■		■		■	■	■	■												■		■	10–15′	12
4AT	1							■			■	■			■	■	■						■	2′	11

Black square means explicitly investigated within reference studies; white square means not investigated within reference studies but retrievable from different ones. *N* refers to the number of studies of a given cognitive screening test. *4AT* 4 'A's Test, *ACE-III* Addenbrooke's Cognitive Examination III, *ACE-R* Addenbrooke's Cognitive Examination Revised, *ADAS-Cog* Alzheimer's Disease Assessment Scale-Cognitive subscale, *Adj.* adjustment for background variables, *ALS-CBS* ALS Cognitive Behavioral Screen, *AQT* A Quick Test of Cognitive Speed, *ART* Aphasia Rapid Test, *BACS* Brief Assessment of Cognition in Schizophrenia, *BNS* Brief Neuropsychological Screening, *BT* back-translation, *CASP* Cognitive Assessment for Stroke Patients, *CDT* Clock Drawing Test, *C/F* ceiling/floor effects, *CrV* criterion validity, *CSSC* Cognitive Screening Standardization Checklist, *CV* concurrent validity, *DV* divergent validity, *ECAS* Edinburgh Cognitive and Behavioral ALS Screen, *ES* equivalent scores, *FAB* Frontal Assessment Battery, *FS* factorial structure, *GPCOG-It* General Practitioner Cognitive Assessment of Cognition Italian version, *HAMTS* Hodkinson Abbreviated Mental Test, *HDS-IT* HIV Dementia Scale Italian Version, *I-AABT* Italian Aachener Aphasie Bedside Test, *IC* internal consistency, *I-MIG SCOG* Italian version of the MIGraine attacks-Subjective COGnitive impairment scale, *IRR* inter-rater reliability, Itel-MMSE Italian telephone Mini-Mental State Examination, *I-TICS* Italian Telephone Interview for Cognitive Status, *LR−* negative likelihood ratio, *LR+* positive likelihood ratio, *MEPS* Mental Performance in Acute Stroke, *MMP* Mini-Mental Parkinson, *MMSE* Mini-Mental State Examination, *MoCA* Montreal Cognitive Assessment, *MODA* Milan Overall Dementia Assessment, *NPV* negative predictive value, *OCS* Oxford Cognitive Screen, *PANDA* Parkinson Neuropsychometric Dementia Assessment, *PD-CRS* Parkinson's Disease-Cognitive Rating Scale, *PF* parallel forms, *PPV* positive predictive value, *Qmci-I* Italian version of Quick mild cognitive impairment screen, *RMT* Recognition Memory Test, *ROC* receiver-operating characteristics, *SAND* Screening for Aphasia in NeuroDegeneration, *SCEB* Short Cognitive Evaluation Battery, *SCIP* Screen for Cognitive Impairment in Psychiatry, *SCOPA-Cog* Scales for Outcomes in Parkinson's disease-Cognition, *SE* sensitivity, *SP* specificity, *s-PCFT-I* Short forms of Prudhoe Cognitive Function Test Italian version, *STIMA* Short Test for Ideomotor Apraxia, *Time* administration time, *TYM-I* Test Your Memory, *TL* tolerance limits, *TRR* test-retest reliability, *VMMSE* Videoconference Mini-Mental State Examination

Reliability

Overall, reliability of Italian CSTs was frequently assessed, although often either incompletely or inefficiently.

When testing reliability of CSTs, it is worth bearing in mind that internal consistency might be problematic: indeed, different items within the same CST may be meant to measure different facets of cognition, this possibly being even truer for multi-domain tests such as the MMSE. This is an aspect that needs further developments.

By contrast, assessing reliability via test-retest or inter-rater may be generalizable to different CST categories and more practically relevant (e.g., clinicians are interested in knowing whether the CST yields similar outcomes/scores when administered in different conditions).

Furthermore, parallel-form reliability was seldom examined, and no CST came with information on its ability to detect significant change [6, 76]. Indeed, although parallel forms reduce the possibility to have "practice effect" (i.e., systematic performance improvements across consecutive assessments), the lack of appropriate methods for detecting clinically meaningful changes over time unrelated to practice has a crucial (even detrimental) impact whenever CSTs are meant to be used longitudinally to monitor the progress of cognitive functions or dysfunctions for either diagnostic or prognostic purposes. Indeed, without thresholds for significant change it is not possible to ascertain whether observed score variations over repeated measurements could be merely traced back to intrinsic and expected, physiological oscillations of performances, or whether they more likely reflect a true cognitive change (worsening or improvement).

Diagnostic properties

The study of the diagnostic properties of CSTs was often addressed within a nosographic-descriptive framework, which, however, might not always fit cognitive semiology [7]. Indeed, a one-to-one correspondence between cognitive profiles and neurological/neuropsychiatric conditions is often not straightforward [13]. Thereupon, the notion of "target condition" within ROC analyses may happen to be elusive, hence limiting the disease-specificity of certain CSTs [2, 7]. The present work indeed highlights the need for identifying more rigorous statistical methods for deriving the optimal cut-off values (e.g., the Youden statistics, an index identifying the best cut-off at the optimal compromise between sensitivity and specificity).

Moreover, although basic diagnostic properties happened to be investigated, less attention has been given to those selectively relevant to screening aims, such as taking into account disease prevalence (e.g., positive and negative predictive values) and allowing an estimation of post-test probability of cognitive impairment (e.g., positive and negative likelihood ratios) [3].

With respect to Italian CSTs, it is noteworthy that diagnostic properties were almost investigated only for 5 out of 16 disease-specific CSTs. Although evidence of case-control discrimination was frequently provided by means of between-group comparisons (e.g., ALS-CBS), it is recommended that sensitivity, specificity, and derived measures be tested in order to statistically substantiate CST applications to target conditions.

Norms

As far as normative data are concerned, although regression-based and inferential-error-controlling techniques were highly represented, it should be noted that a relatively high heterogeneity in approaching norming methods was detected. First, the Equivalent Score method happened to be embraced "incompletely," by only computing tolerance limits only but not Equivalent Score thresholds [12]. Second, norms were occasionally derived via approaches assuming a normal distribution, this possibly undermining their adequacy as cognitive data often present with overdispersion and skewness [43]. With this last respect, checking for ceiling/floor effects in test scores is encouraged; unfortunately, this was rarely carried out in the studies herewith included.

With a few exceptions, sampling revealed to be overall adequate as far as typical/clinical sample sizes are concerned. However, it has to be noted that geographic coverage of normative samples was often circumscribed. Between-regional differences should nonetheless receive attention as being potential confounders in cognitive testing; this issue has been only recently addressed within the concerning Italian literature [66, 70].

Feasibility

Despite a core feature of a CST is a short administration time [2], it should be noted that several CSTs here reported require up to 20′ to be administered and scored (e.g., ACE-R), in turn limiting their usability in time-restricted settings (e.g., bedside evaluations). By contrast, these "in-depth CSTs" may be more adequate in outpatient settings [2].

With regard to cross-cultural/-linguistic adaptations, it has to be stressed that back-translation approaches have been seldom adopted and culture/language-related issues often not addressed. The latter aspect is of major interest especially for items assessing language functioning, which should instead undergo dedicated, country-specific controls for psycholinguistic predictors (e.g., word frequency for naming tasks) [10].

An overall need for more systematic evidence of acceptability/face validity of CSTs also emerged from the present

work. This indeed would help practitioners select a test based on the administration setting, as for instance the assessment of acute patients would benefit from a short, tolerable CST that is clearly recognizable as such by the patient.

Limitations and perspectives

First, it has to be noted that the goodness of a CST is not exhausted in psychometric/diagnostic properties. Indeed, in order for a CST to be introduced in clinical practice, thorough evidence of its applicability in atypical populations should be provided. Moreover, it should be born in mind that evidence of both psychometric and diagnostic soundness may be inferred from applied studies as well. Thereupon, future studies should focus on reviewing available contributions on the clinical usability of Italian CTSs in order to provide a more comprehensive picture on their statistical/methodological quality.

Furthermore, Italian practitioners might benefit from a future review focused on psychometric/diagnostic properties of qualitative/proxy measures of cognition that were not addressed within the present study for generalizability reasons.

Although beyond the aim of this work, it should be then noted that more detailed item-level analyses (Item Response Theory) were conducted in only one of the records included [66]. As being able to provide insights into adaptive testing as well as to help ease interpretations issues, Item Response Theory-based analyses should be taken into consideration when assessing psychometrics/diagnostic properties of CSTs [66].

Finally, it is important to underline that, to the best of the authors' knowledge, there is not an official, worldwide consensus on the relevant properties to be addressed in cognitive screening, this resulting in the choices being possibly incomplete or selectively reflecting the knowledge of researchers. This latter consideration stresses the importance of developing wider agreement within neurological/neuropsychological societies to ensure higher standards and raise the awareness on the impact of statistical properties on the applicability of CSTs in both applied (e.g., clinical and forensic) and research contexts.

Conclusion

The present work shows that, although available Italian CSTs overall met basic psychometric/diagnostic requirements, their statistical profile often proved to be deficient on several properties that are desirable/needed for clinical applications, with a few exceptions among general and domain-specific CSTs yielding high soundness, namely, the MoCA and ACE-R, and the FAB and SAND, respectively. In particular, this work highlights that:

- psychometric/diagnostic properties of disease-specific CSTs happened were poorly examined;
- construct and criterion validity should be differentiated and assessed separately;
- factorial structure underlying CSTs should be tested for both general and domain-specific ones;
- ecological validity of CSTs need to be addressed to provide information relevant to patients' everyday functioning;
- significant change thresholds and alternate versions of CSTs need to be developed in order to improve their longitudinal usage;
- a general lack of investigations on sensitivity-/specificity-derived diagnostic metrics selectively relevant to screening aims (i.e., positive and negative predictive values and likelihood ratios) was detected;
- a clearer definition of target conditions for a given CST is needed, especially for those thought to be disease-specific;
- information on CST acceptability, face validity, and administration time are desirable, as helping an ad hoc usage by practitioners select.

Acknowledgements The authors would like to thank Dr. Francesca Crespi for her contributions.

Declarations

Ethical approval None.

Conflict of interest The authors declare no competing interests.

References

1. Larner AJ. Introduction to cognitive screening instruments: rationale and desiderata. In A. J. Larner (Ed.). Cognitive screening instruments: a practical approach (2nd ed., pp 3–14). Springer; 2017a.
2. Larner AJ. The usage of cognitive screening instruments: test characteristics and suspected diagnosis. In A. J. Larner (Ed.). Cognitive screening instruments: a practical approach (2nd ed., pp 315–339). Springer; 2017b.
3. Bolin E, Lam W. A review of sensitivity, specificity, and likelihood ratios: evaluating the utility of the electrocardiogram as a screening tool in hypertrophic cardiomyopathy. Congenit Heart Dis 2013;8:406–10.
4. Grigoletto F, Zappalà G, Anderson DW, Lebowitz BD. Norms for the Mini-Mental State Examination in a healthy population. Neurol 1999;53:315–20.
5. Crawford JR, Garthwaite PH. Comparing patients' predicted test scores from a regression equation with their obtained scores: a

significance test and point estimate of abnormality with accompanying confidence limits. Neuropsychology 2006;20:259.
6. Arcara G, Bambini V. A test for the assessment of pragmatic abilities and cognitive substrates (APACS): normative data and psychometric properties. Front Psychol 2016;7:70.
7. Quinn TJ, Takwoingi Y. Assessment of the utility of cognitive screening instruments. In A. J. Larner (Ed.). Cognitive Screening instruments: a practical approach (2nd ed., pp 15–36). Springer; 2017.
8. Rohling ML, Axelrod BN, Langhinrichsen-Rohling J. Fundamental forensic statistics: statistics every forensic neuropsychologist must know. In: Bush SS, Demakis GJ, Rohling ML (eds) APA Handbook of Forensic Neuropsychology. American Psychiatric Association Publishing; 2017. p. 3–22.
9. Pottie K, Rahal R, Jaramillo A, Birtwhistle R, Thombs BD, Singh H, Gorber SC, Dunfield L, Shane A, Bacchus M, Bell N. Recommendations on screening for cognitive impairment in older adults. CMAJ 2016;188:37–46.
10. Uysal-Bozkir Ö, Parlevliet JL, de Rooij SE. Insufficient cross-cultural adaptations and psychometric properties for many translated health assessment scales: a systematic review. J Clin Epidemiol 2013;66:608–18.
11. Pedraza O, Mungas D. Measurement in cross-cultural neuropsychology. Neuropsychol Rev 2008;18:184–93.
12. Bianchi A. L'esame neuropsicologico dell'adulto: Applicazioni cliniche e forensi. Giunti Psychometrics; 2013.
13. Mondini S, Mapelli D, Arcara G. Semeiotica e diagnosi neuropsicologica: metodologia per la valutazione. Carrocci; 2016.
14. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: the PRISMA statement. J Clin Epidemiol 2009;62:1006–12.
15. Mazzoni M, Ferroni L, Lombardi L, Del Torto E, Vista M, Moretti P. Mini-Mental State Examination (MMSE): sensitivity in an Italian sample of patients with dementia. Ital J Neurol Sci 1992;13:323–29.
16. Rocca WA, Bonaiuto S, Lippi A, Luciani P, Pistarelli T, Grandinetti A, Cavarzeran F, Amaducci L. Validation of the Hodkinson abbreviated mental test as a screening instrument for dementia in an Italian population. Neuroepidemiology 1992;11:288–95.
17. Measso G, Cavarzeran F, Zappalà G, Lebowitz BD, Crook TH, Pirozzolo FJ, Amaducci L, Massari D, Grigoletto F. The Mini-Mental State Examination: normative study of an Italian random sample. Dev Neuropsychol 1993;9:77–85.
18. Fioravanti M, Nacca D, Buckley AE, Ferrario E, Varetto O, Mogni P, Fabris F. The Italian version of the Alzheimer's Disease Assessment Scale (ADAS): psychometric and normative characteristics from a normal aged population. Arch Gerontol Geriatr 1994;19:21–30.
19. Magni E, Binetti G, Bianchetti A, Rozzini R, Trabucchi M. Mini-Mental State Examination: a normative study in Italian elderly population. Eur J Neurol 1996;3:198–202.
20. Ferrucci L, Del Lungo I, Guralnik JM, Bandinelli S, Benvenuti E, Salani B, Lamponi M, Ubezio C, Benvenuti F, Baroni A. Is the telephone interview for cognitive status a valid alternative in persons who cannot be evaluated by the Mini Mental State Examination? Aging Clin Exp Res 1998;10:332–38.
21. Cossa FM, Sala SD, Musicco M, Spinnler H, Ubezio MC. The Milan Overall Dementia Assessment and the Mini-Mental State Examination compared: an epidemiological investigation of dementia. Eur J Neurol 1999;6:289–94.
22. Metitieri T, Geroldi C, Pezzini A, Frisoni GB, Bianchetti A, Trabucchi M. The Itel-MMSE: an Italian telephone version of the Mini-Mental State Examination. Int J Geriatr Psychiatry 2001;16:166–67.
23. Iavarone A, Ronga B, Pellegrino L, Loré E, Vitaliano S, Galeone F, Carlomagno S. The Frontal Assessment Battery (FAB): normative data from an Italian sample and performances of patients with Alzheimer's disease and frontotemporal dementia. Funct Neurol 2004;19:191–95.
24. Appollonio I, Leone M, Isella V, Piamarta F, Consoli T, Villa ML, Forapani E, Russo A, Nichelli P. The Frontal Assessment Battery (FAB): normative values in an Italian population sample. Neurol Sci 2005;26:108–16.
25. Dal Forno G, Chiovenda P, Bressi F, Ferreri F, Grossi E, Brandt J, Rossini PM, Pasqualetti P. Use of an Italian version of the telephone interview for cognitive status in Alzheimer's disease. Int J Geriatr Psychiatry 2006;21:126–33.
26. Michieletto F, Binkin N, Saugo M, Boorson S, Scanlan J. Utilizzo del test Mini-Cog come metodo di screening per la demenza nella popolazione italiana: l'esperienza dello studio Argento. Sanità Pubblica 2006;62:159–72.
27. Vanacore N, De Carolis A, Sepe-Monti M, Bomboi G, Stazi A, Bianchetti A, Giubilei F. Validity of the Italian telephone version of the mini-mental state examination in the elderly healthy population. Acta Neurol Belg 2006;106:132–36.
28. Anselmetti S, Poletti S, Ermoli E, Bechi M, Cappa S, Venneri A, Smeraldi E, Cavallaro R. The brief assessment of cognition in schizophrenia. Normative data for the Italian population. Neurol Sci 2008;29:85–92.
29. Lunardelli A, Mengotti P, Pesavento A, Sverzut A, Zadini A. The Brief Neuropsychological Screening (BNS): valuation of its clinical validity. Eur J Phys Rehabil Med 2009;45:85–91.
30. Caffarra P, Gardini S, Zonato F, Concari L, Dieci F, Copelli S, Freedman M, Stracciari A, Venneri A. Italian norms for the Freedman version of the Clock Drawing Test. J Clin Exp Neuropsychol 2011;33:982–88.
31. Girtler N, Brugnolo A, Campus C, De Carli F, Famà F, Ghio L, Robert P, Barbieri MP, Delmonte M, Venuti S, Gabrielli F, Rodriguez G, Nobili F. The short cognitive evaluation battery in cognitive disorders of the elderly-Italian version. Dement Geriatr Cogn Disord 2012;33:255–65.
32. Costa A, Bagoj E, Monaco M, Zabberoni S, De Rosa S, Mundi C, Caltagirone C, Carlesimo GA. Mini mental Parkinson test: standardization and normative data on an Italian sample. Neurol Sci 2013;34:1797–803.
33. Isella V, Mapelli C, Morielli N, De Gaspari D, Siri C, Pezzoli G, Antonini A, Poletti M, Bonuccelli U, Picchi L, Napolitano A, Vista M, Veglia M, Piamarta F, Grassi F, Appollonio IM. Psychometric properties of the Italian version of the Scales for Outcomes in Parkinson's disease-Cognition (SCOPA-Cog). Funct Neurol 2013;28:121.
34. Isella V, Mapelli C, Morielli N, Siri C, De Gaspari D, Pezzoli G, Antonini A, Poletti M, Bonuccelli U, Picchi L, Napolitano A. Diagnosis of possible mild cognitive impairment in Parkinson's disease: validity of the SCOPA-Cog. Parkinsonism Relat Disord 2013;19:1160–63.
35. Timpano F, Pirrotta F, Bonanno L, Marino S, Marra A, Bramanti P, Lanzafame P. Videoconference-based mini mental state examination: a validation study. Telemed J E Health 2013;19:931–37.
36. Bellelli G, Morandi A, Davis DH, Mazzola P, Turco R, Gentile S, Ryan T, Cash H, Guerini F, Torpillesi T, Del Santo F, Trabucchi M, Annoni G, MacLullich AM. Validation of the 4AT, a new instrument for rapid delirium screening: a study in 234 hospitalised older people. Age Ageing 2014;43:496–502.
37. Pignatti R, Bertella L, Scarpina F, Mauro A, Portolani E, Calabrese P. Italian version of the Parkinson Neuropsychometric Dementia Assessment (PANDA): a useful instrument to detect cognitive impairments in Parkinson's disease. J Parkinsons Dis 2014;4:151–160.
38. Santangelo G, Barone P, Abbruzzese G, Ferini-Strambi L, Antonini A, IRIS Study Group. Validation of the Italian version

of Parkinson's disease-cognitive rating scale (PD-CRS). Neurol Sci 2014;35:537–44.
39. Conti S, Bonazzi S, Laiacona M, Masina M, Coralli MV. Montreal Cognitive Assessment (MoCA)-Italian version: regression based norms and equivalent scores. Neurol Sci 2015;36:209–214.
40. Pigliautile M, Chiesi F, Rossetti S, Conestabile Della Staffa M, Ricci M, Federici S, Chiloiro D, Primi C, Mecocci P. Normative data for the ACE-R in an Italian population sample. Neurol Sci 2015;36:2185–90.
41. Pirrotta F, Timpano F, Bonanno L, Nunnari D, Marino S, Bramanti P, Lanzafame P. Italian validation of Montreal cognitive assessment. Eur J Psychol Assess 2015;31:131–37.
42. Santangelo G, Siciliano M, Pedone R, Vitale C, Falco F, Bisogno R, Siano P, Barone P, Grossi D, Santangelo F, Trojano L. Normative data for the Montreal Cognitive Assessment in an Italian population sample. Neurol Sci 2015;36:585–91.
43. Tessari A, Toraldo A, Lunardelli A, Zadini A, Rumiati RI. STIMA: a short screening test for ideo-motor apraxia, selective for action meaning and bodily district. Neurol Sci 2015;36:977–84.
44. Mancuso M, Varalta V, Sardella L, Capitani D, Zoccolotti P, Antonucci G, Italian OCS Group. Italian normative data for a stroke specific cognitive screening tool: the Oxford Cognitive Screen (OCS). Neurol Sci 2016;37:1713–21.
45. Poletti B, Solca F, Carelli L, Madotto F, Lafronza A, Faini A, Monti A, Zago S, Calini D, Tiloca C, Doretti A, Verde F, Ratti A, Ticozzi N, Abrahams S, Silani V. The validation of the Italian Edinburgh Cognitive and Behavioural ALS Screen (ECAS). Amyotrophic lateral Sclerosis & Frontotemporal Degeneration 2016;17:489–98.
46. Ricci M, Pigliautile M, D'Ambrosio V, Ercolani S, Bianchini C, Ruggiero C, Vanacore N, Mecocci P. The clock drawing test as a screening tool in mild cognitive impairment and very mild dementia: a new brief method of scoring and normative data in the elderly. Neurol Sci 2016;37:867–73.
47. Siciliano M, Raimo S, Tufano D, Basile G, Grossi D, Santangelo F, Trojano L, Santangelo G. The Addenbrooke's Cognitive Examination Revised (ACE-R) and its sub-scores: normative values in an Italian population sample. Neurol Sci 2016;37:385–92.
48. Siciliano M, Santangelo G, D'Iorio A, Basile G, Piscopo F, Grossi D, Trojano L. Rouleau version of the Clock Drawing Test: age- and education-adjusted normative data from a wide Italian sample. Clin Neuropsychol 2016;30:1501–16.
49. Bosco A, Spano G, Caffò AO, Lopez A, Grattagliano I, Saracino G, Pinto K, Hoogeveen F, Lancioni GE. Italians do it worse. Montreal Cognitive Assessment (MoCA) optimal cut-off scores for people with probable Alzheimer's disease and with probable cognitive impairment. Aging Clin Exp Res 2017;29:1113–20.
50. Catricalà E, Gobbi E, Battista P, Miozzo A, Polito C, Boschi V, Esposito V, Cuoco S, Barone P, Sorbi S, Cappa SF, Garrard P. SAND: a Screening for Aphasia in NeuroDegeneration. Development and normative data Neurological Sciences 2017;38:1469–83.
51. Santangelo G, Lagravinese G, Battini V, Chiorri C, Siciliano M, Abbruzzese G, Vitale C, Barone P. The Parkinson's Disease-Cognitive Rating Scale (PD-CRS): normative values from 268 healthy Italian individuals. Neurol Sci 2017;38:845–53.
52. Siciliano M, Trojano L, Trojsi F, Greco R, Santoro M, Basile G, Piscopo F, D'Iorio A, Patrone M, Femiano C, Monsurrò M, Tedeschi G, Santangelo G. Edinburgh Cognitive and Behavioural ALS Screen (ECAS)-Italian version: regression based norms and equivalent scores. Neurol Sci 2017;38:1059–68.
53. Battista P, Catricalà E, Piccininni M, Copetti M, Esposito V, Polito C, Miozzo A, Gobbi E, Cuoco S, Boschi V, Picillo M, Sorbi S, Barone P, Iannaccone S, Garrard P, Logroscino G, Cappa SF. Screening for aphasia in neurodegeneration for the diagnosis of patients with primary progressive aphasia: clinical validity and psychometric properties. Dement Geriatr Cogn Disord 2018;46:243–52.
54. Crivelli D, Angelillo MT, Grippa E, Colucci A, Nardulli R, Balconi M. When is a novel psychometric measure needed? A preliminary analysis regarding the Cognitive Assessment for Stroke Patients (CASP) battery compared with MMSE and MoCA. Appl Neuropsychol Adult 2018;25:410–16.
55. Iavarone A, Carpinelli Mazzi M, Russo G, D'Anna F, Peluso S, Mazzeo P, De Luca V, De Michele G, Iaccarino G, Abete P, Milan G, Garofalo E, Musella C, O'Caoimh R, Molloy W, De Joanna G, Manzo V, Ambra FI, Postiglione A, Illario M, Working Group. The Italian version of the quick mild cognitive impairment (Qmci-I) screen: normative study on 307 healthy subjects. Aging Clin Exp Res 2019;31:353–60.
56. Panebianco M, Zavanone C, Dupont S, Perrone P, Pavone A, Napolitano S, Leger A, Samson Y. The inter-rater reliability of the Italian version of Aphasia Rapid Test (ART) for acute ischemic stroke. Neurol Sci 2019;40:2171–74.
57. Pigliautile M, Chiesi F, Stablum F, Rossetti S, Primi C, Chiloiro D, Federici S, Mecocci P. Italian version and normative data of Addenbrooke's Cognitive Examination III. Int Psychogeriatr 2019;31:241–49.
58. Siciliano M, Chiorri C, Passaniti C, Sant'Elia V, Trojano L, Santangelo G. Comparison of alternate and original forms of the Montreal Cognitive Assessment (MoCA): an Italian normative study. Neurol Sci 2019;40:691–702.
59. Smirni D, Smirni P, Di Martino G, Fontana ML, Cipolotti L, Oliveri M, Turriziani P. Early detection of memory impairments in older adults: standardization of a short version of the verbal and nonverbal recognition memory test. Neurol Sci 2019;40:97–103.
60. Belvederi Murri M, Folesani F, Costa S, Morelli AC, Scillitani V, Guaiana G, Biancosino B, Caruso R, Nanni MG, Zerbinati L, Purdon SE, Grassi L. Italian validation of the screen for cognitive impairment in psychiatry. Community Ment Health J 2020;56:1411–18.
61. Muò R, Raimondo S, Martufi F, Cavagna N, Bassi M, Schindler A. Cross-cultural adaptation and validation of the Italian Aachener Aphasie Bedside Test (I-AABT), a tool for Aphasia assessment in the acute phase. Aphasiology. 2020;1–24.
62. Pasotti F, Magnani FG, Gallucci M, Salvato G, Ovadia D, Scotto M, Merolla S, Beretta S, Micieli GR, Agostoni EC, Beretta G, Bottini G. Neuropsychological assessment in acute stroke patients. Neurol Sci 2020;41:1259–66.
63. Russo A, Silvestro M, Garramone F, Tessitore A, Cropano M, Scotto di Clemente F, Trojano L, Tedeschi G, Santangelo G. A subjective cognitive impairments scale for migraine attacks: validation of the Italian version of the MIG-SCOG. Neurol Sci 2020;41(5):1139–43.
64. Tremolizzo L, Lizio A, Santangelo G, Diamanti S, Lunetta C, Gerardi F, Messina S, La Foresta S, Riva N, Falzone Y, Filippi M, Woolley SC, Sansone VA, Siciliano M, Ferrarese C, Appollonio I, ALS-CBS Italian Study Group. ALS Cognitive Behavioral Screen (ALS-CBS): normative values for the Italian population and clinical usability. Neurol Sci 2020;41:835–41.
65. Aiello EN, Esposito A, Gramegna C, Gazzaniga V, Zago S, Difonzo T, Appollonio I, Bolognini N. The Frontal Assessment Battery (FAB) and its sub-scales: validation and updated normative data in an Italian population sample. Neurol Sci 2021a;1–6.
66. Aiello EN, Gramegna C, Esposito A, Zago S, Difonzo T, Maddaluno O, Appollonio I, Bolognini N. The Montreal Cognitive Assessment (MoCA): updated norms and psychometric insights into adaptive testing from healthy individuals in Northern Italy. Aging Clin Exp Res. 2021b;1–8.
67. Barulli MR, Piccininni M, Brugnolo A, Musarò C, Di Dio C, Capozzo R, Tortelli R, Lucca U, Logroscino G. The Italian Version of the Test Your Memory (TYM-I): a tool to detect mild

cognitive impairment in the clinical setting. Front Psychol 2021;11:614920.
68. De Vreese LP, Gomiero T, De Bastiani E, Marangoni A, Weger E, Mantesso U (2021) Short forms of Prudhoe Cognitive Function Test in adults and aging people with intellectual disabilities: Italian validation study. J Intellect Disability Res: JIDR 65:162–72.
69. Montanucci C, Chipi E, Salvadori N, Rinaldi R, Eusebi P, Parnetti L. HIV-Dementia Scale as a screening tool for the detection of subcortical cognitive deficits: validation of the Italian version. J Neurol. 2021;1–7.
70. Carpinelli Mazzi M, Iavarone A, Russo G, Musella C, Milan G, D'Anna F, Garofalo E, Chieffi S, Sannino M, Illario M, De Luca V, Postiglione A, Abete P, with the support of the Working group. Mini-Mental State Examination: new normative values on subjects in Southern Italy. Aging Clin Exp Res 2020;32:699–702.
71. Brazzelli M, Capitani E, Della Sala S, Spinnler H, Zuffi M. A neuropsychological instrument adding to the description of patients with suspected cortical dementia: the Milan overall dementia assessment. J Neurol Neurosurg Psychiatry 1994;57:1510–17.
72. Petrazzuoli F, Palmqvist S, Thulesius H, Buono N, Pirrotta E, Cuffari A, Cambielli M, D'Urso M, Farinaro C, Chiumeo F, Marsala V, Wiig EH. A Quick Test of Cognitive Speed: norm-referenced criteria for 121 Italian adults aged 45 to 90 years. International Psychogeriatrics. 2014;1–8.
73. Pigliautile M, Ricci M, Mioshi E, Ercolani S, Mangialasche F, Monastero R, Croce MF, Federici S, Mecocci P. Validation study of the Italian Addenbrooke's Cognitive Examination Revised in a young-old and old-old population. Dement Geriatr Cogn Disord 2011;32:301–7.
74. Pirani A, Brodaty H, Martini E, Zaccherini D, Neviani F, Neri M. The validation of the Italian version of the GPCOG (GPCOG-It): a contribution to cross-national implementation of a screening test for dementia in general practice. Int Psychogeriatr 2010;22:82–90.
75. Taule T, Søvik M, Lein RK, Wehling E, Aßmus J, Rekand T. Psychometric properties of cognitive assessment in amyotrophic lateral sclerosis: a systematic review. Patient Related Outcome Measures 2020;11:181.
76. Jacobson NS, Roberts LJ, Berns SB, McGlinchey JB. Methods for defining and determining the clinical significance of treatment effects: description, application, and alternatives. J Consult Clin Psychol 1999;67:300.