Systematic Review

The Effect of Translation and Cultural Adaptations on Diagnostic Accuracy and Test Performance in Dementia Cognitive Screening Tools: A Systematic Review

Natasha C. Czerwinski-Alley[a], Tamara Chithiramohan[b], Hari Subramaniam[a,b], Lucy Beishon[c,d] and Elizabeta B. Mukaetova-Ladinska[a,b,*]

[a]*Department of Psychology and Visual Sciences, University of Leicester, Leicester, UK*
[b]*Leicestershire Partnership NHS Trust, Leicester, UK*
[c]*Department of Cardiovascular Sciences, University of Leicester, Leicester, UK*
[d]*NIHR Leicester Biomedical Research Centre, British Heart Foundation Cardiovascular Research Centre, Glenfield Hospital, Leicester, UK*

Abstract.
Background: The current cognitive tests have been developed based on and standardized against Western constructs and normative data. With older people of minority ethnic background increasing across Western countries, there is a need for cognitive screening tests to address factors which influence performance bias and timely diagnostic dementia accuracy. The diagnostic accuracy in translated and culturally adapted cognitive screening tests and their impact on test performance in diverse populations have not been well addressed to date.
Objective: This review aims to highlight considerations relating to the adaptation processes, language, cultural influences, impact of immigration, and level of education to assess for dementia in non-Western and/or non-English speaking populations.
Methods: We conducted a systematic search for studies addressing the effects of translation and cultural adaptations of cognitive screening tests (developed in a Western context) upon their diagnostic accuracy and test performance across diverse populations. Four electronic databases and manual searches were conducted, using a predefined search strategy. A narrative synthesis of findings was conducted.
Results: Search strategy yielded 2,890 articles, and seventeen studies (4,463 participants) met the inclusion criteria. There was variability in the sensitivity and specificity of cognitive tests, irrespective of whether they were translated only, culturally adapted only, or both. Cognitive test performance was affected by education, linguistic ability, and aspects of acculturation.
Conclusions: We highlight the importance of translating and culturally adapting tests that have been developed in the Western context. However, these findings should be interpreted with caution as results varied due to the broad selection of included cognitive tests.

Keywords: Alzheimer's disease, cognitive test, dementia, ethnic minority, mild cognitive impairment, translation

*Correspondence to: Elizabeta B. Mukaetova-Ladinska, Department of Psychology and Visual Sciences, University of Leicester, Leicester LE1 7RH, UK. E-mail: eml12@le.ac.uk.

INTRODUCTION

Over the last 50 years, the impact of globalization has led to an increase in people from culturally diverse backgrounds migrating to various destinations, including Western countries [1]. According to the census data from the Office for National Statistics (2021), approximately 18% of the United Kingdom (UK) population identifies as an ethnic minority, which includes those for whom English is not their first language [2].

Older people from ethnic minorities are at increased risk of receiving a diagnosis of dementia [3]. Physical health related factors, such as hypertension and diabetes, are partially responsible for this [4]. The higher likelihood of a dementia diagnosis could also be attributed to the reduced specificity of some tests, especially brief cognitive assessment tools [5], resulting in higher incidents of false positives, whereby dementia and mild cognitive impairment (MCI) are incorrectly identified by a test in a cognitively healthy (CH) individual.

Screening tests assist the diagnostic process of assessing cognitive impairment. Such tests are known to be influenced by multiple factors including age, years/level of education, language, culture, and literacy rates [6]. Eurocentric bias is also well-documented in the cognitive assessment process [7]. Many cognitive screening tests have been constructed in Western countries, i.e., developed predominantly in the United Kingdom (UK) and United States of America (USA), and therefore standardized on primarily Western, Caucasian, and English-speaking samples and cultures [8]. At the core, this situates European norms as the standard to which other cultures are being compared against [9].

Even though cultural factors may influence screening tools with dominant Western characteristics and, thus, can affect test performance, only few cognitive tests have considered this [10]. Furthermore, some argue that emphasis is placed on different cognitive skills within different cultures [11]. For example, the Western model of education typically exposes individuals to test conditions from a fairly young age [12] which could be likened to that of cognitive assessments. Where schooling is embedded within the culture, individuals are more likely to develop skills relating to finishing tasks within an allotted timeframe and being familiar with the demand for continuous attention. Therefore, individuals migrating from a developing country may possess potential sources of underperformance such as lower level of education, lack of familiarity with the language and cultural differences, such as culture-specific test items which assume familiarity to a different culture in which the population had been raised [13]. In these circumstances, such a testing process may raise a person's anxiety levels and, thus, not be an accurate reflection of an individual's true ability.

There has been growing evidence that cultural variables can exert a powerful effect on test performance [14] and, therefore, pose a challenge when interpreting scores in different cultural contexts. Even within the same country, ethnic minorities may also be influenced by cultural factors. One factor which is thought to have a potential impact on test performance is acculturation, defined as the process of cultural learning and incorporation of language, values, beliefs, and customs of a new country in which a person with an immigrant background resides [15, 16]. Some studies have identified associations between the degree of acculturation and the performance on cognitive tests designed in the country in which they reside, i.e., nativity status (place of birth and length of residence in the new country) and language acculturation (documented whether participants were bilingual or spoke primarily Spanish or English at home) [17], or quality of education (reading level) [18]. However, acculturation except language ability (understanding reading, writing and speaking, use and preference) comprises of different variables including the quality and the degree of formal education, and the pattern of abilities and values developed as a consequence of the cultural background [19], including social contacts (relations, friendships, way of dealing with people, communication style, family life), daily living habits, use of media, access to education, work, marriage, childrearing practices, celebrations and holidays, general knowledge and knowledge of world affairs, as well as specific cultural habits and customs [20]. It is, therefore, important to understand exactly what aspects of acculturation may impact performance.

Various screening tests have been validated through translated adaptations to increase diagnostic accuracy in culturally diverse populations. However, translation alone can be problematic as it does not address other matters concerning cultural differences [21]. Translation may also create issues when measuring aspects such as fluency, as it will not be testing the same factor if translated and could manifest as longer response times [22]. The impact of culture may affect the validity of cognitive screening tests used to identify dementia and MCI through biases

that may be related to the test itself, the administration and/or the interpretation of scores [8]. For example, where an item in a test may be ambiguous due to poor translation or lack of meaning in a particular cultural context, this may affect an individual's understanding, their interpretation [23] and, as a result, impact a test's diagnostic accuracy. This may be through different biases including that of construct (non-equivalent constructs across cultural groups) and item (inadequate translation from incorrect word choice) [24].

Universal cut-offs can also lead to sampling biases, whereby a cut-off may not match with the individuals being assessed [25] (e.g., if a target population has a lower level of education or higher rates of illiteracy). As such, this could lead to false conclusions made about an individual's cognitive ability which could have certain consequences, such as leading to inaccurate diagnosis. Additionally, this could also be problematic for the interpretation of scores by healthcare professionals which may impact clinical decision making.

Systematic reviews have explored cultural adaptations of specific screening tests including the Montreal Cognitive Assessment (MoCA) [26] and the Rowland Universal Dementia Assessment (RUDAS) [27] in specific populations e.g., literate versus illiterate [28], or specific to certain countries [29]. However, the translation and cultural adaptation of several dementia screening tests more broadly across a variety of populations has not been addressed. The aim of this paper is to explore the effects of translation and cultural adaptations in a range of cognitive screening tests developed in a Western context and investigate the impact this has on their diagnostic accuracy and test performance across diverse populations. This review aims to highlight considerations relating to the adaptation processes, language, cultural influences, impact of immigration, and level of education when using a range of different Western-designed cognitive screening tools originally in English to assess for dementia in non-Western and/or non-English speaking populations.

METHODS

Search strategy

Four databases were searched, including: APA PsycInfo, Medline, the Cumulative Index to Nursing and Allied Health Literature (CINAHL), and SCOPUS. Databases were searched from inception

Table 1
Concept and corresponding search terms table

Concept	Search terms
Individuals from culturally diverse populations**	ethnic* group* or ethnic* or rac* or cultur* or marginali*ed background* or minorit* or racial background* or cultural background* or racial differen* or ethnic minorit* or ethnic* divers*
Dementia	dementia* or alzheimer* or memory loss or cognitive impairment* or vascular dementia or frontotemporal
Cognitive screening tests	cognitive assessment* or cognitive tool* or cognitive test* or brief cognitive assessment* or cognitive screening* or cognitive screening test*

The Boolean operator 'AND' was used to combine search terms for the final search. *Truncation symbol used at the end of search terms to find any string of characters in that position, for example, ethnic* would identify ethnic, ethnicity, ethnicities etc. **The term 'culturally diverse populations' was used to refer to non-English speaking participants and/or those from non-Western countries for whom English was not their first language.

to 30 June 2023. Database searches were additionally supplemented with manual searches [i.e., The Clinical Neuropsychologist 2017–2023; Archives of Clinical Psychology 2020–2023; Neurological Sciences 2022; British Journal of Psychiatry Open 2022 (an online-only open access journal); Journal of International Neuropsychological Society 2022, etc.].

Development of search terms

The key search concepts for this review were identified through a focus on the population and intervention elements of the PICO framework [30]. The individual terms for each concept were guided by consideration of terms used in previous reviews within the field (Table 1).

Article selection summary

A total of 2,890 records were identified. Duplicates ($n = 801$) were removed using the reference management tool, Endnote ($n = 460$), as well as manually ($n = 341$) (NCC-A). The remaining records ($n = 2,089$) were screened based on title and abstract independently by three reviewers (NCC-A, TC, and EBM-L) against the inclusion and exclusion criteria (Table 2). The PRISMA diagram [31] outlines the search results and screening process, including the reasons for exclusion at the full-text screen (Fig. 1). The study is compliant with the PRISMA

Table 2

Inclusion and exclusion criteria. Cultural diversity was defined according to Lin (2020) as 'an open-ended term, which generally refers to a reality of coexistence of diverse knowledge, beliefs, arts, morals, laws, customs, religions, languages, abilities and disabilities, genders, ethnicities, races, nationalities, sexual orientations, etc.'

Inclusion criteria	Exclusion criteria
Sample involving participants assessed for the detection of dementia including Alzheimer disease, vascular dementia, frontotemporal dementia, Parkinson disease, rare dementias, or mild cognitive impairment in any clinical or community setting	Studies that were written in another language other than English
	Studies conducted on participants devoid of cognitive complaints only
	Qualitative studies, case studies, focus groups
Adults aged ≥60*	Studies that include behavioral or functional assessments only
Studies written in English	Studies involving only participants devoid of cognitive complaints
Cognitive assessment tools which have been culturally adapted or translated from English	Cognitive impairment or neuropsychological profile being assessed for/or associated with another condition/ co-morbidity (e.g., stroke, TIA, acquired brain injury, heart failure, HIV, diabetes, epilepsy etc.)
	Studies which use only screening tools designed to be "culturally fair"

*As based on mean age reported; for some studies 95% confidence interval (CI) was also calculated, to ensure that 95% of participants had the true mean of age ≥60 years. HIV, human immunodeficiency virus infection; TIA, transient ischemic attack.

2020 27-item checklist for reporting systematic reviews [31].

Data extraction and synthesis

Data was extracted on the following: population, setting, specification of diagnosis, comparison (dementia/MCI versus controls and cultural group), age, country, education level, adaptations (cultural/ translations versus original), cut-off scores, sensitivity, and specificity (Supplementary Material 1). The extraction of data and quality assessment was performed independently by two authors (extraction of data: NCC-A and TC, and TC and EBM-L; quality assessment NCC-A and EBM-L and EBM-L and HS). Disagreements were internally resolved between the reviewers. Due to limited studies identified, their heterogeneity, as based on dementia subtypes and clinical cognitive assessment tools used, a meta-analysis was precluded, and a narrative synthesis was conducted instead. Studies characteristics and key outcomes were summarized as tables and a figure. Data files/information were available for download from the journal website or a data repositories.

Sensitivity and specificity for optimal cut offs (as indicated by the authors) were extracted into Microsoft Excel, and confidence intervals at 95% were calculated for 14 studies. Sensitivity/specificity data was plotted onto two separate forest plots for the relevant category: differentiating MCI from CH and dementia from CH. Where sensitivity and specificity values were not used ($n = 3$), area under curve (AUC), correlation coefficients or other relevant data is presented. Where studies have included comparison between English-speaking and non-English speaking or immigrant population versus non-immigrant population groups, comparative data relating to test performance is also presented. Heterogeneity was accessed via visual examination of risk bias, descriptive data and forest plots. Factors which were thought to possibly contribute to heterogeneity related to co-variates including selection of patients, education, linguistic ability, and immigration status.

Quality appraisal

The quality of the eligible studies was assessed using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool (Table 1, Supplementary Material 2A and 2B) [32].

RESULTS

Study characteristics

Studies were published between 1999–2023 (Table 3). The studies recruited samples from

```
┌──────────────┐
│Identification│
└──────────────┘
```

Reports identified
through database and
manual searching
(n=2,890)

→ Duplicates removed
(*n*=801)

```
┌─────────┐
│Screening│
└─────────┘
```

Reports screened based
on title and abstract
(*n*=2,089)

→ Reports excluded based on title and
abstract (*n*=1,993)

Full text articles
assessed for eligibility
(n=96)

→ Full text articles excluded (*n*=79)

Reasons for exclusion:

Non dementia related group (n=35)
Specifically developed cultural fair
screening tool (n=9)
No reference standard (n=5)
Non ethnic minority group (n=5)
Systematic review (n=6)
Scoping review (n=1)
Narrative review (n=4)
Non peer review (n=3)
Qualitative study (n=3)
Not relevant (n=6)
Non-English (n=2)

```
┌────────┐
│Included│
└────────┘
```
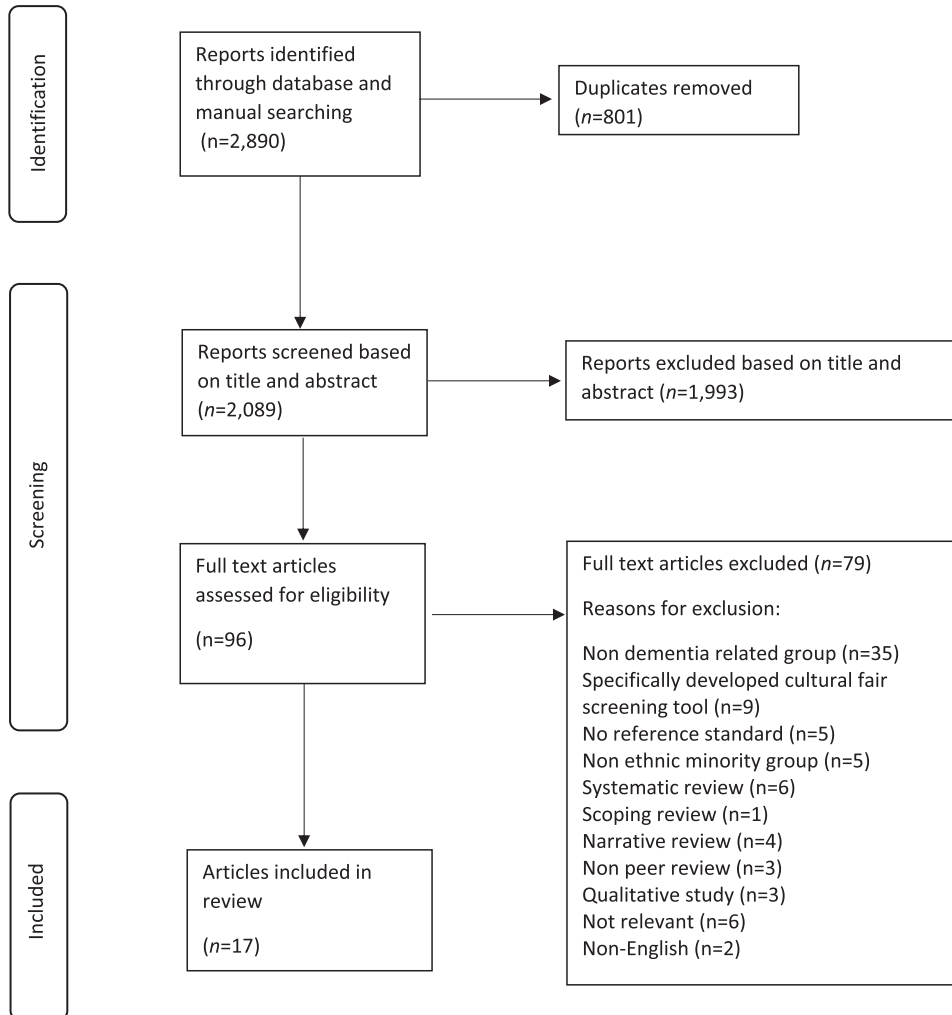
Articles included in
review
(*n*=17)

Fig. 1. PRISMA flowchart for systematic reviews with included searches of databases and identified reports.

clinical and community settings, and came from Middle East (Turkey $n = 2$, Lebanon $n = 1$), North America (Mexico $n = 1$, USA $n = 1$, Canada, $n = 1$), Europe (one study each from Greece, Belgium/Denmark/Norway/Sweden/Germany, Czech Republic, Netherlands, Germany, Sweden), one study each from South America (Argentina) and Southern/ Eastern Asia (India, Sri Lanka, Thailand and Japan). Most studies used the Diagnostic and Statistical Manual of Mental Disorders (DSM) as a reference standard ($n = 9$), followed by the NINCDS-ADRDA (National Institute of Neurological and Communicative Disorders and Stroke-Alzheimer's Disease and Related Disorders Association; $n = 6$). The most prevalent cognitive tools were RUDAS ($n = 5$) followed by MoCA ($n = 4$) (see Supplementary Material 3 for further study characteristics which specify cut-off scores for each study).

Participant characteristics

The total number of participants was 4,463, with sample sizes ranging from 56 [33] to 874 [34]. The overall total sample comprised of 55% females and 45% males, with a mean age 71.8 years. Mean years of education varied across studies from 7 [33] to 14.6 years [35], with few studies including participants with no education [10] or no formal education (52%) [36]. Where studies included further information on the level of education in populations with an immigrant background, a summary of this is presented in Supplementary Material 4.

Table 3
Summary of eligible studies

Authors	Location of study	Sample size (*n*)	Age (*M* ± SD))	Sex (Female %)	Education, y (M)	Cognitive status	Criteria	Screening tool(s) assessed
Aguilar-Navarro et al. (2019) [45]	Mexico	215	79 ± 7.3	59%	9.9	AD and MCI	DSM-V	5-WT (cultural adaptation; five nouns from the Spanish language were selected)
Bezdicek et al. (2020) [34]	Czech Republic/USA	874	71 ± 13.3	56%	14	PD and PD-MCI	U.K. PD Society Brain Bank criteria	MoCA-CZ (translation)
Borson et al. (1999) [44]	USA	295	73.7 ± 9.2	68%	11.4	Probable dementia	CERAD	CDT (Adaptation/translation not required)
Celik et al. (2022) [47]	Germany	65	72.1 ± 7.6	55%	8.7*	AD	NINCDS-ADRDA	RUDAS and MMSE (translation)
Chaaya et al. (2016) [36]	Lebanon	232	79.1 ± 8.1	71%	52% no formal education	Dementia (not specified)	DSM-IV	A-RUDAS (back translated) and A-MMSE (translation)
Franzen et al. (2019) [10]	Netherlands	87	67.5 ± 11.1	51%	Range (*y*) = 0–>primary education	Dementia (unspecified)	NINCDS-ADRDA NINDS-AIREN	VAT and mVAT (modified VAT using colored photographs instead of line drawings)
Karunaratne et al. (2011) [38]	Sri Lanka	98	64 ± 10.3	35%	10.4	AD	DSM-V	MoCA-S (back translation and cultural adaptation)
Kaya et al. (2016) [39]	Turkey	274	74 ± 7.2	62%	8.1	MCI and AD	NINCDS-ADRDA	SLUMS-T (back translation)
Lakshminarayanan et al. (2022) [41]	India	107	72.3 ± unknown	59%	11	Mild-moderate dementia	DSM-V	ADAS-Cog (back translation and cultural adaptation)
Manjavong et al. (2021) [25]	Thailand	150	68.4 ± unknown	52%	Range (*y*) = 6.5–15	MCI	DSM-V	RUDAS-Thai MoCA-Thai (translation)

Study	Country	N	Age (M ± SD)	% Female	Education (y)	Diagnosis	Diagnostic criteria	Tool (adaptation)
Mavioglu et al. (2006) [33]	Turkey	56	69.5 ± 7.4	45%	7	AD	DSM-IV; NINCDS-ADRDA	ADAS-Cog (translation and cultural adaptation)
Nielsen et al. (2019) [40]	Germany, Belgium, Sweden, Denmark, Norway, Greece	280	71 ± 9.4	57%	8	Dementia (various)	DSM-IV-TR	RUDAS (back translation)
Serrano et al. (2020) [50]	Argentina	399	73.4 ± 6.9	70%	10.3	MCI and dementia	DSM-V	MoCA-A (translation and cultural adaptation)
Statucka et al. (2021) [35]	USA	714	63.1 ± 6.6	33%	14.6*	PD-MCI	MDS PD-MCI Level II criteria	DRS-2 (translation)
Torkpoor et al, 2022 [43]	Sweden	125	71.6 ± 12.5	54%	9.2 ± 4.1 Range (y) = 1–20	No dementia (n = 62; (SCI, MCI, depression, PTSD etc.) dementia (n = 63; AD, VaD, Mixed (AD and VaD), DLB, FTD, PDD and nonspecific dementia)	ICD-10	MMSE-S, RUDAS-S, FAQ (translation)
Tsolaki et al. (2000) [37]	Greece	250	79 ± 3.8	46%	5.4	Mild-severe dementia (various types)	DSM-V; NINCDS-ADRDA	CAMCOG (back translation)
Yoshida et al. (2012) [42]	Japan	242	72 ± 8.3	64%	11.6	MCI and dementia	NINCDS-ADRDA	ACE-R-J (translation and modification so the test be linguistically and culturally equivalent to the original English version)

ACE-R, Addenbrooke's Cognitive Examination-Revised; ACE-R-J, Japanese version of Addenbrookes Cognitive Examination – Revised; AD, Alzheimer's disease; ADAS-Cog, Alzheimer's Disease Assessment Scale-Cognitive subscale; CAMCOG, Cambridge Cognitive Examination; CDT, Clock Drawing Test; CERAD, Consortium to Establish a Registry for Alzheimer's Disease; DRS, Mattis Dementia Rating Scale; DSM, Diagnostic and Statistical Manual of Mental Disorders; DLB, Dementia with Lewy Body; 5-WT, Five-Word Test; FDT, Frontotemporal lobe dementia; FAQ, Functional Assessment Questionnaire; M, Mean; MCI, Mild Cognitive Impairment; MMSE, Mini Mental State Examination; A-MMSE, MMSE in the Arabic language; mMMSE, Modified Mini Mental Status Examination; MMSE-S, MMSE in the Swedish language; MoCA, Montreal Cognitive Assessment; MoCA-CZ, Czech version of MoCA; NINCDS-ADRDA, National Institute of Neurological and Communicative Disorders and Stroke-Alzheimer's Disease and Related Disorders Association; PD, Parkinson's disease; PDD, Parkinson's Disease Dementia; PD-MCI, Parkinson's disease-Mild Cognitive Impairment; PTSD, Posttraumatic Stress Disorder; RUDAS, Rowland Universal Dementia Assessment Scale; A-RUDAS, RUDAS in the Arabic language; RUDAS-S, RUDAS in the Swedish language; SCI, Subjective Cognitive Impairment; SD, Standard Deviation; SLUMS, Saint Louis University Mental Status; VaD, Vascular Dementia; VAT, Visual Association Test; mVAT, modified Visual Association Test.

Cultural and translation adaptations

Translation process

Six studies [36–41] adopted the backwards translation process to translate the screening tests (Table 3). These studies demonstrated high sensitivity (84%–99%) [39, 42] and high specificity, from 80% [38] to 99% [42]. In contrast, studies that did not use back translation, demonstrated much higher heterogeneity in differentiating dementia and MCI from cognitively healthy individuals, with sensitivity and specificity as low as 65% [43] and 40% [35], respectively.

Some studies did not warrant a backwards translation process [10, 44]. For example, Borson et al. (1999) explored the utility of the Clock Drawing Test (CDT) in a multi-ethnic sample which involved minimal language requirements [44]. Instead, an interpreter specifically trained for the study administered the test instructions in the participants' native language. Results demonstrated high levels of sensitivity (94%) and specificity (85%) for differentiating between probable dementia and CH controls in low educated and non-English speaking participants. Similar high accuracy to differentiate people with dementia and MCI from cognitively intact individuals was also demonstrated with the Visual Association Test (VAT), irrespective of whether the six-line drawings of pairs of interacting objects (association cards) were presented in black and white or in color [10].

Cultural adaptations

Seven of the included studies made cultural adaptations to fit more specifically within a country's cultural norms. For example, the cultural adaptation of the five-word test (5WT) [45], was undertaken by five linguistic experts from Mexico. Five words from the Spanish language were chosen based on criteria to make the test comparable to the original: word length of two to three syllables and beginning with different letters of the alphabet. The understanding of the chosen words was then assessed by small groups of older adult volunteers from the memory service of the study site. The test was validated on volunteers from the site in which the study was conducted, which could increase risk of bias and lead to concerns regarding applicability to other regions within Mexico thus, impacting on generalizability. Results demonstrated sensitivity of 89% and specificity of 98% for differentiating between Alzheimer's disease

(AD) and CH. For MCI, sensitivity and specificity were 66% and 77%, respectively.

Other studies used a backwards translation process as well as making cultural adaptations. For example, following the translation process, Mavioglu et al. (2006) adapted words from the ADAS-Cog to fit better with Turkish culture, i.e., "ocean" was replaced with "sea" and "lobster" with "fish" [33]. With these modifications, the test differentiated between probable AD and CH ($p < 0.001$). In Lakshminarayanan et al. (2022) study [41], the ADAS-Cog "Word Recall' was replaced with the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) word list, used in other adaptation studies conducted in low and middle-income countries [46]. Likewise, the "Naming Objects and Fingers" item was substituted with culturally appropriate alternatives, with the authors providing an example of replacing "mask" (which they found difficult for participants to identify) with "battery" from the list of alternative words.

Franzen et al. (2019) adapted the original Visual Association Test (VAT) by changing the black and white line drawings to color photographs [10]. Both the original VAT and the modified version (mVAT) were administered to non-Western immigrants from a memory clinic setting and a control sample of CH individuals with an immigrant background. Although both tests discriminated well between individuals with and without dementia; participants performed higher when completing the mVAT (VAT, AUC = 0.77–0.88; mVAT, AUC = 0.85–0.95). This suggests that performance of non-Western immigrants could be underestimated when using black and white line drawings.

Immigrant populations

In addition to investigating the accuracy of cultural and translational adaptations of dementia cognitive screening tests, six studies explored screening tests in identifying dementia or MCI in those from an immigrant population. Three of the studies' findings suggested that both educational background and linguistic ability of those from a non-Western first-generation immigrant background need to be considered when characterizing cognitive profiles [35, 40, 47]. However, four of these studies did not translate or culturally adapt the respective cognitive screening tests and instead investigated diagnostic accuracy and test performance in relation to immigrant versus non-immigrant groups. For example, in Nielsen et al.'s study. (2019), 57% of partic-

ipants had an immigrant background [40]. The findings in this group demonstrated that RUDAS scores were mainly affected by level of education as opposed to immigrant status. When cut-off scores were adjusted to account for education, this slightly improved diagnostic accuracy (before adjustment, AUC of 0.93; after adjustment, AUC of 0.95). On the other hand, Torkpoor et al. (2022) study, based on a smaller group of immigrants, reported that educational background and immigrant status do not influence the RUDAS, but the MMSE cognitive performance [43]. In contrast, Celik et al. (2022) found no difference when adjustments were made to account for years of education in RUDAS scores ($p = 0.622$) but did find that years of education were significantly associated with total MMSE scores ($p < 0.001$), where German natives performed better [47]. Thus, it could be important to consider to what extent level of education may be correlated with immigrant background, i.e., some studies have found lower test performance scores to be associated with lower levels of education related to lack of opportunity available within lower economically developed countries (LEDCs) [48]. This is supported by Statucka et al. (2021) [35] who reported that immigrant status alone did not significantly impact results ($p = 0.560$) but instead was related to other variables around the Historical Index of Human Development ($p < 0.001$). This latter measure captures three dimensions of human development including a long and healthy life (life expectancy), access to knowledge (years of schooling) and decent standard of living (gross income, adjusted for price level of country) [49].

The impact of bilingualism in cognitive testing also appeared to be relevant. Thus, bilingual people performed significantly better on the MMSE as opposed to those who were monolingual ($p = 0.046$) [47]. However, when bilinguals were removed from the analyses, results demonstrated that Turkish (TR) and Turkish immigrant (TI) groups performed significantly worse than German participants (TR, $p = 0.021$; TI, $p = <0.001$).

Risk of bias across studies

A summary of the methodological quality using the QUADAS-2 is presented in Table 4.

The quality within some domains (patient selection, index test, reference standard and flow and timing) was sometimes difficult to assess due to information not being clearly reported within the studies.
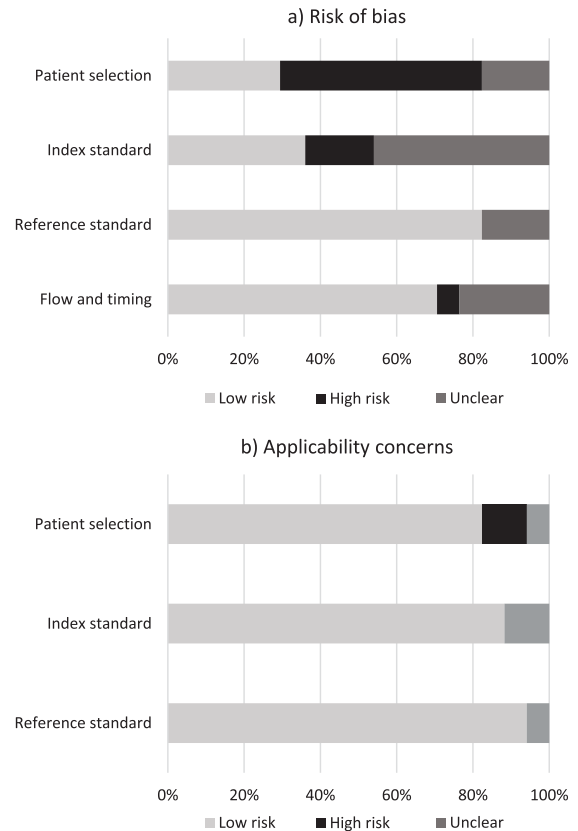

Fig. 2. Risk of bias and concerns about applicability. Proportion of studies with low, high, and unclear risk of bias (a) and applicability (b). Values represent percentage of studies ($n = 17$).

The ratings for each domain are presented as percentages in Fig. 2. The only study that demonstrated low bias on all domains was the one by Tsolaki et al. (2000) in which participants came from a memory clinic setting [37]. The authors included a heterogeneous group of participants from different cultural and socio-economic environments and extent of cognitive problems, ranging from subjective and functional cognitive problems to those with MCI and unselected dementias, so that they represented as many variables of the general population as possible in the analyzed sample. This was not the case for other studies that had a high level of bias in respect to patient selection, due to the nature of some of the case-control designs.

For index test, 7 studies were marked as 'unclear' due to the absence or lack of clarity around test administration. Three studies [38, 40, 43] were rated as being at 'high' risk of bias due to the assessors' knowledge of the participant's cognitive and/or

Table 4
Methodological quality of data based on the Quality Assessment of Diagnostic
Accuracy Studies (QUADAS-2) tool [32]

Study	Risk of bias				Concerns regarding applicability		
	Patient selection	Index test	Reference standard	Flow and Timing	Patient selection	Index test	Reference standard
Aguilar-Navarro et al. (2019) [45]	Unclear	Unclear	Low	Unclear	Low	Unclear	Low
Bezdicek et al. (2020) [34]	High	Unclear	Low	Low	High	Low	Low
Borson et al. (1999) [44]	Low	Low	Low	Low	Low	Low	Low
Celik et al. (2022) [47]	Unclear	Unclear	Unclear	Low	Low	Low	Low
Chaaya et al. (2016) [36]	Low	Unclear	Low	Low	Low	Low	Low
Franzen et al. (2019) [10]	Unclear	Low	Low	Low	Low	Low	Unclear
Karunaratne et al. (2011) [38]	High	High	Low	Unclear	Low	Low	Low
Kaya et al. (2016) [39]	High	Low	Low	Unclear	Unclear	Low	Low
Lakshminarayanan et al. (2022) [41]	High	Unclear	Low	Low	Low	Low	Low
Manjavong et al. (2021) [25]	Low	Low	Low	Low	High	Unclear	Low
Mavioglu et al. (2006) [33]	High	Low	Low	Low	Low	Low	Low
Nielsen et al. (2019) [40]	Low	High	Unclear	Low	Low	Low	Low
Serrano et al. (2020) [50]	High	Low	Low	Low	Low	Low	Low
Statucka et al. (2021) [35]	High	Low	Low	High	Low	Low	Low
Torkpoor et al. (2022) [43]	Low	High	Low	Unclear	Low	Low	Low
Tsolaki et al. (2000) [37]	Low	Low	Low	Low	Low	Low	Low
Yoshida et al. (2012) [42]	High	Unclear	Low	Unclear	Low	Low	Low

Key: ■ High   □ Low   ▨ Unclear

immigrant status and potential for overestimation of diagnostic accuracy.

For reference standard, all but one of the studies used a recognized reference standard that was likely to correctly classify the target condition. Over 80% of studies were rated as 'low' risk of bias and 'low' in terms of concerns about applicability. Two studies [40, 47] were marked as 'unclear' due to lack of clarity around the use of masking.

For flow and timing, four studies were marked as 'unclear' due to lack of information reported in relation to the time intervals between the index test and reference standard. One study [35] was rated 'high' in level of bias due to lack of justification around a time of approximately 12 months between reference standard and index test. Conditions may have worsened during this timeframe, thus contributing towards unreliable diagnostic accuracy data.

Mild cognitive impairment versus cognitive health studies

Seven studies ($n = 1,158$; Table 3) investigated the ability of the included cognitive screening tests [5WT, Czech version of the MoCA (MoCA-CZ), Turkish version of Saint Louis University Mental Status (SLUMS-T), RUDAS-Thai, Argentine version of the MoCA (MoCA-A), Dementia Rating Scale (DRS), Japanese version of Addenbrookes Cognitive Examination - Revised (ACE-R-J)] to differentiate MCI from CH. Sensitivity ranged from 66% (at a specificity of 77%) to 87% (at a specificity of 97%), whereas their specificity ranged from 40% (at a sensitivity of 82%) to 92% (at a sensitivity of 87%) (Fig. 3). Two of the included studies by Bezdicek et al. (2020) [34] and Statucka et al. (2021) [35] investigated the ability of two separate cognitive screening tests (MoCA-CZ and DRS-2) to differentiate between
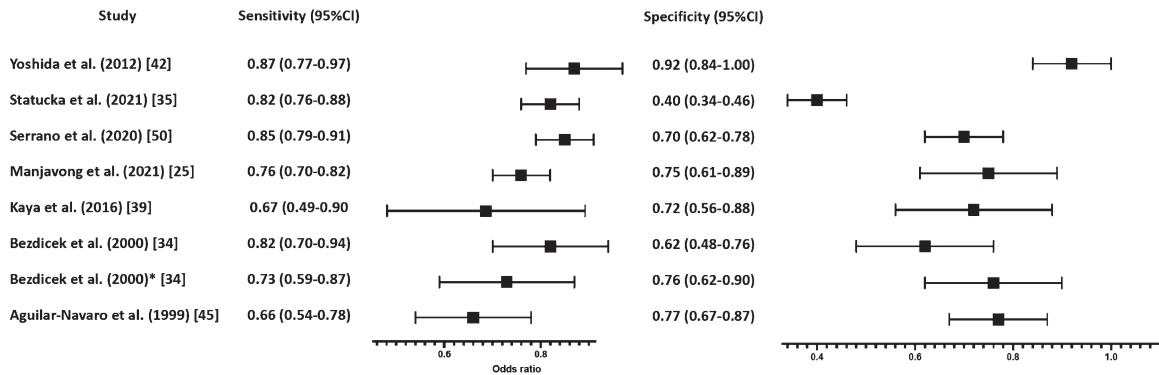
Study	Sensitivity (95%CI)		Specificity (95%CI)	
Yoshida et al. (2012) [42]	0.87 (0.77-0.97)		0.92 (0.84-1.00)	
Statucka et al. (2021) [35]	0.82 (0.76-0.88)		0.40 (0.34-0.46)	
Serrano et al. (2020) [50]	0.85 (0.79-0.91)		0.70 (0.62-0.78)	
Manjavong et al. (2021) [25]	0.76 (0.70-0.82)		0.75 (0.61-0.89)	
Kaya et al. (2016) [39]	0.67 (0.49-0.90)		0.72 (0.56-0.88)	
Bezdicek et al. (2000) [34]	0.82 (0.70-0.94)		0.62 (0.48-0.76)	
Bezdicek et al. (2000)* [34]	0.73 (0.59-0.87)		0.76 (0.62-0.90)	
Aguilar-Navaro et al. (1999) [45]	0.66 (0.54-0.78)		0.77 (0.67-0.87)	

Fig. 3. Forest plot differentiating between MCI and CH. *These studies differentiated between PD-MCI.

Parkinson's disease (PD) MCI (PD-MCI) and PD cognitively healthy (PD-CH) people. Bezdicek et al. (2020) explored the MoCA-CZ which had a sensitivity of 73% and specificity at 76% [34]. In comparison, the same study also differentiated PD-MCI from CH at a sensitivity of 82% and specificity of 62%, suggesting that the MoCA-CZ was more sensitive in differentiating between PD-MCI and CH, but had higher false positive rates when compared to the PD-MCI and PD-CH group.

These results are in contrast with Statucka et al. (2021) who reported DRS-2 to have a sensitivity of 82% and specificity of 40% when differentiating between the international PD-MCI versus PD-HC group [35]. However, as the two studies investigated different cognitive screening tests, one which was translated and culturally adapted and one which was not (MoCA-CZ versus DRS-2, respectively), direct comparisons between these two studies could not be made. Nevertheless, this study highlighted the importance of cultural and translated adaptations of a test which could possibly increase the overall diagnostic accuracy in culturally diverse groups leading to lower false positive rates.

Dementia versus CH

Eleven studies with 1,296 participants investigated the ability of cognitive screening tests (ACE-R-J, ADAS-Cog-Tamil, CAMCOG, CDT, MoCA-A, MoCA-S, RUDAS, A-RUDAS, RUDAS-S, SLUMS-T, and 5WT) to differentiate dementia from CH (Table 3). Dementia types included AD ($n = 6$), vascular dementia, frontotemporal dementia, mixed dementia, dementia with Lewy bodies (two studies each), PD dementia ($n = 3$) and non-specified demen-

tia type ($n = 6$). Sensitivity ranged from 80% (at a specificity of 90%; [40]) to 100% (at a specificity of 70%; [50]) whereas specificity ranged from 70% (at a sensitivity of 100%; [50]) to 99% (at a sensitivity of 99%; [42]) (Fig. 4).

Overall, the findings across studies demonstrated cognitive screening tests to have higher sensitivity and specificity in differentiating dementia versus CH subjects.

Heterogeneity

Some studies rated at 'high risk' of bias in patient selection had higher levels of accuracy reported than those of low risk and unclear ratings. For example, in the study conducted by Statucka et al. (2021), the DRS-2 showed high sensitivity (82%) but low specificity (40%) when differentiating between PD-MCI from PD-HC [35]. This may be due to the difference in disease severity within and between groups. For instance, both groups comprised of participants diagnosed with PD, with MCI being a common feature, and could have been associated with older age at assessment and depression amongst other factors [51], thus, potentially resulting in higher false positive rates within this population.

DISCUSSION

Dementia cognitive screening tests in older culturally diverse populations were widely variable across studies, especially in terms of cut-offs, with optimal cut offs based on authors' own data without external validation, and sensitivity and specificity scores. Findings were multi-faceted which may provide insight into different factors when considering
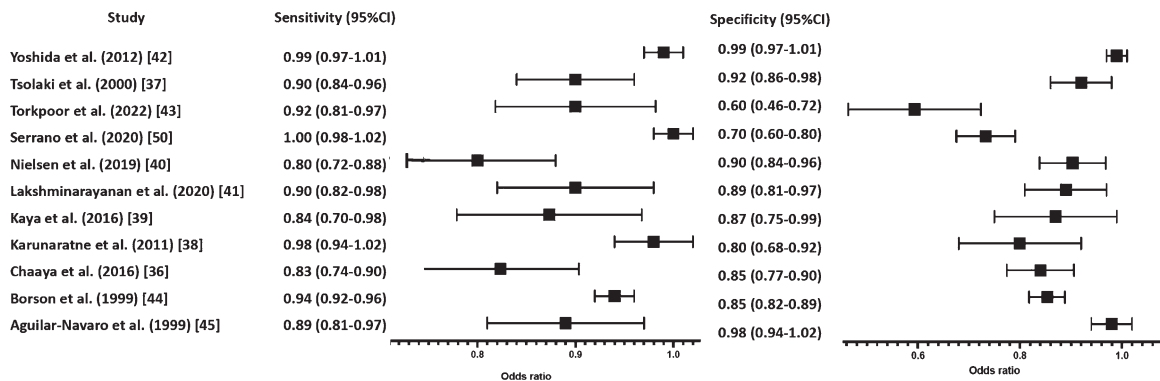
Fig. 4. Forest plot differentiating between dementia and CH.

the effect that both translation and cultural adaptations of tests have on diagnostic accuracy and test performance in diverse populations.

The studies included within this review varied in sample size, demographics, settings, education level, as well as variations in the population (e.g., severity of disease). This could have included inconsistency in the severity of symptoms, education level, linguistic ability, and the level of anxiety experienced in those completing tests in clinical settings compared to in participants' own homes within the community. Additionally, the specified cut-offs in studies were adjusted dependent on different factors within the study populations, such as years of education and/or severity of disease. This could have impacted on the diagnostic variability across studies and subsequently affected the validity and reliability, sensitivity, and specificity of the included cognitive screening tests in accurately identifying MCI and dementia.

General findings indicated that sensitivity and specificity was higher when differentiating between dementia and CH in comparison to MCI versus CH. Furthermore, depending on the cognitive tests, the differentiation between dementia and CH groups had the lowest sensitivity and specificity when translated-only cognitive tests were used [43], followed by back translation [36, 37, 39, 40], with the highest accuracy demonstrated on those tests that were based on both back translation and cultural adaptation [38, 41, 42, 50], or cultural adaptation only [45]. On the other hand, all the tests used to differentiate MCI from CH were based only on translation [25, 34, 35] or back translation [39], with only two studies addressing the cultural adaptation of the used cognitive tests, with variable success [42, 45]. One of the reasons why cognitive tests may have different accuracy in detecting dementia and MCI may lie not only in the quality

of the translated and adapted cognitive tools, but also their ability to aid the MCI diagnosis in culturally diverse clinical settings. Moreover, MCI itself is also heterogenous and an unstable construct which can lead to variable patterns and diagnosis rates.

Consideration should be taken when interpreting our findings as many of the samples in the included studies compared groups with confirmed diagnoses against participants who were deemed as "cognitively healthy". This could potentially over-estimated diagnostic accuracy as some studies have excluded participants who may have been considered "difficult to diagnose" or those with unconfirmed diagnoses, suggesting that in clinical settings the sensitivity and specificity of these cognitive tests may be lower. Thus, all these studies, except for Tsolaki et al. (2020) study [37] that used CAMCOG, did not reflect a real clinical setting with a whole spectrum of memory problems, including functional cognitive impairment (i.e., memory problems but no dementia). Indeed, if the latter were included, the sensitivity and specificity of cognitive tests (RUDAS) was substantially lower (0.717 and 0.583, respectively) with accuracy of <70% [52]. This raises the question as to whether differentiating between those with a confirmed diagnosis and those deemed as 'cognitively healthy' could be overestimating diagnostic accuracy in typical clinical scenarios, and, therefore, introduce an element of bias by not testing an overall sample including those with 'unconfirmed' diagnoses. Additionally, healthcare professionals need to be mindful of high false positive rates as this could lead not only to further testing in those that may not necessarily require it but also impacts patients' daily life, in terms of driving, work, insurance, anxiety etc.

Cognitive screening tests, detailed above, can be biased against those from a non-English speak-

ing and/or non-Western country. This is particularly apparent in studies which include tests that have not been translated or culturally adapted [35, 40, 47]. Whilst cognitive processes are thought to be generally universal, a person's culture can impact on this which can also extend to cognitive test performance [53]. Factors concerning socioenvironmental context were thought to have contributed to poorer performance in those from a culturally diverse population when compared to English speaking or Western populations. Characteristics including country of origin, level of education and economic status of a country, are variables thought to have led to discrepancy amongst test performance. It is reported that those from LEDCs have less opportunity to receive higher quality education [54] and with cognitive tests being predominantly Eurocentric in their approach, culturally diverse populations may be at an increased disadvantage.

There were mixed findings relating to the effect of education across studies. Some findings reported that adjusted cut-off scores to account for education were helpful in improving diagnostic accuracy [40], whereas other reported no difference when scores were adjusted for the same test [47]. However, it appeared that higher performance on some tests (such as the MMSE) were significantly associated with years of education, and this seems to be mediated by the Historical Index of Human Development of the participants' country of birth (which reflects economic, health, and educational potential of a country at the time of birth) [8]. In addition, the quality of the cultural adaptation of the cognitive tests plays a role. Namely, the test should not only account for the language and culture of the person being assessed but should also maintain the integrity of the concepts being assessed [55]. It is necessary to bear in mind that some tests will be more influenced by education than others (e.g., MMSE versus RUDAS) due to the construct of the tests and certain items that may be more problematic for diverse populations such as verbal fluency tasks in those with lower levels of literacy.

Cognitive tests which involved minimal language requirements were more acceptable in those from a non-English speaking and/or non-Western country and demonstrated good levels of sensitivity and specificity. Tests including the CDT [44] and the mVAT [10] demonstrated promising results in differentiating between those with either dementia or MCI and control groups. Higher scores on the mVAT were thought to be associated with the added information that the

colored photograph provided, as opposed to decoding black and white line drawings which is thought to be a skill acquired through education [56]. This is important to note, as many cognitive screening tests use visuo-constructional tasks (e.g., the copying of geometric shapes as seen in the ACE) which are reportedly the least culturally adapted items in cognitive screening tests as it is thought that visuospatial ability is not strongly associated with culture [21]. However, the results from Franzen et al. (2019) did not seem to be consistent with this theory [10].

One of the limitations of this study is the lack of more detailed information about the acculturation variables we have analyzed. Thus, although the educational level (i.e., formal level of education) was noted in all studies, there was neither explicit information about the language ability nor social values of participants in the analyzed studies. Furthermore, the reviewed literature does not specify the model of acculturation they have used, or participants' acculturation groups/stages. Studies largely concentrated on the separation stage/approach (high origin-culture affiliation, with low new-culture affiliation), based on the account the language proficiency is necessary for completion of cognitive testing. A previous study conducted in the USA [57] specifically addressed and investigated the relationship between language and acculturation and identified specific relevant variables. However, according to these authors, these acculturation variables were only applicable to foreign-born individuals and/or those who spoke English as a second language. In their empirical study, Brauer Boone et al. (2007) [57] extracted the following language and acculturation data from the patient files: 1) whether subjects learned English as a first language (or concurrent with another language) versus English learned as a second language (ESL); 2) age at which English was first learned; 3) number of years resided in the United States (subtracted from total age); and finally, 4) number of years educated in the United States (subtracted from total years of education completed). These data, sadly, were not available for the migrant population in the studies we included in this review to make comparison.

Despite some evidence which implies that higher degrees of acculturation could positively impact test performance [58], the findings from this review did not come to the same conclusion. On the contrary, the level of acculturation was not related to total scores in one study [47]. However, this was the only study within the review which formally assessed the impact of acculturation using the Frankfurt Acculturation

Scale (FRAKK) [59], and therefore warrants further investigation. Furthermore, acculturation is said to comprise of several variables (including language), and bilingualism was reported to have had a positive effect on test performance in some studies. Therefore, it would be helpful to understand what aspects of acculturation aid test performance.

The translation process aims to provide a means of quality control to ensure the translation possesses the same meaning when moved back to the original source language [60]. The importance of a good translation process can help increase the accuracy of data derived from a translated measure. Out of the six studies which adopted a backwards translation process, two involved medical professionals including neurologists, psychologists, and psychiatrists [33, 41] alongside linguistic experts. Involving qualified professionals is ideal from a clinical and psychometric perspective as a native speaker may not necessarily have knowledge of cognitive tests and this could result in an inaccurate translation or misinterpretation, leading to construct bias. The other four studies used native linguists and/or bilinguals to translate the studies. Three studies involved patients and controls or caregivers to aid with the adaptation and check for cultural appropriateness which is helpful to increase face validity. Face validity is defined as whether test items are appropriate and relevant to a specific target population [61]. Therefore, the target population themselves are said to be best placed to determine the applicability of an adapted test which could help to increase this [62]. A recent study conducted in Maori (*Māori*; indigenous peoples of Aotearoa New Zealand) stressed the importance of cultural values to both improve the understanding of dementia and inform the cognitive assessments for those with dementia [63]. Interestingly, the authors highlighted that for those without cognitive deficit, a different framework is in place, arguing for more indebt understanding of distinct cultural aspects of physiological ageing and those related to dementia that may help inform both design of novel and adaptation of already available cognitive tests.

To conclude, this review highlights the importance of translating and culturally adapting tests that have been developed in the Western context and standardized against Western normative data, as it may improve diagnostic accuracy within diverse populations. Our review highlights consideration around certain variables that clinicians should be mindful of when interpreting individual test scores, including education, linguistic ability, and aspects of acculturation. However, caution should be taken when interpreting the findings of this review as data is varied due to the broad selection of cognitive tests that are included. This makes it difficult to compare sensitivity and specificity data across different studies and populations. Pooled sensitivity and specificity data could provide a more comprehensive assessment of the diagnostic accuracy across studies.

CONFLICT OF INTEREST

The authors have no conflict of interest to report.

DATA AVAILABILITY

The data supporting the findings of this study are available within the article and/or its supplementary material.

SUPPLEMENTARY MATERIAL

The supplementary material is available in the electronic version of this article: https://dx.doi.org/10.3233/ADR-230198.

REFERENCES

[1] Czaika M, de Haas H (2014) The globalization of migration: Has the world become more migratory? *Int Migr Rev* **48**, 283-323.

[2] Office for National Statistics: 2021 Census. https://www.ons.gov.uk/census. Last updated 2021, Accessed December 3, 2023.

[3] Tuerk R, Sauer J (2015) Dementia in a black and minority ethnic population: Characteristics of presentation to an inner London memory service. *BJPsych Bull* **39**, 162-166.

[4] Shiekh S, Cadogan S, Lin L-Y, Mathur R, Smeeth L, Warren-Gash C (2021) Ethnic differences in dementia risk: A systematic review and meta-analysis. *J Alzheimers Dis* **80**, 337-355.

[5] Ranson JM, Kuźma E, Hamilton W, Muniz-Terrera G, Langa KM, Llewellyn DJ (2019) Predictors of dementia misclassification when using brief cognitive assessments. *Neurol Clin Pract* **9**, 109-117.

[6] Alarcón RD (2009) Culture, cultural factors and psychiatric diagnosis: Review and projections. *World J Psychiatry* **8**, 131-139.

[7] Snowden, LR (2003) Bias in mental health assessment and intervention: Theory and evidence. *Am J Public Health* **93**, 239-243.

[8] Statucka M, Cohn M (2019) Origins Matter: Culture impacts cognitive testing in Parkinson's disease. *Front Hum Neurosci* **13**, 269.

[9] Helms JE (1989) Eurocentricism strikes in strange ways and in unusual places. *Counsg Psycholt* **17**, 643-647.

[10] Franzen S, van den Berg E, Kalkisim Y, van de Wiel L, Harkes M, van Bruchem-Visser RL, de Jong FJ, Jiskoot LC, Papma JM (2019) Assessment of visual association memory in low-educated, non-western immigrants with the modified visual association test. Dement Geriatr Cogn Disord **47**, 345-354.

[11] Kuwabara M, Smith LB (2012) Cross-cultural differences in cognitive development: Attention to relations and objects. *J Exp Child Psychol* **113**, 20-35.

[12] McCaffery J (2016) Western education: Suitable for everyone? Education for travelling, marginalized, and indigenous communities in the West. *Prospects* **46**, 491-505.

[13] Parker C, Philp I (2004) Screening for cognitive impairment among older people in black and minority ethnic groups. *Age Ageing* **33**, 447-452.

[14] Fernández AL, Abe J (2018) Bias in cross-cultural neuropsychological testing: Problems and possible solutions. *Cult Brain* **6**, 1-35.

[15] McDermott-Levy R (2009) Acculturation. *Holist Nurs Pract* **23**, 282-288.

[16] Zea MC (2004) Acculturation. In *Encyclopedia of Women's Health*. Loue S, Sajatovic M, eds. Springer, Boston, MA.

[17] Alam RB, Singleton CR, Aguiñaga S, Chodzko-Zajko W, Jahan NA, Oke A, Schwingel A (2022) Is acculturation associated with the cognitive performance of older Hispanics? *J Alzheimers Dis* **85**, 535–544.

[18] Manly JJ, Byrd DA, Touradji P, Stern Y (2004) Acculturation, reading level, and neuropsychological test performance among African American elders. *Appl Neuropsychol* **11**, 37-46.

[19] Ardila A (2020) Cross-cultural neuropsychology: History and prospects. *RUDN J Psychol Pedagogics* **17**, 64–78.

[20] Arends-Tóth J, van de Vijver FJR (2012) Assessment of psychological acculturation. In *The Cambridge Handbook of Acculturation Psychology*. Sam DL, Berry JW, eds. Cambridge University Press.

[21] Khan G, Mirza N, Waheed W (2022) Developing guidelines for the translation and cultural adaptation of the Montreal Cognitive Assessment: Scoping review and qualitative synthesis. *BJPsych Open* **8**, e21.

[22] Saz O, Lin Y, Eskenazi M (2015) Measuring the impact of translation on the accuracy and fluency of vocabulary acquisition of English. *Comput Speech Lang* **31**, 49-64.

[23] Schwieter JW, Prior A (2020). Translation ambiguity. In Bilingual lexical ambiguity resolution. Heredia RR, Cieślicka AB, eds Cambridge University Press.

[24] van de Vijver F, Hambleton RK (1996) Translating tests: Some practical guidelines. *Eur Psychol* **1**, 89-99.

[25] Manjavong M, Limpawattana P, Sawanyawisuth K (2021) Can RUDAS be an alternate test for detecting Mild Cognitive Impairment in older adults, Thailand? *Geriatrics (Basel)* **6**, 16.

[26] O'Driscoll C, Shaikh M (2017) Cross-cultural applicability of the Montreal Cognitive Assessment (MoCA): A systematic review. *J Alzheimers Dis* **58**, 789-801.

[27] Nielsen TR, Jorgensen K (2020) Cross-cultural dementia screening using the Rowland Universal Dementia Assessment Scale: A systematic review and meta-analysis. *Int Psychogeriatr* **32**, 1031-1044.

[28] Paddick SM, Gray WK, McGuire J, Richardson J, Dotchin C, Walker RW (2017) Cognitive screening tools for identification of dementia in illiterate and low-educated older adults, a systematic review and meta-analysis. *Int Psychogeriatr* **29**, 897-929.

[29] Santiago-Bravo G, Sudo FK, Assunção N, Drummond C, Mattos P (2019) Dementia screening in Brazil: A systematic review of normative data for the mini-mental state examination. *Clinics (Sao Paulo)* **74**, e971.

[30] Richardson WS, Wilson MC, Nishikawa J, Hayward RS (1995) The well-built clinical question: A key to evidence-based decisions. *ACP J Club* **123**, A12-A13.

[31] Page MJ, Moher D, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R, Glanville J, Grimshaw JM, Hróbjartsson A, Lalu MM, Li T, Loder EW, Mayo-Wilson E, McDonald S, McGuinness LA, Stewart LA, Thomas J, Tricco AC, Welch VA, Whiting P, McKenzie JE (2021) PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ* **372**, n160.

[32] Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MMG, Sterne JAC, Bossuyt PMM (2011) QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* **155**, 529-536.

[33] Mavioglu H, Gedizlioglu M, Akyel S, Aslaner T, Eser E (2006) The validity and reliability of the Turkish version of Alzheimer's Disease Assessment Scale-Cognitive Subscale (ADAS-Cog) in patients with mild and moderate Alzheimer's disease and normal subjects. *Int J Geriatr Psychiatry* **21**, 259-265.

[34] Bezdicek O, Cervenkova M, Moore TM, Stepankova Georgi H, Sulc Z, Wolk DA, Weintraub DA, Moberg PJ, Jech R, Kopecek M, Roalf DR (2020) Determining a Short Form Montreal Cognitive Assessment (s-MoCA) Czech Version: Validity in mild cognitive impairment Parkinson's fisease and cross-cultural comparison. *Assessment* **27**, 1960-1970.

[35] Statucka M, Cherian K, Fasano A, Munhoz RP, Cohn M (2021) Multiculturalism: A challenge for cognitive screeners in Parkinson's disease. *Mov Disord Clin Pract* **8**, 733-742.

[36] Chaaya M, Phung TKT, El Asmar K, Atweh S, Ghusn H, Khoury RM, Prince M, Nielsen TR, Waldemar G (2016) Validation of the Arabic Rowland Universal Dementia Assessment Scale (A-RUDAS) in elderly with mild and moderate dementia. *Aging Ment Health* **20**, 880-887.

[37] Tsolaki M, Fountoulakis KN, Chantzi H, Kazis A (2000) The Cambridge Cognitive Examination (CAMCOG): A validation study in outpatients suffering from dementia and non-demented elderly subjects (including age associated cognitive decline patients) in Greece. Am J Alzheimers Dis Other Demen **15**, 269-276.

[38] Karunaratne S, Hanwella R, de Silva V (2011) Validation of the Sinhala version of the Montreal Cognitive Assessment in screening for dementia. *Ceylon Med J* **56**, 147-153.

[39] Kaya D, Isik AT, Usarel C, Soysal P, Ellidokuz H, Grossberg GT (2016) The Saint Louis University Mental Status Examination Is better than the Mini-Mental State Examination to determine the cognitive impairment in Turkish elderly people. *Am Med Dir Assoc* **17**, 370.e11-5.

[40] Nielsen TR, Segers K, Vanderaspoilden V, Bekkhus-Wetterberg P, Bjorklof G H, Beinhoff U, Minthon L, Pissiota A, Tsolaki M, Gkioka M, Waldemar G (2019) Validation of the Rowland Universal Dementia Assessment Scale (RUDAS) in a multicultural sample across five Western European countries: Diagnostic accuracy and normative data. *Intl Psychogeriatr* **31**, 287-296.

[41] Lakshminarayanan M, Vaitheswaran S, Srinivasan N, Nagarajan G, Ganesh A, Shaji KS, Cahndra M, Krishna M, Spector A (2022) Cultural adaptation of Alzheimer's disease assessment scale–cognitive subscale for use in India and validation of the Tamil version for South Indian population. *Aging Ment Health* **26**, 423-430.

[42] Yoshida H, Terada S, Honda H, Kishimoto Y, Takeda N, Oshima E, Hirayama K, Yokota O, Uchitomi Y (2012) Validation of the revised Addenbrooke's Cognitive Examination (ACE-R) for detecting mild cognitive impairment and dementia in a Japanese population. *Int Psychogeriatr* **24**, 28-37.

[43] Torkpoor R, Kristin Frolich K, Nielsen TR, Londos E (2022) Diagnostic accuracy of the Swedish version of the Rowland Universal Dementia Assessment Scale (RUDAS-S) for multicultural cognitive screening in Swedish memory clinics. *J Alzheimers Dis* **89**, 865-876.

[44] Borson, S, Brush M, Gil E, Scanlan J, Vitaliano P, Chen J, Cashman J, Sta Maria M.M, Barnhart R, Rogues J (1999) The Clock Drawing Test: Utility for dementia detection in multiethnic elders. *J Gerontol A Biol Sci Med Sci* **54**, M534-M540.

[45] Aguilar-Navarro SG, Mimenza-Alvarado AJ, Martínez-Sánchez R, Mejía-Arango S, Gutiérrez-Gutiérrez L, Ávila-Funes JA (2019) Spanish validation and cultural adaptation of the Five-Word Test for the detection of the neurocognitive disorder in older adults. *Salud Ment* **42**, 281-287.

[46] Paddick S-M, Kisoli A, Mkenda S, Mbowe G, Gray WK, Dotchin C, Ogunniyi A, Kisima J, Olakehinde O, Mushi D, Walker RW (2017) Adaptation and validation of the Alzheimer's Disease Assessment Scale - Cognitive (ADAS-Cog) in a low-literacy setting in sub-Saharan Africa. *Acta Neuropsychiatr* **29**, 244–251.

[47] Celik S, Onur O, Yener G, Kessler J, Ozbek Y, Meyer P, Frölich ML, Teichmann B (2022) Cross-cultural comparison of MMSE and RUDAS in German and Turkish patients with Alzheimer's disease. *Neuropsychology* **36**, 195-205.

[48] Stephan BCM, Pakpahan E, Siervo M, Licher S, Muniz-Terrera G, Mohan D, Acosta D, Llibre-Rodriguez JJ, Prince M, Robinson L, Prina M (2020) Prediction of dementia risk in low-income and middle-income countries (the 10/66 Study): An independent external validation of existing models. *Lancet Glob Health* **8**, e524-e535.

[49] Prados de la Escosura L (2015) World human development: 1870–2007. *Rev Income Wealth* **61**, 220-247.

[50] Serrano CM, Sorbara M, Minond A, Finlay JB, Arizaga RL, Iturry M, Martinez P, Heinemann G, Gagliardi C, Serra A, Magliano FC, Yacovino D, Esnola y Rojas MM, Ruiz AS, Graviotto HG (2020) Validation of the Argentine version of the Montreal Cognitive Assessment Test (MOCA): A screening tool for mild cognitive impairment and mild dementia in elderly. *Dement Neuropsychol* **14**, 145-152.

[51] Aarsland D, Bronnick K, Williams-Gray C, Weintraub D, Marder K, Kulisevsky J, Burn D, Barone P, Pagonabarraga J, Allcock L, Santangelo G, Foltynie T, Janvin C, Larsen JP, Barker RA, Emre M (2010) Mild cognitive impairment in Parkinson disease: A multicenter pooled analysis. *Neurology* **75**, 1062-1069.

[52] Mukaetova-Ladinska EB, Abdullah S, Critchfield M, Maltby J (2022) Suspected dementia in young adults: Cognitive screening tools for use in primary care. *J Alzheimers Dis* **86**, 333-341.

[53] Rivera Mindt M, Byrd D, Saez P, Manly J (2010) Increasing culturally competent neuropsychological services for ethnic minority populations: A call to action. *Clin Neuropsychol* **24**, 429-453.

[54] Schmand B, Lindeboom J, Hooijer C, Jonker C. (1995) Relation between education and dementia: The role of test bias revisited. *J Neurol Neurosurg Psychiatry* **59**, 170–174.

[55] Mirza N, Wali Waheed M, Waheed W (2023) A new tool for assessing the cultural adaptation of cognitive tests: Demonstrating the utility of the Manchester Translation Evaluation Checklist (MTEC) through the Mini-Mental State Examination Urdu. *BJPsych Open* **9**, e5.

[56] Reis A, Faísca L, Ingvar M, Petersson KM (2006) Color makes a difference: Two-dimensional object naming in literate and illiterate subjects. *Brain Cogn* **60**, 49-54.

[57] Brauer Boone K, Victor TL, Wen J, Razani J, Pontón M (2007) The association between neuropsychological scores and ethnicity, language, and acculturation variables in a large patient population. *Arch Clin Neuropsychol* **22**, 355-365.

[58] Razani J, Burciaga J, Madore M, Wong J (2007) Effects of acculturation on tests of attention and information processing in an ethnically diverse group. *Arch Clin Neuropsychol* **22**, 333-341.

[59] Bongard S, Kelava A, Sabic M, Aazami-Gilan D, Kim Y-B (2007) Akkulturation und gesundheitliche Beschwerden bei drei Migrantenstichproben in Deutschland. *Beiträge Gesundheits Psychol* **29**, 53.

[60] Colina S, Marrone N, Ingram M, Sánchez D (2017) Translation quality assessment in health research: A functionalist

alternative to back-translation. *Eval Health Prof* **40**, 267-293.

[61] Thomas SD, Hathaway DK, Arheart KL (1992) Face validity. *West J Nurs Res* **14**, 109-112.

[62] Dingwall KM, Gray AO, McCarthy AR, Delima JF, Bowden SC (2017) Exploring the reliability and acceptability of cognitive tests for Indigenous Australians: A pilot study. *BMC Psychol* **5**, 26.

[63] Menzies O, Dudley M, Garrett N, Elder H, Daniels P, Wilson D (2022) *He Tūhononga Whaiaro*: A Kaupapa Māori approach to *Mate Wareware* (Dementia) and cognitive assessment of older Māori. *J Appl Gerontol* **41**, 1066-1073.