

Validity of Cognitive Assessment Tools for Older Adult Hispanics: A Systematic Review

Sandra P. Arévalo, PhD,^{*†}  Jennifer Kress, PhD,^{†‡} and Francisca S. Rodriguez, PhD^{†§||} 

OBJECTIVES: A higher prevalence and incidence of dementia is found in Hispanic/Latino older adults. Therefore, valid instruments are necessary to assess cognitive functioning in this population group. Our aim was to review existing articles that have examined and reported on the validity of cognitive assessment tools in Hispanic/Latino population groups in the United States.

DESIGN: Systematic literature review according to the Preferred Reporting Items for Systematic Reviews and Meta-analysis.

MEASUREMENTS: We systematically searched in the PubMed and Web of Science databases and assessed the quality of the search results using the Standards for the Reporting of Diagnostic Accuracy Studies. We included evidence from within the United States as well as from Spanish-speaking countries of origin (Mexico, Central and South America, and the Caribbean).

RESULTS: The literature search revealed 27 studies with adequate quality that investigated 13 instruments. The Mini-Mental Status Examination (MMSE) was the most frequently investigated instrument in Hispanic/Latino groups in the United States with high sensitivity for dementia but also with significant differences for ethnicity and education. The Addenbrooke Cognitive Examination-Revised, Montreal Cognitive Assessment, 10/66 short diagnostic schedule, clock-drawing test, Phototest, Eurotest, and Executive Battery 25 had good diagnostic performance in Spanish-speaking

countries. The naming test and verbal fluency tests have a higher risk of misclassifying US Hispanics/Latinos who have dementia.

CONCLUSION: Evidence on validity suggests that the MMSE may be an appropriate cognitive assessment tool for Hispanics. More research is needed to confirm the validity of cognitive tools to assess Hispanic/Latino groups for Alzheimer's disease and other related dementias in the United States to reduce current trends of culturally biased under- or overdiagnosis of cognitive impairments. *J Am Geriatr Soc* 68:882-888, 2020.

Key words: Latinos; Hispanics; Alzheimer's disease; cognitive functioning; assessment

Alzheimer's disease and related dementias are leading causes of morbidity and disability among older adults.¹ Prevalence and incidence of dementia is higher among diverse racial/ethnic groups in the United States.^{2,3} With the aging of society, the number of people with dementia will rise accordingly. For the Hispanic/Latino (H/L) population, the percentage of people older than 65 years is projected to increase from 6.7% today to 18.1% in 2060.⁴ This trend adds to the already higher prevalence of dementia in this population group. In addition, H/Ls have a longer median survival with dementia than whites,⁵ adding to the burden due to dementia and highlighting the relevance of appropriate assessment tools and treatment plans for this population group.

Substantial challenges lie ahead for finding appropriate solutions. One challenge is that many H/Ls have a low socioeconomic status⁶ that comes with a lower use of healthcare services.⁷ Barriers for medical care use include lack of transportation and disablement⁸ but also language barriers and cultural uncertainties about the appropriateness of care.^{9,10} Further, the understanding of clinical diagnostic criteria can differ depending on the cultural context in the interchange between healthcare providers, patients, and translators. In addition, translators such as family members communicate symptoms to

From the *California State University, Long Beach, Department of Human Development, Long Beach, California; †USC Edward R. Roybal Institute on Aging, University of Southern California, Los Angeles, California; ‡Darmstadt University of Applied Science, Darmstadt, Germany; §German Center for Neurodegenerative Diseases (DZNE), Greifswald, Germany; ¶Center for Cognitive Science, University of Kaiserslautern, Kaiserslautern, Germany; and the ||Institute of Social Medicine, Occupational Health and Public Health (ISAP), University of Leipzig, Leipzig, Germany.

Address correspondence to Francisca S. Rodriguez, PhD, German Center for Neurodegenerative Diseases (DZNE), Research Group "Psychosocial Epidemiology and Public Health," Ellernholzstr. 1-2, 17498 Greifswald, Germany. E-mail: francisca-saveria.rodriguez@dzne.de

Previous presentation: This article was presented at the annual scientific meeting of the Gerontological Society of America 2018.

[Correction added on 23 October 2020, after first online publication: Projekt Deal funding statement has been added.]

DOI: 10.1111/jgs.16300

the doctor that may not yield the most accurate fidelity of information exchange or interpretation of meaning with medical specialists in standard care settings.

Moreover, symptom presentation seems to be different in H/L groups. Once H/Ls appear in the clinical setting, they tend to present more severe clinical symptoms with a diverse symptom presentation and a concurrent presentation of other forms of dementia that in turn may affect diagnostic validity by reducing the sensitivity and specificity of available psychometric diagnostic tools.¹¹ Behavioral psychiatric symptoms, such as depression and anxiety, appear to be more prominent in H/Ls with dementia than in other ethnic groups.^{12,13} H/L patients also seem to express lower rates of apathy than whites.¹⁴ Further, studies observed that the level of performance in verbal conceptual thinking and visual attention is not always comparable between H/L and other ethnic groups.¹⁵ Accordingly, the development of valid diagnostic tools to screen and identify Latino older persons with Alzheimer's disease and related dementia at earlier stages is a great challenge. Current cognitive assessment tools in the United States are designed around the needs of native English-speaking and white populations, and the availability of assessments for the H/L population has been lagging behind.

This article review existing reports that have examined and reported on the validity of assessment tools examining cognitive functioning in H/L population groups in the United States, that is, individuals of Spanish-speaking Latin American ancestry according to the US Census categorization.¹⁶ To get a broader sense of available assessment instruments for this population group, we included evidence from Spanish-speaking population groups in Mexico, Central and South America, and the Caribbean.

METHODS

This systematic review was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-analysis, a set of items for the evaluation and reporting of studies to ensure a transparent and complete reporting of results developed by an international group of epidemiologists, clinicians, statisticians, and editors.

Inclusion and Exclusion Criteria

Criteria for study selection were determined using the Population, Intervention, Comparison, Outcome (PICO) model.¹⁷

Population

We considered study samples from the general population aged 45 years and older. The rationale for this age limit was our goal to focus on age-related cognitive impairments without risking the exclusion of an important study because it included younger participants. To specify age in the literature search, we adapted the terms suggested by Kastner and colleagues¹⁸ (search terms listed in Supplementary Table S1). Studies in the community as well as the clinical setting were considered relevant if they included H/L groups in the United States and/or Spanish-speaking groups from other countries not including Spain. We defined Hispanic/Latino populations according to

the classification of Hispanic or Latino used by the US Census that refers to a person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin regardless of race.¹⁹ Because European Spaniards have a different culture that may not be comparable with H/L groups living in the United States, we excluded studies reporting on cognitive assessments of older people residing in Spain.

Intervention or Prognostic Factor

Search terms on cognitive functioning, cognitive status, or Alzheimer's and age-related cognitive disorders were included (Supplementary File 1, Table S1). We focused on studies using instruments that assessed cognitive functioning itself and did not include behavioral rating scales, such as the Neuropsychiatry Inventory, or functional rating scales, such as the Functional Assessment Questionnaire.

Comparison

We considered two types of comparisons as relevant for validity: comparison between individuals with and without cognitive impairment such as Alzheimer's disease, and comparison between H/L and non-H/L groups.

Outcome of Interest

The outcome refers to the validity of the instrument. We chose search terms recommended by the Cochrane Handbook including either construct- or content-related validity.²⁰ We included studies that examined the validity of a construct (ie, Alzheimer's disease or related age-related cognitive disorder) as well as studies that examined the validity of the content (ie, cognitive functioning between H/L and non-H/L groups). We did not include studies that compared one cognitive assessment with another because no cognitive assessment for Latinos can be referred to as the gold standard. We also did not include studies that tested the internal factor structure of an assessment. The results were not restricted to a specific study design.

Search Strategy

We performed a literature search using PubMed and Web of Science on November 1, 2017. Search terms (Supplementary File 1, Table S1) were developed based on the previously described PICO model. We used many concept-related terms in the literature search that gave our search a wider focus to identify as many relevant studies as possible. The search was first conducted for each PICO category separately and then were combined with the operator "AND" to obtain all the relevant articles. We filtered the results for "has an abstract" to exclude all comments, editorial, book chapters, and so on, and we filtered for English or Spanish language only. In addition to the systematic literature search, a manual search was conducted using the references of the identified articles and related review articles as well as personal contacts.

Data Extraction

Search results were imported into the reference software End-Note. Two reviewers (S.P.A. and F.S.R.) systematically screened the search results for inclusion and exclusion criteria.

In a four-step process, we first screened for duplicates, then we screened the titles of the search results, and in a third step the abstracts of the remaining search results. Independent ratings by the reviewers were compared and discussed. Every study for which no consensus was achieved remained in the evaluation process. In the fourth step, the full texts of the studies were screened by all three authors. Besides verifying coherence with the inclusion and exclusion criteria, the quality of the studies was checked using the Standards for the Reporting of Diagnostic Accuracy Studies standards checklist.²¹ The evaluations of each reviewer were discussed in consensus discussions with all the reviewers. Each article was discussed in up to three meetings. Studies with minor limitations (eg, recruitment of control group from communicating with visitors in the clinic, small sample size) were included in the results if they met all the inclusion criteria. Studies with major limitations, such as severe bias in the design (eg, lack of statistics on the comparison of performance in cognitive testing, no control group, comparison focused on language only) were excluded.

Data Synthesis

Results were first categorized by type of validity testing: clinical vs nonclinical (within-group comparison for construct validity regarding Alzheimer's and dementia) and H/L vs non-H/L groups (between-group comparison for

content validity regarding ethnicity). Under these two headings, we present results first by geographic location (ie, studies conducted in the United States vs other Spanish-speaking countries) and then by assessment tool. The effect estimates for validity were summarized narratively because the heterogeneity of the studies did not allow us to perform a comprehensive meta-analysis. Differences between estimates for validity were interpreted based on the limitations of each study including participant selection and characteristics, sample size, methods for identifying clinical symptoms, cultural background, and statistical methods.

RESULTS

The search algorithm identified 4980 articles in PubMed and 1108 in Web of Science (N = 6088). One additional article was identified from the reference list of relevant studies. After removing 669 duplicates, we eliminated 5201 studies in the title screening (Figure 1). A total of 219 abstracts were evaluated. Only 94 articles met the inclusion and exclusion criteria and underwent the full-text screening. The quality screening led to the exclusion of 67 studies because they did not have a full text (1), did not investigate H/L groups (18), or investigated a young or special disease cohort (4), did not focus on cognitive functioning as outcome (6), did not use statistical methods for validity comparison (18), or did not investigate

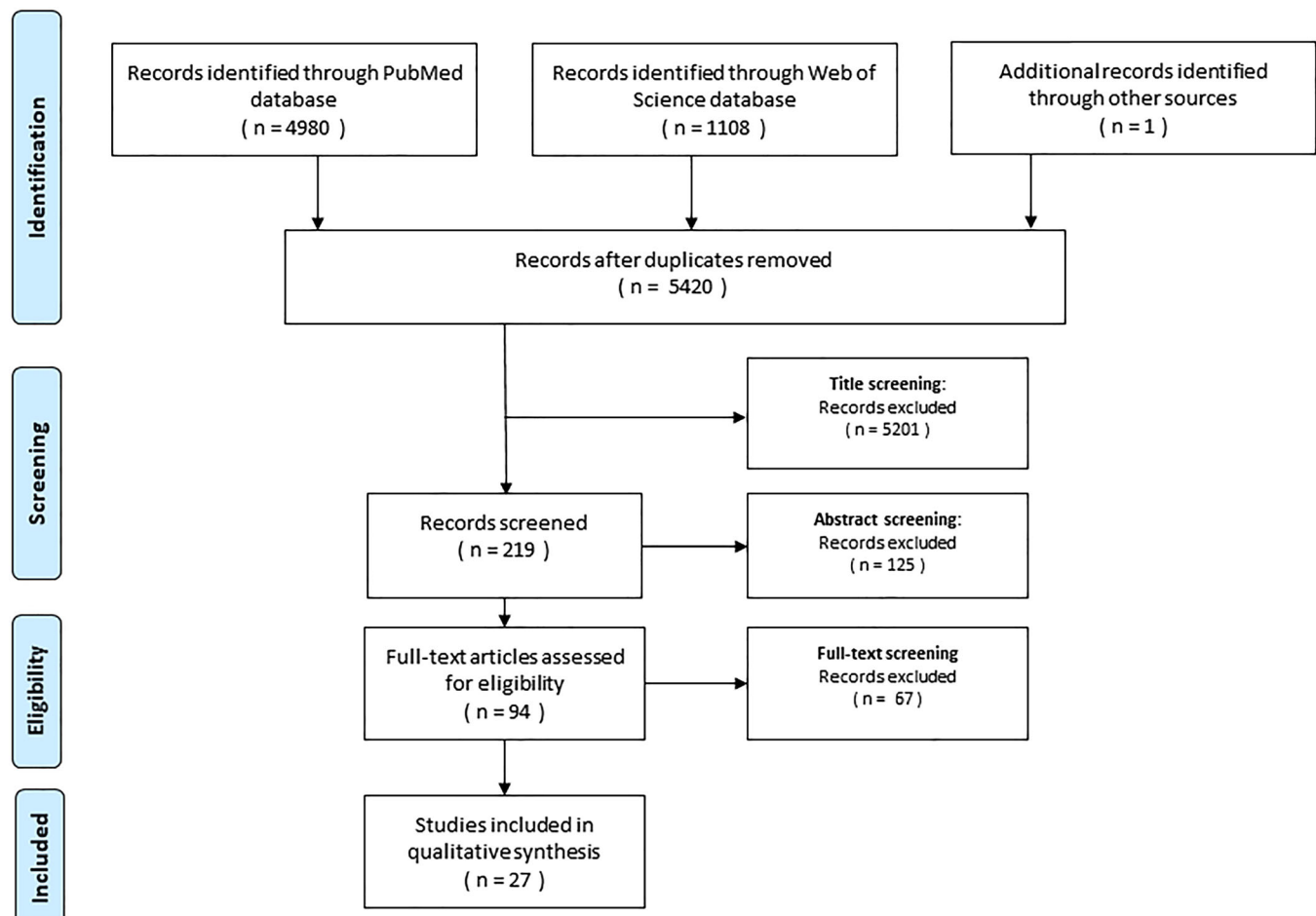


Figure 1. Flow diagram of the literature search process.

validity with respect to dementia diagnosis in H/L groups (20). A total of 27 met the inclusion criteria and quality standards. Three studies investigated more than one cognitive assessment.

Studies that met our inclusion criteria are shown in Supplementary File 2, Table S2. A total of eight studies examined the validity of the Mini-Mental Status Examination (MMSE), three studies the Addenbrooke Cognitive Examination-Revised (ACE-R), three studies the Spanish and English Neuropsychological Assessment Scales (SENAS), three the Montreal Cognitive Assessment (MoCA), four studies a naming test, two the verbal fluency test, one the clock-drawing test, one the Syndrom-Kurztest (SKT), one the 10/66 short diagnostic schedule (community Screening Instrument for Dementia), one the Executive Battery 25, one the Phototest, one the Eurotest, and one for four scales of the Wechsler Adult Intelligence Scale-Revised (WAIS-R) (description of assessments in Supplementary file 1, page 2).

The quality of the articles was good but not excellent. A flaw of many studies was that only a small amount of information on the participants' health status was provided; information on clinical diagnoses other than dementia is generally lacking. Studies with convenience samples should have at least controlled for the health status of their controls. Further, recruiting controls from within the clinic can lead to a selection bias because people who show up at the clinic are a special selection from the general population. Most studies are missing sample size estimations. Because some of the studies have sample sizes smaller than 100 subjects, we cannot be sure whether the sample size was big enough. We summarize results by method of comparison: clinical vs nonclinical (construct validity = 19) and H/L vs non-H/L groups (content validity = 11) comparisons.

Clinical vs Nonclinical: Hispanics/Latinos with Alzheimer's Disease or Related Cognitive Disorders Compared with Healthy Hispanics/Latinos

This section reports on studies that investigated construct validity of cognitive assessment tools by comparing clinical groups (H/L individuals with a clinical diagnosis of Alzheimer's or dementia) with nonclinical groups (healthy H/L individuals). Four studies were conducted in the United States and 15 studies outside the country.

Studies Conducted in the United States

Mini-Mental Status Examination (MMSE): Three of four studies that compared clinical with nonclinical participants examined the validity of the MMSE. The MMSE's sensitivity and specificity ranged from 74.5% to 100% and 97.9% to 98%, respectively²²⁻²⁴ (cutoff of 21; see Supplementary File 2, Table S2). Sensitivity was higher in the sample of the Aging, Demographics and Memory Study (a nationally representative subsample of the Health and Retirement Study) than in the sample from the Alzheimer's Disease Research Center (ADRC). The study that adjusted the MMSE score for age and education achieved the highest sensitivity but came with a lower specificity, running the risk of classifying healthy H/L individuals as having dementia.²²

The Naming Test: One study compared the Texas Spanish Naming Test in clinical vs nonclinical participants from the United States (Texas), Colombia, and Spain. It reported significantly lower scores in clinical patient participants vs cognitive

healthy people, suggesting the test might be able to differentiate these two groups²⁵ (Supplementary File 2, Table S2).

Studies Conducted in Hispanic/Latino Countries

Mini-Mental Status Examination

A study in Chile used the US-approved cutoff of 21 and reported a sensitivity of 94% and a specificity of 46%²⁶ (Supplementary File 2, Table S2). The low specificity comes with a high risk of falsely identifying a healthy Chilean as sick.

Naming Test

Two studies from Argentina investigated the validity of the naming test. The first one used the Confrontation Naming Test (*z*-score cutoff) and reported a sensitivity of 74% and a specificity of 77%.²⁷ The second one used the Boston Naming Test with education-specific cutoff scores (Supplementary Table S2) and reported a sensitivity of 39% and a specificity of 89%.²⁸ The low sensitivity may result in some Argentinians with dementia remaining undiagnosed.

Addenbrooke Cognitive Examination-Revised

Three studies investigated the validity of the ACE-R with country-specific modifications in convenience samples (Supplementary File 2, Table S2). The Peruvian version used a cutoff of 86 and reported a sensitivity of 100% and a specificity of 100%,²⁹ the Chilean version (cutoff = 76), revealed a sensitivity of 92% and a specificity of 93%,³⁰ and the Argentinean (cutoff = 86) revealed a sensitivity of 92% and a sensitivity of 96%.³¹

Montreal Cognitive Assessment

Three studies investigated the validity of the MoCA in convenience samples (Supplementary File 2, Table S2). With respect to dementia, one study in Chile used a cutoff of 21 that was adjusted for education (+1 point for 8-12 y of education, +2 points for <8 y of education) and revealed a sensitivity of 75% and a specificity of 90%.³² A second study in Mexico used a cutoff of 24 and showed a sensitivity of 98% and specificity of 93%.³³ With respect to mild cognitive impairment, a study in Mexico (cutoff = 26) revealed a sensitivity of 80% and a sensitivity of 75%,³³ a Colombian study (cutoff = 23) a sensitivity of 89% and a specificity of 79%,³⁴ and a Chilean study (cutoff = 21) a sensitivity of 81% and a specificity of 90%.³²

Clock-Drawing Test

One study in Peru assessed the validity of the clock-drawing test and obtained a sensitivity of 99% and a specificity of 83%³⁵ (Supplementary File 2, Table S2).

Syndrom-Kurztest

A study conducted by the World Health Organization used a random population sample in Chile to investigate the validity of the SKT. Results show significant differences in the scores between cognitively normal people and people with dementia³⁶ (Supplementary File 2, Table S2).

10/66 Short Diagnostic Schedule (Community Screening Instrument for Dementia)

The 10/66 short diagnostic schedule was developed in the 10/66 study³⁷ and tested in stratified random population samples in Cuba, Dominican Republic, Peru, Mexico, Venezuela, India, China, and Puerto Rico. Results indicated a sensitivity of 94% for identifying dementia patients³⁸ (Supplementary File 2, Table S2).

Executive Battery 25

A study in Argentina revealed a sensitivity of 94% and a specificity of 100% (cutoff = 15) for the Executive Battery 25³⁹ (Supplementary File 2, Table S2).

Phototest

Findings from Argentina on the Phototest revealed a sensitivity of 89% and a specificity of 97% (cutoff = 27)⁴⁰ (Supplementary File 2, Table S2).

Eurotest

A study in Peru indicated a sensitivity of 91% and a specificity of 83% (cutoff = 24) for the Eurotest⁴¹ (Supplementary File 2, Table S2).

Hispanics/Latino vs Non-Hispanics/Latinos

Eleven studies conducted in the United States tested the validity of cognitive assessments by comparing the performance of H/Ls with the performance of non-H/Ls.

Mini-Mental Status Examination

Three longitudinal aging studies reported significant differences in performance in the MMSE between the H/L and Non-H/L groups after adjusting for important confounders such as education and socioeconomic status⁴²⁻⁴⁴ (odds ratio for H/Ls vs non-H/Ls performing under the cutoff = 2.12; *P* values for difference in performance between H/Ls and non-H/Ls between <.001 and .012; Supplementary File 2, Table S2). Another study used the sample from the ADRC in La Jolla, California, and did not observe significant differences in performance (*P* = .66).⁴⁵ Dissimilarities in the recruitment procedure might have influenced the findings.

Naming Test

One study investigated the Boston Naming Test and did not find significant differences in performance by H/Ls compared with non-H/L groups⁴⁶ (Supplementary File 2, Table S2).

Spanish and English Neuropsychological Assessment Scales

Three studies investigated the validity of the SENAS and reported significant differences by ethnicity with non-H/Ls outperforming H/Ls (all *P* values <.02)^{15,44,47} (Supplementary File 2, Table S2). Adjusting for language abilities and other socioeconomic factors attenuated this effect.²²

Verbal Fluency Test

One study investigated the verbal fluency test using the sample of the ADRC and found no significant difference in performance (animals and vegetables) between H/Ls and non-H/Ls.⁴⁶ The study with random population sampling in New Mexico, however, found significant differences in test performance (same-sex first names) for H/L ethnicity⁴⁸ (Supplementary File 2, Table S2).

Scales of the WAIS-R

One study investigated the validity of the WAIS-R using a sample from New Mexico (Supplementary File 2, Table S2). Overall, H/Ls performed significantly poorer in the digit forward, full retrieval, recall, and color trail test than non-H/Ls (even if the test was conducted in Spanish), an effect that depended on education and English-language proficiency.⁴⁸

DISCUSSION

The study aim was to review existing articles that have examined and reported on the validity of cognitive assessment tools in Hispanic/Latino (H/L) population groups in the United States¹⁶ as well as Spanish-speaking population groups in Mexico, Central America, and South America. We identified 27 studies that tested the validity of 13 different instruments. Studies investigated either construct validity by comparing cognitively healthy H/L individuals with H/Ls with dementia or content validity by comparing ethnic groups. Findings related to ethnic group comparisons indicate that performance in the tests is significantly different within samples with a very low education level. Findings with respect to construct validity also indicate that education is an important factor that influences validity.

Our results indicate good diagnostic accuracy of the MMSE for H/L groups in the United States because it was tested in different samples and achieved sensitivities from 75% to 100%.²²⁻²⁴ However, because performance is sensitive to educational attainment and socioeconomic status,⁴²⁻⁴⁴ it may falsely classify H/L individuals with very low education levels or H/L individuals outside the United States as having dementia. Alternative cognitive assessments such as the SENAS and WAIS-R have the same bias for education. Good diagnostic performance was reported for culturally adapted versions of the ACE^{29-31,49} and the MoCA,³²⁻³⁴ but all the studies were carried out in H/L countries with convenience samples. These findings must be validated with H/Ls living in the United States before clinician can rely on them.

Based on the evidence available for this review, we recommend clinicians use the MMSE for cognitive testing of H/L individuals, mainly because it was the only instrument tested with large population samples in the United States. Because this assessment tool is sensitive to education, we recommend clinicians who work with people with H/L with very low education levels to conduct further cognitive tests to validate a dementia diagnosis. Brief assessments designed for illiterate individuals, such as the clock-drawing test, the Phototest, and the Eurotest, can offer additional information on the patients' cognitive status. Yet even though studies indicate a high validity outside the United States, further studies with H/L individuals inside the country are required.

For the 10/66 short diagnostic schedule, the level of evidence is similar, but the quality is higher. It was developed by an international team of experts and validated with stratified random population samples in six different countries. We encourage clinicians to try using the 10/66 short diagnostic schedule for H/L in the United States and report their experiences.

Validity testing for H/L groups is urgently needed. One of the greatest challenges is to identify appropriate normative data to interpret results. Current norms based on monolingual English-speaking white samples may lead to an inaccurate dementia diagnosis when administered to H/L individuals.⁵⁰ In addition, other behavioral and contextual factors need to be accounted for that may have a disproportionate effect on the cognitive health of H/Ls. Clinical diagnosis depends on symptom presentation that is not always clearly communicated due to language barriers, patient-physician communication dynamics, a lack of cultural and linguistically competent testing methods, and lack of cultural sensitivity from healthcare providers.^{10,51} Patients' demographic and sociocultural factors, such as quality of education, literacy level, language proficiency, acculturation level, and distinct presentation of symptoms^{11,52,53} have an effect on cognitive test performance and therefore need to be considered when creating norms to assess H/L group's cognitive status in the United States. The high sensitivity and specificity values obtained in studies from Spanish-speaking countries are an indication for the importance of social and cultural adaptation of cognitive assessment tools.

National efforts to address the needs of H/L groups in the United States have focused on translating and adapting existing English measures (eg, the National Institutes of Health Toolbox Cognition Battery in Spanish⁵⁴). This is an important first step, but we need to move forward translating cultural aspects that affect cognitive performance into our diagnostic tools. Differences found in performance in the verbal fluency test and the naming test may be related to ethnocultural influences on the familiarity, salience, and cultural relevance. For instance, different levels of exposure to vegetables due to culturally influenced kitchen behaviors and norms may result in differences in test performance. In addition, bilingual H/L individuals with proficient language skills may mask an underlying dementia pathology due to better performance on tests because bilingualism seems to have a protective effect against dementia symptoms.⁵⁵

Conclusions derived from this review have limitations. First, a common problem of most studies was that no sample size calculation was provided. Optimal sample sizes offer stronger implications on validity. Second, many studies did not describe how they dealt with missing data, a relevant source of bias considering only a few studies were undertaken for each assessment. Further, it was not possible to run a meta-analysis due to a great variety of cognitive assessments and a limited number of studies on each assessment. Furthermore, there could be a publication bias of significant findings that might have influenced our interpretation. Other limitations include the lack of ability to compare results with respect to their setting (in the clinic vs in the community), differences in the sampling methods, and the demographic and cultural characteristics of the study samples that are not always comparable. Finally, we cannot guarantee that we might have missed a publication.

In conclusion, evidence on validity indicates that the MMSE is an appropriate cognitive assessment tool for H/Ls with at least a moderate education. For H/L individuals with low education levels, additional assessment tools should be used that were designed for illiterate patients (eg, the clock-drawing test, the Phototest, the Eurotest, or the 10/66 short diagnostic schedule). These tests, however, have not yet been validated for H/L groups in the United States. To reduce current trends of culturally biased under- or overdiagnosis of cognitive impairments, more research is needed to adapt cognitive tools to assess H/Ls for Alzheimer's disease and other related dementias in the United States, for instance, by taking into account differences in English-language proficiency, bilingualism, number of years living in the United States, and acculturation level.

ACKNOWLEDGMENTS

Many thanks to the teams of researchers who worked on the study that we reviewed. Open access funding enabled and organized by Projekt DEAL.

Financial Disclosure: This report was supported by the German Research Foundation (DFG; no. TH2137/3-1) and the Hans and Ilse Breuer Foundation.

Conflicts of Interest: The authors have declared no conflicts of interest for this article.

Author Contributions: *Conceptualization of the study:* Arévalo and Rodriguez. *Literature search:* Arévalo and Rodriguez. *Screening of the search results:* Arévalo, Kress, and Rodriguez. *Interpretation of the results:* Arévalo and Rodriguez. *Drafting and revising the article:* Arévalo and Rodriguez. *Approving the final version of the article for submission:* All authors.

Sponsor's Role: The sponsors had no role in the planning of the study, the analysis, or the results.

REFERENCES

1. World Health Organization. World Report on Aging and Health. Geneva, Switzerland: World Health Organization; 2015.
2. Mehta KM, Yeo GW. Systematic review of dementia prevalence and incidence in United States race/ethnic populations. *Alzheimers Dement*. 2017; 13(1):72-83.
3. Mayeda ER, Glymour MM, Quesenberry CP, Whitmer RA. Inequalities in dementia incidence between six racial and ethnic groups over 14 years. *Alzheimers Dement*. 2016;12(3):216-224.
4. Wu S, Vega WA, Resendez J, Jin H. Latinos & Alzheimer's Disease: New Numbers Behind the Crisis. Los Angeles, CA: USC Edward Roybal Institute on Aging; 2016. https://health.ucdavis.edu/latinoaging/images/pdf/Latinos-and-AD_USC_UsA2-Impact-Report.pdf. Accessed May 18, 2018.
5. Mayeda ER, Glymour MM, Quesenberry CP, Johnson JK, Perez-Stable EJ, Whitmer RA. Survival after dementia diagnosis in five racial/ethnic groups. *Alzheimers Dement*. 2017;13(7):761-769.
6. DeNavas-Walt C, Proctor BD, Smith JC. Income and Poverty in the United States: 2014. Washington, DC: US Census Bureau; 2015.
7. Colón M, Lyke J. Comparison of hospice use and demographics among European Americans, African Americans, and Latinos. *Am J Hosp Palliat Care*. 2003;20(3):182-190.
8. Cutts T, Langdon S, Meza FR, et al. Community health asset mapping partnership engages Hispanic/Latino health seekers and providers. *N C Med J*. 2016;77(3):160-167.
9. Sayegh P, Knight BG. Cross-cultural differences in dementia: the sociocultural health belief model. *Int Psychogeriatr*. 2013;25(4):517-530.
10. Watson JL, Ryan L, Silverberg N, Cahan V, Bernard MA. Obstacles and opportunities in Alzheimer's clinical trial recruitment. *Health Aff (Millwood)*. 2014;33(4):574-579.

11. Chin AL, Negash S, Hamilton R. Diversity and disparity in dementia: the impact of ethnoracial differences in Alzheimer's disease. *Alzheimer Dis Assoc Disord*. 2011;25(3):187-195.
12. Ortiz F, Fitten LJ, Cummings JL, Hwang S, Fonseca M. Neuropsychiatric and behavioral symptoms in a community sample of Hispanics with Alzheimer's disease. *Am J Alzheimers Dis Other Dement*. 2006;21(4):263-273.
13. Sayegh P, Knight BG. Assessment and diagnosis of dementia in Hispanic and non-Hispanic white outpatients. *Gerontologist*. 2013;53(5):760-769.
14. Hargrave R, Stoeklin M, Haan M, Reed B. Clinical aspects of dementia in African-American, Hispanic, and white patients. *J Natl Med Assoc*. 2000;92(1):15-21.
15. Mungas D, Widaman KF, Reed BR, Tomaszewski Farias S. Measurement invariance of neuropsychological tests in diverse older persons. *Neuropsychology*. 2011;25(2):260-269.
16. Suárez-Orozco MM, Pérez MM. Introduction: the research agenda. In: Suárez MM, Pérez MM, eds. *Latinos: Remaking America*. Berkeley: University of California Press; 2002:1-37.
17. Huang X, Lin J, Demner-Fushman D. Evaluation of PICO as a knowledge representation for clinical questions. *AMIA Annu Symp Proc*. 2006;2006:359-363.
18. Kastner M, Wilczynski NL, Walker-Dilks C, McKibbin KA, Haynes B. Age-specific search strategies for Medline. *J Med Internet Res*. 2006;8(4):e25.
19. US Census Bureau. Overview of Race and Hispanic Origin: 2010 Census Briefs. <https://www.census.gov/prod/cen2010/briefs/c2010br-02.pdf>. Accessed May 1, 2018.
20. Cochrane Handbook for Systematic Reviews of Interventions. Chichester, UK: John Wiley & Sons; 2008.
21. Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Clin Chem Lab Med*. 2003;41(1):68-73.
22. Mungas D, Marshall SC, Weldon M, Haan M, Reed BR. Age and education correction of Mini-Mental State Examination for English and Spanish-speaking elderly. *Neurology*. 1996;46(3):700-706.
23. Rideaux T, Beaudreau SA, Fernandez S, O'Hara R. Utility of the abbreviated full object memory evaluation and MMSE for detection of dementia and cognitive impairment not dementia in diverse ethnic groups. *J Alzheimers Dis*. 2012;31(2):371-386.
24. Taussig IM, Mack WJ, Henderson VW. Concurrent validity of Spanish-language versions of the Mini-Mental State Examination, Mental Status Questionnaire, Information-Memory-Concentration test, and Orientation-Memory-Concentration test: Alzheimer's disease patients and nondemented elderly comparison subjects. *J Int Neuropsychol Soc*. 1996;2(4):286-298.
25. de La PCM, Arango-Lasprilla JC, Alegret M, et al. Item analysis of three Spanish naming tests: a cross-cultural investigation. *NeuroRehabilitation*. 2009;24(1):75-85.
26. Quiroga P, Albala C, Klaasen G. Validation of a screening test for age associated cognitive impairment, in Chile [in Spanish]. *Rev Med Chil*. 2004;132(4):467-478.
27. Fernandez AL. Development of a confrontation naming test for Spanish-speakers: the Cordoba naming test. *Clin Neuropsychol*. 2013;27(7):1179-1198.
28. Fernandez AL, Fulbright RL. Construct and concurrent validity of the Spanish adaptation of the Boston naming test. *Appl Neuropsychol Adult*. 2015;22(5):355-362.
29. Custodio N, Lira D, Montesinos R, Gleichgerricht E, Manes F. Usefulness of the Addenbrooke's cognitive examination (Spanish version) in Peruvian patients with Alzheimer's disease and frontotemporal dementia. *Vertex*. 2012;23(103):165-172.
30. Munoz-Neira C, Henriquez Ch F, Ihnen JJ, Sanchez CM, Flores MP, Slachevsky Ch A. Psychometric properties and diagnostic usefulness of the Addenbrooke's cognitive examination-revised in a Chilean elderly sample. *Rev Med Chil*. 2012;140(8):1006-1013.
31. Sarasola D, de Luján-Calcagno M, Sabe L, et al. validity of the Spanish version of the Addenbrooke's cognitive examination for the diagnosis of dementia and to differentiate Alzheimer's disease and frontotemporal dementia. *Rev Neurol*. 2005;41(12):717-721.
32. Delgado C, Araneda A, Behrens MI. Validation of the Spanish-language version of the Montreal Cognitive Assessment test in adults older than 60 years. *Neurologia*. 2019;34(6):376-385.
33. Aguilar-Navarro SG, Mimenza-Alvarado AJ, Palacios-Garcia AA, Samudio-Cruz A, Gutierrez-Gutierrez LA, Avila-Funes JA. Validity and reliability of the Spanish version of the Montreal Cognitive Assessment (MoCA) for the detection of cognitive impairment in Mexico. *Rev Colomb Psiquiatr*. 2018;47(4):237-243.
34. Gil L, Ruiz de Sanchez C, Gil F, Romero SJ, Pretelt Burgos F. Validation of the Montreal Cognitive Assessment (MoCA) in Spanish as a screening tool for mild cognitive impairment and mild dementia in patients over 65 years old in Bogota, Colombia. *Int J Geriatr Psychiatry*. 2015;30(6):655-662.
35. Custodio N, Garcia A, Montesinos R, Lira D, Bendezu L. Validation of the clock drawing test - Manos' version - as a screening test for detection of dementia in older persons of Lima, Peru [in Spanish]. *Rev Peru Med Exp Salud Publica*. 2011;28(1):29-34.
36. Fornazzari L, Cumsille F, Quevedo F, et al. Spanish validation of the Syndrom Kurztest (SKT). *Alzheimer Dis Assoc Disord*. 2001;15(4):211-215.
37. Prince M. Dementia in developing countries. A consensus statement from the 10/66 Dementia Research Group. *Int J Geriatr Psychiatry*. 2000;15(1):14-20.
38. Stewart R, Guerchet M, Prince M. Development of a brief assessment and algorithm for ascertaining dementia in low-income and middle-income countries: the 10/66 short dementia diagnostic schedule. *BMJ Open*. 2016;6(5). <https://doi.org/10.1136/bmjopen-2015-010712>.
39. Serrani Azcurra DJL. Spanish translation and validation of an executive battery 25 (EB25) and its shortened version (ABE12) for executive dysfunction screening in dementia. *Neurologia*. 2013;28(8):457-476.
40. Russo MJ, Iturry M, Sraka MA, Bartoloni L, Carnero Pardo C, Allegri RF. Diagnostic accuracy of the Phototest for cognitive impairment and dementia in Argentina. *Clin Neuropsychol*. 2014;28(5):826-840.
41. Oscanoa TJ, Cieza E, Parodi JF, Paredes N. Evaluation of Peruvian money test in screening of cognitive impairment among older adults [in Spanish]. *Rev Peru Med Exp Salud Publica*. 2016;33(1):67-73.
42. Espino DV, Lichtenstein MJ, Palmer RF, Hazuda HP. Ethnic differences in Mini-Mental State Examination (MMSE) scores: where you live makes a difference. *J Am Geriatr Soc*. 2001;49(5):538-548.
43. Mulgrew CL, Morgenstern N, Shetterly SM, Baxter J, Baron AE, Hamman RF. Cognitive functioning and impairment among rural elderly Hispanics and non-Hispanic whites as assessed by the Mini-Mental State Examination. *J Gerontol B Psychol Sci Soc Sci*. 1999;54(4):P223-P230.
44. Mungas D, Reed BR, Farias ST, DeCarli C. Criterion-referenced validity of a neuropsychological test battery: equivalent performance in elderly Hispanics and non-Hispanic whites. *J Int Neuropsychol Soc*. 2005;11(5):620-630.
45. Hohl U, Grundman M, Salmon DP, Thomas RG, Thal LJ. Mini-Mental State Examination and Mattis Dementia Rating Scale performance differs in Hispanic and non-Hispanic Alzheimer's disease patients. *J Int Neuropsychol Soc*. 1999;5(4):301-307.
46. Sayegh P. Neuropsychological language tests in dementia diagnosis in English-speaking Hispanic and non-Hispanic white outpatients. *Appl Neuropsychol Adult*. 2015;22(6):435-444.
47. Mungas D, Reed BR, Haan MN, Gonzalez H. Spanish and English neuropsychological assessment scales: relationship to demographics, language, cognition, and independent function. *Neuropsychology*. 2005;19(4):466-475.
48. La Rue A, Romero LJ, Ortiz IE, Liang HC, Lindeman RD. Neuropsychological performance of Hispanic and non-Hispanic older adults: an epidemiologic survey. *Clin Neuropsychol*. 1999;13(4):474-486.
49. Carvalho VA, Barbosa MT, Caramelli P. Brazilian version of the Addenbrooke Cognitive Examination-revised in the diagnosis of mild Alzheimer disease. *Cogn Behav Neurol*. 2010;23(1):8-13.
50. Gasquoin PG. Research in clinical neuropsychology with Hispanic American participants: a review. *Clin Neuropsychol*. 2001;15(1):2-12.
51. Valle R. Culture-fair behavioral symptom differential assessment and intervention in dementing illness. *Alzheimer Dis Assoc Disord*. 1994;8:21-45.
52. Manly JJ, Espino DV. Cultural influences on dementia recognition and management. *Clin Geriatr Med*. 2004;20(1):93-119.
53. Therrien M, Ramirez RR. The Hispanic Population in the United States. Current Population Reports. Washington, DC: US Department of Commerce; 2000.
54. Flores I, Casaleto KB, Marquine MJ, et al. Performance of Hispanics and non-Hispanic whites on the NIH Toolbox Cognition Battery: the roles of ethnicity and language backgrounds. *Clin Neuropsychol*. 2017;31(4):783-797.
55. Schweizer TA, Ware J, Fischer CE, Im Craik F, Bialystok E. Bilingualism as a contributor to cognitive reserve: evidence from brain atrophy in Alzheimer's disease. *Cortex*. 2012;48(8):991-996.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Table S1: Search Terms Used in the Database Search.

Table S2: Validity Estimates of Instruments to Assess Cognitive Functioning in Hispanics/Latinos as Presented in the Studies Identified in the Systematic Literature Search.