Diagnosis, Assessment
& Disease Monitoring

RESEARCH ARTICLE

Cognitive impairment in racially/ethnically diverse older adults: Accounting for sources of diagnostic bias

Dan Mungas[1] | Crystal Shaw[2] | Eleanor Hayes-Larson[2] | Charles DeCarli[1] |
Sarah Tomaszewski Farias[1] | John Olichney[1] | Hector Hernandez Saucedo[1] |
Paola Gilsanz[3] | M Maria Glymour[4] | Rachel A Whitmer[5] | Elizabeth Rose Mayeda[2]

[1] Department of Neurology, University of California, Davis, Sacramento, California, USA

[2] Department of Epidemiology, Fielding School of Public Health, University of California, Los Angeles, California, USA

[3] Kaiser Permanente Division of Research, Oakland, California, USA

[4] Department of Epidemiology and Biostatistics, University of California, San Francisco, California, USA

[5] Department of Public Health Sciences, University of CaliforniaDavis, Davis, California, USA

Correspondence
Dan Mungas, Department of Neurology, University of California, Davis, Sacramento, California, USA.
Email: dmmungas@ucdavis.edu

Funding information
the National Institute on Aging (NIA), Grant/Award Numbers: RF1 AG052132, R00 AG053410, F31 AG071191

Abstract

Introduction: The Kaiser Healthy Aging and Diverse Life Experiences (KHANDLE) study enrolled Asian, Black, Latino, and White adults ages 65+ without prior dementia diagnosis (N = 1709). We evaluated the prevalence of cognitive impairment (mild cognitive impairment or dementia) accounting for potential biases.

Methods: A random subgroup (N = 541) received clinical evaluation and others were evaluated if they failed a cognitive screen. Diagnoses were made under two conditions: (1) demographics-blind, based on clinical exam and demographically adjusted neuropsychological test scores; and (2) all available information (clinical exam, demographics, and adjusted and unadjusted test scores).

Results: Cognitive impairment prevalence was 28% for blinded-adjusted diagnosis and 25% using all available information. Black participants had higher impairment rates than White (both conditions) and Latino (blinded-adjusted diagnosis) participants. Incomplete assessments negatively biased prevalence estimates for White participants.

Discussion: Racial/ethnic disparities in cognitive impairment were amplified by attrition bias in White participants but were unaffected by type of test norms and diagnosticians' knowledge of demographics.

KEYWORDS
aging, cognitive impairment, dementia, ethnicity, mild cognitive impairment, race

1 | BACKGROUND

The US older adult population is becoming increasingly diverse.[1] There is evidence of racial/ethnic disparities in cognitive impairment and dementia[2-5 and a] need for studies that address these disorders in racially/ethnically diverse samples. However, few previous studies have had the sample diversity available to simultaneously examine and compare rates of cognitive impairment in multiple racial/ethnic groups using the same methods.

Comparative studies of cognitive health across diverse groups present special challenges. Cognitive test results play an important role in evaluating cognitive impairment. Racial/ethnic differences in cross-sectional test scores are well documented[6-10] and failure to account for normative group differences can result in misdi-

agnosing cognitively normal individuals from some groups with cognitive impairment.[11–13] Racial/ethnic group-specific norms are one approach to minimize this source of bias, but this is not without controversy.[11,14–16] Attrition/drop-out is a second threat to validity of group comparisons. Those with cognitive impairment may be less likely to complete cognitive and clinical assessments. Estimates of cognitive impairment based on those who completed assessments might underestimate true impairment rates, and group comparisons can be further biased if there are differential drop-out rates across groups. Implicit bias—subconscious associations that promote negative evaluations of persons based on race, gender, or other characteristics[17]—is another potential threat to diagnostic accuracy that has not been well studied in the context of age-related cognitive impairment.

The Kaiser Healthy Aging and Diverse Life Experiences (KHANDLE) study was designed to evaluate how life-course experiences and sociocultural factors influence late-life brain health and cognitive decline and contribute to racial/ethnic disparities in cognitive impairment and dementia. KHANDLE is unique in that it enrolled non-demented older adults from four racial/ethnic groups who were recruited from a defined population. The clinical evaluation protocol includes rigorous, standardized clinical assessment procedures that yield comprehensive clinical data. This makes it possible to systematically vary conditions under which study clinicians make diagnoses to examine how different information influences group differences in diagnoses. In this study, we examined (1) the prevalence of cognitive impairment in this racially/ethnically diverse sample, and (2) whether group differences in cognitive impairment rates were affected by three methodological factors: use of demographically adjusted versus unadjusted neuropsychological test scores, diagnostician knowledge of demographic characteristics, and attrition bias related to failure to complete assessments.

2 | METHODS

2.1 | Participants

KHANDLE participants were long-term members of the Kaiser Permanente Northern California (KPNC) integrated health care system, ages 65+ years in 2017, spoke English or Spanish, and had participated previously in Kaiser Permanente multiphasic health checkups (MHCs) between 1964 and 1985. KHANDLE recruited approximately equal numbers of Asian, Black, Latino, and White older adults who were diverse in educational attainment. This was implemented with stratified random sampling by race/ethnicity and educational attainment applied to the population of MHC participants who did not have a medical records diagnosis of dementia (≈30,000 available in 2017). Exclusion criteria included health conditions that would impede participation in study interviews (eg, hospice activity in the past 12 months, history of severe chronic obstructive pulmonary disease in the past 6 months, congestive heart failure hospitalizations in the past 6 months, history of end-stage renal disease or dialysis in the past 12 months). Three of 1712 participants enrolled at baseline reported Native Amer-

RESEARCH IN CONTEXT

1. **Systematic review**: We reviewed the literature on studies examining prevalent cognitive impairment across diverse racial/ethnic groups using traditional (eg, PubMed) sources. We advance this literature by examining how cognitive impairment differences across groups are affected by sources of bias in clinical diagnosis.

2. **Interpretation**: Cognitive impairment rates were greater than 20% all in four racial/ethnic groups (Asian, Black, Latino, and White). Black participants had the highest prevalence of cognitive impairment and differed significantly from White (both diagnosis conditions) and Latino (one condition) participants. Group differences were amplified by attrition bias in White participants but were not affected by knowledge of demographics or use of demographically adjusted test scores.

3. **Future directions**: Future research should examine dementia incidence and continuous cognitive decline across racially/ethnically diverse groups of older adults. This study can inform approaches to identify and control for factors that can bias results and distort estimation of group differences in cognitive decline and dementia.

ican or unknown race and were dropped from the analyses due to the small cell size, yielding an analytic sample of 1709.

2.1.1 | KHANDLE study design

KHANDLE was a collaboration between the KPNC Division of Research and UC Davis Alzheimer's Disease Center. The study was approved by Kaiser and UC Davis Institutional Review Boards and all enrolled participants provided informed consent. A schematic of the study design is presented in Figure 1. The Kaiser team recruited participants and administered a study interview, which included two cognitive test batteries: the Spanish and English Neuropsychological Assessment Scales (SENAS)[9] and the National Institutes of Health (NIH) Toolbox Cognitive Health Battery (NIHTB-CHB).[18–20] At the study interview, participants were asked to consent to referral to UC Davis, where the clinical evaluation component was completed.

The sample was divided into a Random Selection group (N = 541) and a Screen Selection group (N = 1168). Assignment to these groups was random within each racial/ethnic group. All Random Selection group participants were invited to receive a full clinical evaluation, including clinical neuropsychological testing and a clinical exam. Based on a screening algorithm developed using results from the Random Selection group (details in Supplementary Materials), Screen Selection group participants were selected for clinical neuropsychological test-
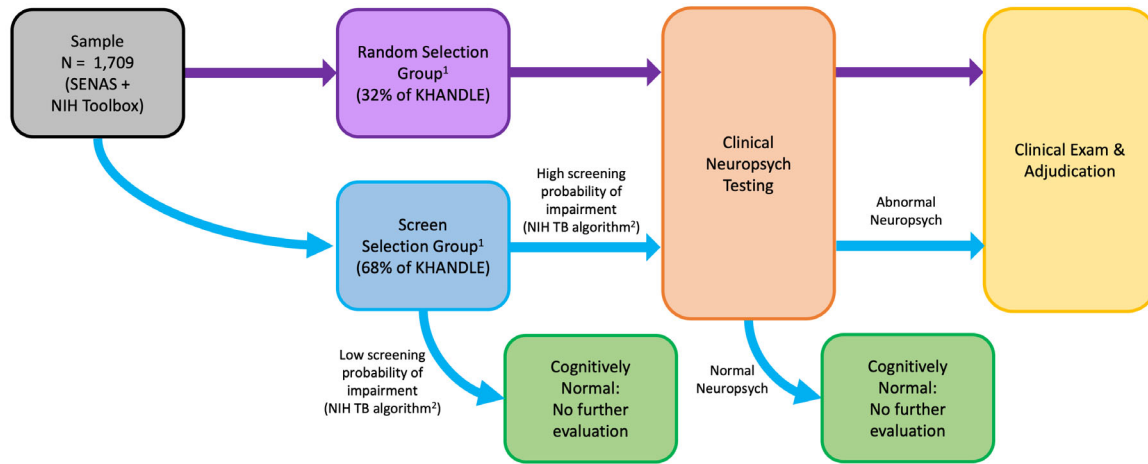
[1] Stratified by race/ethnicity

[2] Logistic regression model: Cognitive Impairment/Dementia vs. Normal ~ NIH TB (flanker, list sorting, sequence memory, picture vocabulary) + age + education + gender + race/ethnicity

FIGURE 1 Study design [All participants received SENAS and NIHTB-CHB cognitive testing (excluded from consideration for diagnosis). The Random Selection group was selected randomly to receive full clinical evaluation (clinical neuropsychological testing and clinical exam). Individuals in the Screen Selection group were selected for clinical neuropsychological testing if they failed the cognitive screen, and received clinical exam and adjudicated diagnosis if neuropsychological results were abnormal

ing and were referred for a clinical exam if clinical neuropsychological test results were abnormal.

2.2 | Clinical evaluation and diagnosis

2.2.1 | Clinical evaluation components

Clinical neuropsychological testing was performed by a trained psychometrist; the battery consisted of version 3 of the Uniform Dataset of the National Institute on Aging Alzheimer's Disease Centers program[21] supplemented by the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) list learning test and CERAD drawing copy and delayed recall.[22] The clinical exam was administered by a physician trained in clinical dementia assessment and included a medical history and history of cognitive problems, physical and neurological exam, mental status testing using the Montreal Cognitive Assessment (MoCA),[23] and the Clinical Dementia Rating (CDR) scale.[24]

2.2.2 | Clinical neuropsychological test scoring and adjudication

Unadjusted and demographically adjusted scores were calculated for each clinical neuropsychological test. Adjusted scores controlled for race/ethnicity, education, and gender (details in Supplementary Materials). Briefly, linear regression methods applied to baseline test scores of 810 cognitively normal individuals (343 from the Random Selection group and 467 from the UC Davis Alzheimer's Disease Center Diversity cohort[25]) removed variance related to these demographic characteristics. Normative values for unadjusted and adjusted scores were

derived from cognitively normal KHANDLE Random Selection group individuals.

2.2.3 | Diagnosis adjudication

Clinical diagnosis was adjudicated by three senior clinicians (two neurologists and a neuropsychologist), who had extensive dementia assessment experience and were KHANDLE investigators. Diagnosticians reviewed clinical exam results (MoCA, the CDR, medical history and presentation of symptoms, examining clinician's impression), neuropsychological test results, and a clinical neuropsychological adjudication of test results). The diagnostician made sequential decisions under varied conditions to examine how different information influences diagnosis: (1) demographics-blind classification (normal, mild cognitive impairment [MCI], dementia) based on clinical exam results and demographically adjusted neuropsychological scores ("blinded-adjusted diagnosis"), and (2) final classification based on all available information ("final-all available information diagnosis"). Diagnostic criteria are described in the Supplementary Materials. The reliability of clinical diagnosis was formally evaluated. Due to the small number of dementia diagnoses, we evaluated inter-rater reliability of diagnosis of cognitive impairment (MCI or dementia). Agreement (94%, 97%) and Cohen's Kappa[26] values (0.86, 0.93) were high for both conditions (see Supplementary Materials for more detail).

2.3 | Nondiagnostic cognitive assessments

All participants completed the SENAS and the NIHTB-CHB before the clinical evaluation; these results were excluded from diagnostic

adjudications. SENAS was developed to measure cognition in diverse racial/ethnic groups, has English and Spanish language versions, and assesses three cognitive domains (verbal episodic memory, semantic memory, executive function).[9,27–32] Four NIHTB-CHB measures, administered using the NIH Toolbox iPad App, were used to measure episodic memory (Picture Sequence Memory Test), language (Picture Vocabulary), executive function (Flanker), and working memory (List Sorting). (SENAS and NIHTB-CHB are described in more detail in Supplementary Materials.)

2.4 | Data analysis

2.4.1 | Design and data analysis overview

Our outcomes were (1) blinded-adjusted diagnosis and (2) final-all available information diagnosis. We compared estimates of prevalence of cognitive impairment for these two outcomes to assess whether estimates of racial/ethnic differences were affected by use of demographically adjusted versus unadjusted neuropsychological test scores and diagnostician knowledge of demographic characteristics. For participants who did not receive an adjudicated diagnosis either by design or failure to complete intended assessments, we estimated cognitive impairment using cognitive variables (SENAS and NIHTB-CHB), core demographic variables (age, gender, education), and racial/ethnic group.

2.4.2 | Raw cognitive impairment prevalence

Raw cognitive impairment prevalence (overall and race/ethnicity subgroups) was estimated in the Random Selection group, calculated as number with adjudicated diagnoses of MCI or dementia divided by total number adjudicated.

2.4.3 | Diagnostic prediction model to estimate cognitive impairment prevalence

Model-derived estimates of cognitive impairment prevalence were developed for the blinded-adjusted and final-all available information conditions as follows. Step 1: A logistic regression model was estimated in those with an adjudicated diagnosis. Presence/absence of cognitive impairment was the dependent variable and independent variables included SENAS and NIHTB-CHB cognitive test scores, age, education, gender, race/ethnicity, and race/ethnicity by cognitive test score interactions. Step 2: The model estimated in Step 1 was used to calculate cognitive impairment probability for each individual without an adjudicated diagnosis. Missing values in independent variables were replaced with multiple imputation (25 imputations), and probability of impairment for each individual was averaged across the imputed data sets. Step 3: An aggregated probability of impairment was set as 1 for those diagnosed as impaired, 0 for those diagnosed as nor-

mal, and the estimated probability of impairment for those without an adjudicated diagnosis. Step 4: Group-specific prevalences were calculated as the sum of aggregated individual probabilities of impairment divided by group sample size. Step 5: Bootstrap sampling (10,000 draws with replacement) was used to establish 95% confidence intervals for prevalence estimates and pairwise differences between racial/ethnic groups.

2.4.4 | Attrition bias

We estimated cognitive impairment probability under the final-all information diagnosis condition for Random Selection group individuals using the logistic regression model described in Section 2.5.3. The probability of impairment was Blom transformed[33] to normalize and standardize the distribution. We used linear regression to compare impairment probability within each racial/ethnic group in cases with and without an adjudicated diagnosis.

2.4.5 | Demographic and health associations with cognitive impairment prevalence

We examined how age, education, and gender were related to aggregated probabilities of impairment for the two diagnosis conditions. We used linear regression to examine the associations of diabetes, hypertension, hypercholesterolemia, depression, and anxiety with probabilities of cognitive impairment within the Random Selection group. These are common health conditions that have been implicated as important for brain and cognitive health.

3 | RESULTS

3.1 | Sample characteristics

Table 1 reports sample characteristics stratified by racial/ethnic group—the full sample characteristics are in the upper half of the table, whereas the lower half refers to Random Selection individuals who completed clinical exams. Linear regression was used to examine racial/ethnic group differences in continuous variables and chi-square tests were used for categorical variables. Gender ($P = .001$), age ($P = .007$), education ($P = .001$), and all seven cognitive tests (P's $< .001$) differed across racial/ethnic groups in the full sample. Proportions of Random Selection individuals who completed planned clinical evaluations differed across groups ($P = .019$) and were higher in Asian and White participants than in Black and Latino participants. The prevalence of diabetes ($P = .002$) and hypertension ($P = .017$) differed across groups in the Random Selection sample, with the highest rates in Black and the lowest rates in White participants. Diabetes and hypertension rates are consistent with a recent population-based study.[34] The MoCA total score ($P = .001$) and CDR Sum of boxes ($P = .046$) differed across groups.

TABLE 1 Sample characteristics

|	Asian	Black	Latino	White	Total
Full Sample - mean (SD)	415 (100.0%)	443 (100.0%)	348 (100.0%)	503 (100.0%)	1,709 (100.0%)
Random/Screening - Random Selection	125 (30.1%)	129 (29.1%)	160 (46.0%)	127 (25.2%)	541 (31.7%)
Random/Screening - Screen Selection	290 (69.9%)	314 (70.9%)	188 (54.0%)	376 (74.8%)	1168 (68.3%)
Gender - female	221 (53.3%)	299 (67.5%)	205 (58.9%)	291 (57.9%)	1016 (59.4%)
Age (years) - mean (SD)	75.2 (±7.0)	75.1 (±7.1)	75.7 (±6.7)	76.6 (±7.5)	75.7 (±7.1)
Education (years) - mean (SD)	15.6 (±2.6)	14.2 (±2.8)	13.1 (±4.1)	15.2 (±2.9)	14.6 (±3.2)
Episodic Memory (SENAS) - mean (SD)	0.2 (±1.0)	-0.1 (±0.9)	-0.2 (±1.0)	0.1 (±1.0)	0.0 (±1.0)
Semantic Memory (SENAS) - mean (SD)	-0.2 (±1.0)	-0.5 (±0.8)	0.0 (±0.9)	0.6 (±0.9)	0.0 (±1.0)
Executive Function (SENAS) - mean (SD)	-0.2 (±0.9)	-0.3 (±0.9)	-0.2 (±0.9)	0.5 (±1.1)	0.0 (±1.0)
Flanker (NIHTB) - mean (SD)	0.2 (±1.0)	-0.4 (±1.0)	0.0 (±1.0)	0.2 (±0.9)	0.0 (±1.0)
Picture Sequence Memory (NIHTB) - mean (SD)	0.0 (±1.0)	-0.3 (±0.9)	0.1 (±0.9)	0.1 (±1.0)	0.0 (±1.0)
List Sorting (NIHTB) - mean (SD)	0.0 (±0.9)	-0.3 (±1.0)	0.0 (±0.9)	0.2 (±1.0)	0.0 (±1.0)
Picture Vocabulary (NIHTB) - mean (SD)	-0.1 (±1.0)	-0.4 (±0.9)	-0.2 (±0.9)	0.6 (±0.9)	0.0 (±1.0)
Random Selection - Clinical Exam	102 (81.6%)	90 (69.8%)	116 (72.5%)	106 (83.5%)	414 (76.5%)
Diabetes* - Yes	19 (18.6%)	30 (33.3%)	23 (19.8%)	13 (12.3%)	85 (20.5%)
Hypertension* - Yes	53 (52.0%)	62 (68.9%)	64 (55.2%)	52 (49.1%)	231 (55.8%)
Hypercholesteremia* - Yes	59 (57.8%)	52 (57.8%)	49 (42.2%)	58 (54.7%)	218 (52.7%)
Depression (last 2 years)* - Yes	1 (1.0%)	4 (4.4%)	6 (5.2%)	6 (5.7%)	17 (4.1%)
Anxiety* - Yes	2 (2.0%)	2 (2.2%)	8 (6.9%)	3 (2.8%)	15 (3.6%)
MoCA Total Score* - mean (SD)	23.1 (±4.4)	21.6 (±4.4)	22.5 (±3.8)	25.0 (±3.7)	23.1 (±4.3)
CDR Sum of Boxes* - mean (SD)	0.2 (±0.7)	0.5 (±0.9)	0.2 (±0.7)	0.2 (±0.6)	0.3 (±0.7)

Asians included 222 Chinese, 81 Japanese, 67 Filipino, 26 Southeast Asian, and 3 Pacific Islander; Latino included 169 Mexican, 66 Central/South American, 22 Caribbean, and 76 Other Latino. * = From Random Selection participants with completed clinical evaluations (N = 414).]
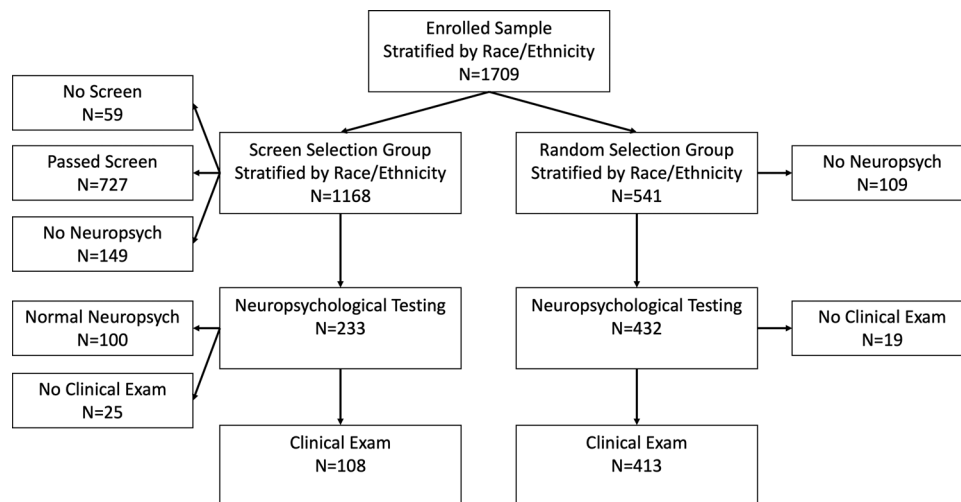
FIGURE 2 Study participation flow chart

3.2 | Recruitment and response rate

Figure 2 shows numbers of participants by assessment. For the Random Selection group, 109 (20.1%) declined consent for referral to UC Davis, declined a clinical evaluation by the UC Davis team, or could not be contacted after multiple attempts. Clinical neuropsychological testing was completed by 432 (79.9%) and 413 of the 432 (95.6%) completed a clinical exam. For the Screen Selection group, 59 (5.1%) did not complete the NIHTB-CHB screening measures, 149 (12.8%) declined clinical evaluation or could not be contacted by the UC Davis team, and
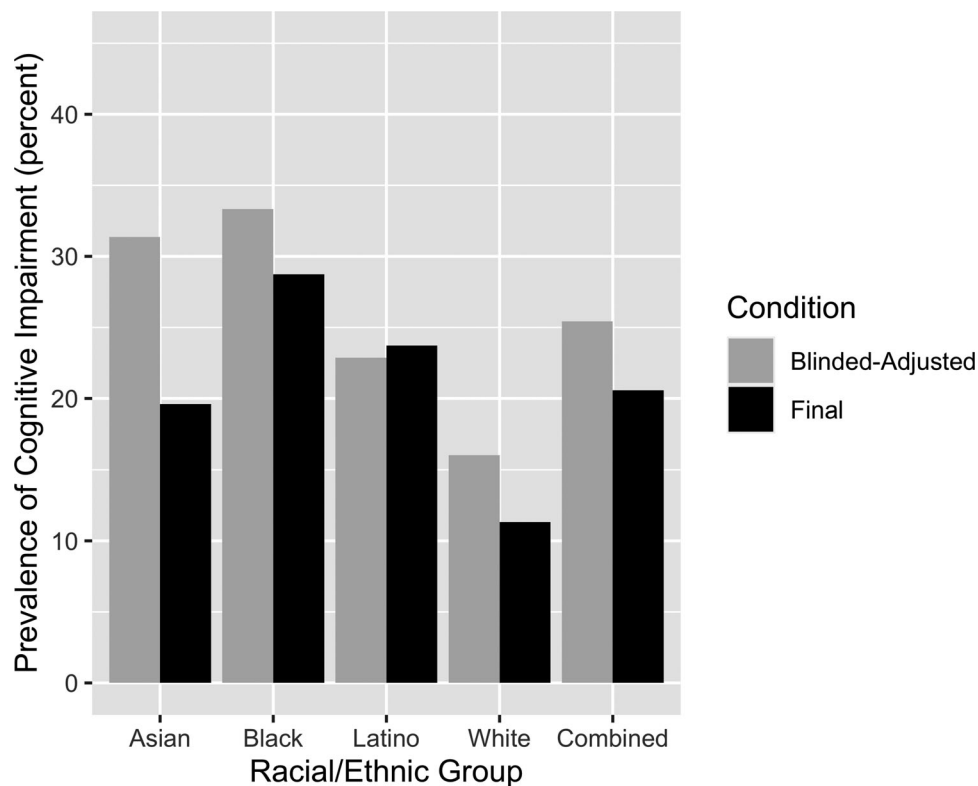
FIGURE 3   Raw prevalence of cognitive impairment by racial/ethnic group; Random Selection group (N = 413). Cognitive impairment = mild cognitive impairment or dementia, blinded-adjusted = diagnosis based on clinical exam results and demographically adjusted neuropsychological scores (adjusted for race/ethnicity, gender, and education), blind to the demographic characteristics of the individual. Final = diagnosis based on all available information (clinical exam results, demographically adjusted neuropsychological scores, unadjusted neuropsychological scores, and participant's demographic characteristics). (Estimated values are presented in Supplementary Table 2, Supplementary Materials)

727 (62.2%) had normal cognitive screen results. Two hundred thirty-three (20.0%) received clinical neuropsychological testing, 100 (42.9% of this group) had normal neuropsychological test results, 108 (46.4%) completed a clinical exam, and 25 (10.7%) did not complete a clinical exam.

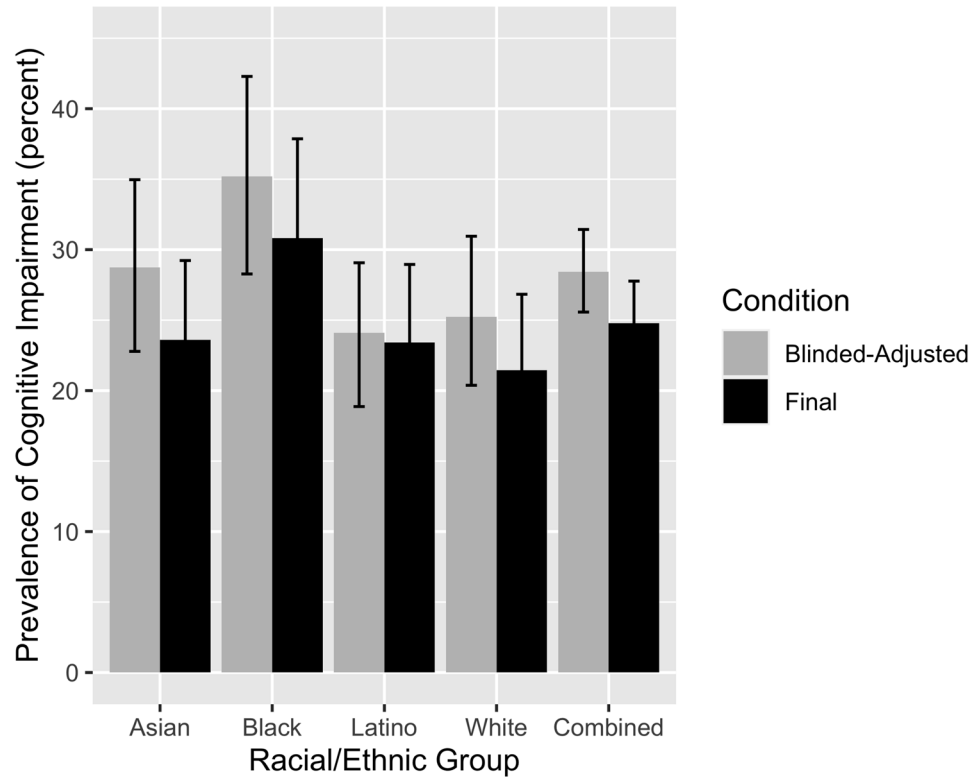### 3.3 | Prevalence of cognitive impairment based on adjudicated diagnoses

Fourteen individuals (0.8%) received a blinded-adjusted diagnosis of dementia and 13 (0.8%) had a final all-information available diagnosis of dementia (11 with dementia under both conditions); there were 184 (10.8%) adjudicated blinded-adjusted diagnoses of MCI and 153 (9%) final diagnoses of MCI. These numbers underestimate true numbers of cognitive impairment cases due to attrition from planned exams and imperfect sensitivity of the screening algorithm. However, the dementia prevalence in this sample was quite low, and subsequent results combine MCI and dementia into a single category of "cognitive impairment."

Figure 3 shows the prevalence of race/ethnicity-stratified cognitive impairment prevalence in the Random Selection group. The prevalence was higher for blinded-adjusted diagnosis than final-all available infor-

mation diagnosis, except in Latino participants. The prevalence was highest for Black participants and lowest for White participants for both conditions.

3.4 | Diagnostic prediction model estimates, full sample

Model-estimated prevalence rates in the full sample (Figure 4) showed similar patterns to raw prevalence estimates from the Random Selection group (Figure 3) in two important ways: (1) Cognitive impairment prevalence was highest among Black and lowest among White participants, (2) prevalences for the blinded-adjusted condition were higher for all groups except Latino participants. Model-estimated prevalences were 8% to 9% higher than raw estimates for White participants. Differences between raw and model-derived prevalences did not exceed 4% for the other three groups. The range of racial/ethnic group differences in model-estimated prevalence was smaller than for raw prevalences, which was largely due to higher model estimated prevalences for White participants. Group differences in prevalences is addressed in Figure 5, which shows bootstrap-estimated prevalence differences and 95% confidence intervals. Under both diagnostic conditions, cognitive impairment prevalence was significantly higher among Black

FIGURE 4 Diagnostic prediction model estimated prevalence of cognitive impairment and 95% confidence intervals by racial/ethnic group, all KHANDLE participants (n = 1709). Cognitive impairment = mild cognitive impairment or dementia, blinded-adjusted = diagnosis based on clinical exam results and demographically adjusted neuropsychological scores (adjusted for race/ethnicity, gender, and education), blind to demographic characteristics of the individual. Final = diagnosis based on all available information (clinical exam results, demographically adjusted neuropsychological scores, unadjusted neuropsychological scores, and participant's demographic characteristics). (Estimated values are presented in Supplementary Table 3, Supplementary Materials)

than White participants. Prevalence estimates based on the blinded-adjusted diagnosis were significantly higher in Black than Latino participants.

3.5 | Attrition bias

Impairment probability differed significantly between White participants in the Random Selection group who did and did not receive an adjudicated diagnosis (diagnosed mean = 0.167, standard error [SE] = 0.065, undiagnosed mean = 0.339, SE = 0.06, p = 0.005). For Black participants, average impairment probability was 0.086 lower for diagnosed cases than for undiagnosed cases, but this difference was not statistically significant (*P* = .065). For Asian and Latino participants, average impairment probabilities for diagnosed and undiagnosed cases differed by less than 0.021 (*P* > .488).

3.6 | Demographic and health associations with cognitive impairment

Age, education, and gender were related to cognitive impairment under both diagnosis conditions. For example, cognitive impairment

rates were five-fold higher among those 85 years or older (64–66%) versus 65- to 69-year-olds (10–12%). Estimated education effects were smaller for the blinded-adjusted condition (30% for less than high school vs 25% for college graduates) than the final-all information condition (29% vs 19%). Cognitive impairment prevalence was higher among male than female participants under both conditions (36% vs 31% blinded-adjusted; 33% vs 27% final-all information). Hypertension was the only health condition associated with higher probability of cognitive impairment in the Random Selection group. (More detailed results are available in the Supplementary Materials.)

4 | DISCUSSION

KHANDLE obtained comprehensive clinical assessments of older adults from four racial/ethnic groups selected from a common sampling frame. Cognitive impairment prevalence was 25% for the final diagnosis based on all available information and was slightly higher (28%) for blinded-adjusted diagnoses. Prevalence was highest among Black participants under both conditions (31% and 35%). Black-White prevalence differences were statistically significant for both conditions, and Black-Latino prevalence differences were statistically significant for the blinded-adjusted condition; all other racial/ethnic group pairwise
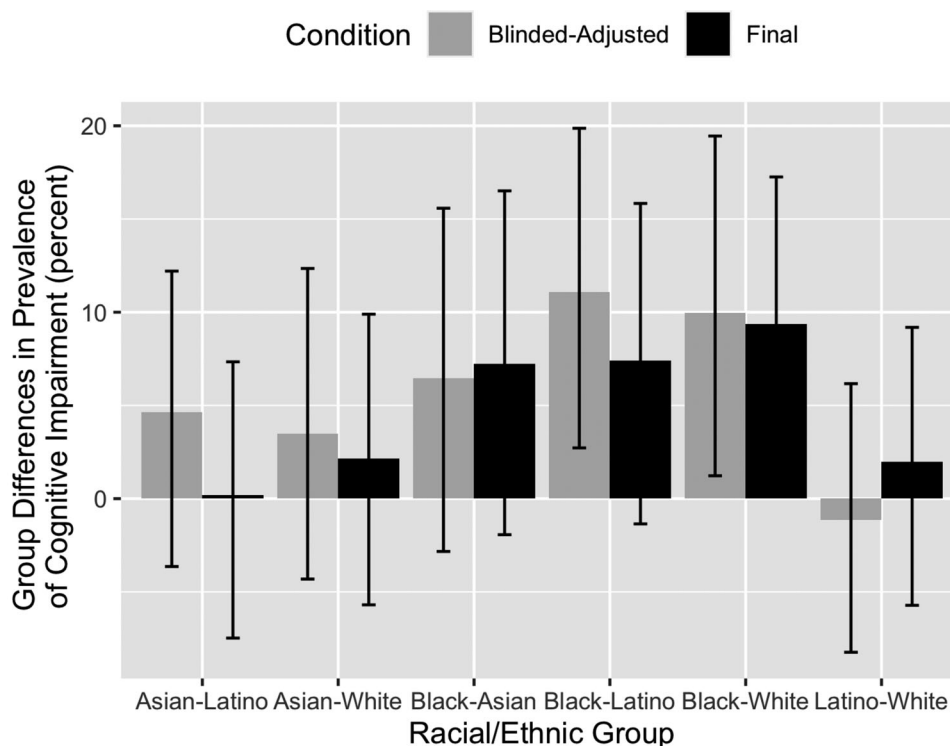
FIGURE 5 Diagnosis Prediction Model estimated racial/ethnic differences in prevalence of cognitive impairment and 95% confidence intervals; all KHANDLE participants (n = 1709). Group differences in prevalence are calculated as the prevalence in the first group minus the prevalence in the second group. Cognitive impairment = mild cognitive impairment or dementia, blinded-adjusted = diagnosis based on clinical exam results and demographically adjusted neuropsychological scores (adjusted for race/ethnicity, gender, and education), blind to demographic characteristics of the individual. Final = diagnosis based on all available information (clinical exam results, demographically adjusted neuropsychological scores, unadjusted neuropsychological scores, and participant's demographic characteristics). (Estimated values are presented in Supplementary Table 4, Supplementary Materials)

comparisons were not statistically significant. White participants who failed to complete clinical evaluations had significantly higher cognitive impairment probability than those who were evaluated.

There is a large volume of literature on racial/ethnic differences in prevalence of dementia and cognitive impairment, with substantial variability in methods used across studies.[35–52] This study predominantly addresses MCI because there were few cases with dementia. MCI prevalence varies substantially as a function of how MCI is defined and operationalized but also varies across samples using similar definitions.[53,54] Prevalence estimates in this study showed substantial cognitive impairment (>20%) in all four racial/ethnic groups, and these rates are consistent with prevalence estimates in studies using similarly broad definitions of MCI.[53] Cognitive impairment rates in this study also were similar to MCI rates from a previous multi-racial/ethnic study[44] that reported 28% prevalence of MCI, with similar rates among Black, Latino, and White participants. As expected, cognitive impairment was strongly associated with age. Higher education was associated with less cognitive impairment. This association was less pronounced for the blinded-adjusted diagnosis condition, which is not surprising, since education is related to cognitive test scores in cognitively normal individuals,[55] and education-normed performance was used under the blinded-adjusted condition. Dementia prevalence was ≈1%.

Because a medical record dementia diagnosis excluded enrollment, this suggests a low prevalence of undetected dementia cases in KPNC.

We examined whether methodological differences in how diagnoses were obtained influenced group differences in cognitive impairment prevalence. The blinded-adjusted diagnosis condition controlled for normative differences in test scores across groups and was designed to limit implicit bias. The final-all information diagnosis condition gave clinicians access to the most comprehensive information, but there was also greater exposure to potential biases. Final-all information prevalence estimates were lower than blinded-adjusted condition estimates in all groups except for Latinos. It is important to note that prevalence differences across groups were very similar for the two conditions. This suggests that test score norms and knowledge of participant demographic characteristics did not differentially affect prevalence estimates across racial/ethnic groups. However, adjusting for failure to complete planned assessments resulted in higher estimated prevalence in White participants but not in other groups; conversely, not adjusting for attrition amplified differences between White participants and other groups.

The literature on how different biases affect estimates of cognitive impairment prevalence is limited. The impact of using race/ethnicity group-specific neuropsychological test norms has received the most attention. There is evidence that cognitively normal individuals from

racial/ethnic minority groups have lower average test scores, and that failure to account for normative differences can lead to high rates of false-positive diagnoses of cognitive impairment.[11,13] But there also is concern that group-specific norms might lead to higher false-negative rates in groups that have lower normative scores.[14,15] The effects of types of test norms on diagnosis from a comprehensive clinical evaluation have not been well studied. Similarly, the effects of implicit bias and attrition bias on the diagnosis of cognitive impairment have received limited attention.

Type of test norms and implicit bias did not appear to affect group differences in cognitive impairment in this study, but this remains an important question for future research. Test norms in this study were based on the KHANDLE sample in which 29% were non-Latino Whites, and consequently, demographically adjusted and unadjusted test scores were being compared to an unusually diverse normative sample. Norms used in clinical practice often are predominantly based on non-Latino Whites, and even for instruments like the NIHTB CHB that have population-representative norming, White individuals account for more than 70% of the normative sample. This may have been a factor contributing to the lack of an effect of the type of test norms in our study. Diagnosticians in this study were three clinicians who had extensive, shared experience working with a racially/ethnically diverse populations in a dementia research setting. It will be important to address diagnostic bias in clinicians from more diverse settings with different levels of training and experience.

4.1 | Strengths and limitations

This study has important strengths. KHANDLE includes four racial/ethnic groups from a common sampling frame. It conducted gold standard, comprehensive assessments with a randomly selected subgroup, augmented by a two-stage design applied to the rest of the sample to identify those at highest risk for cognitive impairment. KHANDLE participants received additional, high-quality cognitive assessments (SENAS and NIHTB-CHB) that were excluded from diagnosis. We were able to use this additional cognitive data to estimate probability of cognitive impairment in those who did not receive full clinical evaluations. Individuals who failed to complete clinical evaluations still contributed to the prevalence estimates thus reducing the effects of attrition bias.

There also are relevant weaknesses. Although the KHANDLE sample is diverse, all participants have had consistent health care access over many decades, which is not the norm for the population of northern California. A second important limitation is that diagnosticians were a small sample of convenience. Third, only two diagnosis conditions (blinded-adjusted and final all-information) were examined and adjusted test scores were available under both conditions. This design does not permit definitive comparisons of unadjusted versus adjusted test scores. Finally, we did not directly measure implicit bias in the diagnosticians.

4.2 | Implications for future studies

This study provides a foundation and demonstrates an approach for future work to examine sources of bias in a more systematic and refined manner. It will be important to evaluate these influences in a more heterogenous sample of diagnosticians and to broaden the types of norms evaluated. We plan future studies within KHANDLE that will include larger numbers of diagnosticians with greater heterogeneity in training and experience. We will compare rates of cognitive impairment when clinicians have access to only demographically adjusted versus only unadjusted scores, and will vary whether norms come from diverse or predominantly non-Latino White samples. We will also directly measure implicit bias in individual clinicians by providing systematically different demographic characteristics for the same observed clinical data and tabulating differences in diagnostic outcomes. This will facilitate examining how individual-level implicit bias influences diagnostic decisions. The current study represents an important step toward unbiased diagnosis, and ultimately, better understanding of racial/ethnic differences in late-life cognitive health.

CONFLICTS OF INTEREST

The authors have no conflicts of interest related to this work.

REFERENCES

1. Vespa J, Medina L, Armstrong DM. Demographic turning points for the United States: population projections for 2020 to 2060 population estimates and projections current population reports. *US Census Bureau*. 2020:1-15.
2. Chen C, Zissimopoulos JM. Racial and ethnic differences in trends in dementia prevalence and risk factors in the United States. *Alzheimer's Dement Transl Res Clin Interv*. 2018;4:510-520. https://doi.org/10.1016/j.trci.2018.08.009.
3. Babulal GM, Quiroz YT, Albensi BC, et al. Perspectives on ethnic and racial disparities in Alzheimer's disease and related dementias: update and areas of immediate need. *Alzheimer's Dement*. 2019;15(2):292-312. https://doi.org/10.1016/j.jalz.2018.09.009.
4. Weuve J, Barnes LL, Mendes de Leon CF, et al. Cognitive aging in black and white Americans: cognition, cognitive decline, and incidence of Alzheimer disease dementia. *Epidemiology (Cambridge, Mass)*. 2018;29(1):151-159. https://doi.org/10.1097/EDE.0000000000000747.
5. Mayeda ER, Glymour MM, Quesenberry CP, Whitmer RA. Inequalities in dementia incidence between six racial and ethnic groups over 14 years. *Alzheimer's Dement*. 2016;12(3):216-224. https://doi.org/10.1016/j.jalz.2015.12.007.
6. Díaz-Venegas C, Downer B, Langa KM, Wong R. Racial and ethnic differences in cognitive function among older adults in the USA. *Int J Geriatr Psychiatry*. 2016;31(9):1004-1012. https://doi.org/10.1002/gps.4410.
7. Ng T-P, Niti M, Chiam P-C, Kua E-H. Ethnic and educational differences in cognitive test performance on mini-mental state examination

in Asians. *Am J Geriatr Psychiatry*. 2007;15(2):130-139. https://doi.org/10.1097/01.JGP.0000235710.17450.9a.

8. Bohnstedt M, Fox PJ, Kohatsu ND. Correlates of mini-mental status examination scores among elderly demented patients: the influence of race-ethnicity. *J Clin Epidemiol*. 1994;47(12):1381-1387. https://doi.org/10.1016/0895-4356(94)90082-5.

9. Mungas D, Reed BR, Crane PK, Haan MN, Gonzalez H. Spanish and English Neuropsychological Assessment Scales (SENAS): further development and psychometric characteristics. *Psychol Assess*. 2004;16(4):347-359. https://doi.org/10.1037/1040-3590.16.4.347.

10. Lucas JA, Ivnik RJ, Willis FB, et al. Mayo's Older African Americans Normative Studies: normative data for commonly used clinical neuropsychological measures. *Clinical Neuropsychologist*. 2005;19(2):162-183. https://doi.org/10.1080/13854040590945265.

11. Gasquoine PG. Race-Norming of Neuropsychological Tests. *Neuropsychol Rev*. 2009;19:250-262. https://doi.org/10.1007/s11065-009-9090-5.

12. Manly JJ, Jacobs DM. Future directions in neuropsychological assessment with African Americans. In: Ferraro FR, ed. *Minority and Cross-Cultural Aspects of Neuropsychological Assessment*. Lisse, Netherlands: Swets and Zeitlinger; 2001:79-96.

13. Werry AE, Daniel M, Bergström B. Group differences in normal neuropsychological test performance for older non-Hispanic White and Black/African American adults. *Neuropsychology*. 2019;33(8):1089-1100. https://doi.org/10.1037/neu0000579.

14. Brandt J. 2005 INS Presidential Address: neuropsychological Crimes and Misdemeanors. *Clin Neuropsychol*. 2007;21:553-568. https://doi.org/10.1080/13854040600770800.

15. Manly JJ. Advantages and disadvantages of separate norms for African Americans. *Clin Neuropsychol*. 2005;19(2):270-275. https://doi.org/10.1080/13854040590945346.

16. Manly JJ, Echemendia RJ. Race-specific norms: using the model of hypertension to understand issues of race, culture, and education in neuropsychology. *Arch Clin Neuropsychol*. 2007;22(3):319-325.

17. FitzGerald C, Hurst S. Implicit bias in healthcare professionals: a systematic review. *BMC Med Ethics*. 2017;18(1):19. https://doi.org/10.1186/s12910-017-0179-8.

18. Gershon RC, Wagster MV, Hendrie HC, Fox NA, Cook KF, Nowinski CJ. NIH toolbox for assessment of neurological and behavioral function. *Neurology*. 2013;80(11 Suppl 3):S2-S6. https://doi.org/10.1212/WNL.0b013e3182872e5f.

19. Weintraub S, Dikmen SS, Heaton RK, et al. Cognition assessment using the NIH Toolbox. *Neurology*. 2013;80(11 Suppl 3):S54-64. https://doi.org/10.1212/WNL.0b013e3182872ded.

20. Weintraub S, Dikmen SS, Heaton RK, et al. The cognition battery of the NIH toolbox for assessment of neurological and behavioral function: validation in an adult sample. *Journal of the International Neuropsychological Society*. 2014;20(6):567-578. https://doi.org/10.1017/S1355617714000320.

21. Besser L, Kukull W, Knopman DS, et al. Version 3 of the National Alzheimer's Coordinating Center's Uniform Data Set. *Alzheimer disease and associated disorders*. 2018;32(4):351-358. https://doi.org/10.1097/WAD.0000000000000279.

22. Morris JC, Heyman A, Mohs RC, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's Disease. *Neurology*. 1989;39:1159-1165.

23. Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*. 2005;53(4):695-699. https://doi.org/10.1111/j.1532-5415.2005.53221.x.

24. Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology*. 1993;43:2412-2414.

25. Hinton L, Carter K, Reed BR, et al. Recruitment of a community-based cohort for research on diversity and risk of dementia. *Alzheimer Dis Assoc Disord*. 2010;24(3):234-241. https://doi.org/10.1097/WAD.0b013e3181c1ee01.

26. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20(1):37-46. https://doi.org/10.1177/001316446002000104.

27. 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.

28. Mungas D, Reed BR, Marshall SC, Gonzalez HM. Development of psychometrically matched English and Spanish language neuropsychological tests for older persons. *Neuropsychology*. 2000;14(2):209-223.

29. Mungas D, Reed BR, DeCarli C. Criterion-Referenced validity of a neuropsychological test battery: equivalent performance in elderly Hispanics and non-Hispanic Whites. *J Int Neuropsychol Soc*. 2005;11:620-630.

30. Mungas D, Beckett L, Harvey D, et al. Heterogeneity of cognitive trajectories in diverse older persons. *Psychol Aging*. 2010;25(3). https://doi.org/10.1037/a0019502.

31. Early DR, Widaman KF, Harvey D, et al. Demographic predictors of cognitive change in ethnically diverse older persons. *Psychol Aging*. 2013;28(3):633-645. https://doi.org/10.1037/a0031645.

32. Mungas D, Widaman KF, Reed BR. Tomaszewski Farias S. Measurement invariance of neuropsychological tests in diverse older persons. *Neuropsychology*. 2011;25(2):260-269. https://doi.org/10.1037/a0021090.

33. Blom G. *Statistical Estimates and Transformed Beta-Variables*. New York: Wiley; 1958.

34. Ware EB, Morataya C, Fu M, Bakulski KM. Type 2 diabetes and cognitive status in the Health and Retirement Study: a Mendelian randomization approach. *Front Genet*. 2021;12:634767. https://doi.org/10.3389/fgene.2021.634767.

35. Gelber RP, Launer LJ, White LR. The Honolulu-Asia Aging Study: epidemiologic and neuropathologic research on cognitive impairment. *Curr Alzheimer Res*. 2012;9(6):672. https://doi.org/10.2174/156720512801322618.

36. Matthews FE, Arthur A, Barnes LE, et al. A two-decade comparison of prevalence of dementia in individuals aged 65 years and older from three geographical areas of England: results of the cognitive function and ageing study I and II. *Lancet*. 2013;382(9902):1405-1412. https://doi.org/10.1016/S0140-6736(13)61570-6.

37. González HM, Tarraf W, Fornage M, et al. A research framework for cognitive aging and Alzheimer's disease among diverse US Latinos: design and implementation of the Hispanic Community Health Study/Study of LatinosInvestigation of Neurocognitive Aging (SOL-INCA). *Alzheimer's Dement*. 2019;15(12):1624-1632. https://doi.org/10.1016/j.jalz.2019.08.192.

38. Hall KS, Gao S, Baiyewu O, et al. Prevalence rates for dementia and Alzheimer's disease in African Americans: 1992 versus 2001. *Alzheimer's Dement*. 2009;5(3):227-233. https://doi.org/10.1016/j.jalz.2009.01.026.

39. Hendrie HC, Osuntokun BO, Hall KS, et al. Prevalence of Alzheimer's disease and dementia in two communities: nigerian Africans and African Americans. *Am J Psychiatry*. 1995;152(10):1485-1492. https://doi.org/10.1176/ajp.152.10.1485.

40. Graves AB, Larson EB, Edland SD, et al. Prevalence of dementia and its subtypes in the Japanese American Population of King County, Washington State: the Kame Project. *Am J Epidemiol*. 1996;144(8):760-771. https://doi.org/10.1093/oxfordjournals.aje.a009000.

41. Haan MN, Mungas DM, Gonzalez HM, Ortiz TA, Acharya A, Jagust WJ. Prevalence of dementia in older Latinos: the influence of type 2 diabetes mellitus, stroke and genetic factors. *J Am Geriatr Soc*. 2003;51(2):169-177. https://doi.org/10.1046/j.1532-5415.2003.51054.x.

42. Carrión-Baralt JR, Suárez-Pérez E, del Rio R, Moore K, Silverman JM. Prevalence of dementia in Puerto rican veterans is higher than among Mainland U.S. Veterans. *J Am Geriatr Soc.* 2010;58(4):798-799. https://doi.org/10.1111/j.1532-5415.2010.02789.x.

43. Galasko D, Salmon D, Gamst A, et al. Prevalence of dementia in Chamorros on Guam: relationship to age, gender, education, and APOE. *Neurology.* 2007;68(21):1772-1781. https://doi.org/10.1212/01.wnl.0000262028.16738.64.

44. Manly JJ, Bell-McGinty S, Tang M-X, Schupf N, Stern Y, Mayeux R. Implementing diagnostic criteria and estimating frequency of mild cognitive impairment in an urban community. *Arch Neurol.* 2005;62(11):1739-1746. https://doi.org/10.1001/archneur.62.11.1739.

45. Bienias JL, Beckett LA, Bennett DA, Wilson RS, Evans DA. Design of the Chicago Health and Aging Project (CHAP). *J Alzheimer's Dis.* 2003;5(5):349-355. https://doi.org/10.3233/JAD-2003-5501.

46. Kuller LH, Lopez OL, Becker JT, Chang Y, Newman AB. Risk of dementia and death in the long term follow up of the Pittsburgh CHS Cognition Study. *Alzheimers Dement.* 2016;12(2):170-183. https://doi.org/10.1016/j.physbeh.2017.03.040.

47. Knopman DS, Gottesman RF, Sharrett AR, et al. Mild cognitive impairment and dementia prevalence: the Atherosclerosis Risk in Communities Neurocognitive Study. Alzheimer's Dement Diagnosis. *Assess Dis Monit.* 2016;2:1-11. https://doi.org/10.1016/j.dadm.2015.12.002.

48. Trittschuh EH, Crane PK, Larson EB, et al. Effects of varying diagnostic criteria on prevalence of mild cognitive impairment in a community based sample. *J Alzheimers Dis.* 2011;25(1):163-173. https://doi.org/10.1161/CIRCULATIONAHA.110.956839.

49. Langa KM, Plassman BL, Wallace RB, et al. The aging, demographics, and memory study: study design and methods. *Neuroepidemiology.* 2005;25:181-191. https://doi.org/10.1159/000087448.

50. Demirovic J, Prineas R, Loewenstein D, et al. Prevalence of dementia in three ethnic groups: the South Florida program on aging and health. *Ann Epidemiol.* 2003;13(6):472-478. https://doi.org/10.1016/S1047-2797(02)00437-4.

51. Fillenbaum GG, Heyman A, Huber MS, et al. The prevalence and 3-year incidence of dementia in older Black and White community residents. *J Clin Epidemiol.* 1998;51(7):587-595. https://doi.org/10.1016/S0895-4356(98)00024-9.

52. Bennett DA, Schneider JA, Buchman AS, Barnes LL, Boyle PA, Wilson RS. Overview and findings from the rush Memory and Aging Project. *Curr Alzheimer Res.* 2012;9(6):663. https://doi.org/10.2174/156720512801322663.

53. Petersen RC, Lopez O, Armstrong MJ, et al. Practice guideline update summary: mild cognitive impairment. *Neurology.* 2018;90(3):126-135. https://doi.org/10.1212/WNL.0000000000004826.

54. Ward A, Arrighi HM, Michels S, Cedarbaum JM. Mild cognitive impairment: disparity of incidence and prevalence estimates. *Alzheimer's & Dementia.* 2012;8(1):14-21. https://doi.org/10.1016/j.jalz.2011.01.002.

55. Ganguli M, Snitz BE, Lee C-W, Vanderbilt J, Saxton JA, Chang C-CH. Age and education effects and norms on a cognitive test battery from a population-based cohort: the MonongahelaYoughiogheny Healthy Aging Team. *Aging Ment Health.* 2010;14(1):100-107. https://doi.org/10.1080/13607860903071014.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.