Diagnosing Mild Cognitive Impairment Among Racially Diverse Older Adults: Comparison of Consensus, Actuarial, and Statistical Methods

Kathryn N. Devlin[a], **Laura Brennan**[b], **Laura Saad**[c], **Tania Giovannetti**[d], **Roy H. Hamilton**[e,f], **David A. Wolk**[e,f], **Sharon X. Xie**[e,g], **Dawn Mechanic-Hamilton**[e,f,*]

[a]Department of Psychological and Brain Sciences, Drexel University, Philadelphia, PA, USA

[b]Department of Neurology, Thomas Jefferson University Hospital, Philadelphia, PA, USA

[c]Department of Psychology, Rutgers University, New Brunswick, NJ, USA

[d]Department of Psychology, Temple University, Philadelphia, PA, USA

[e]Alzheimer's Disease Research Center, University of Pennsylvania, Philadelphia, PA, USA

[f]Department of Neurology, University of Pennsylvania, Philadelphia, PA, USA

[g]Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania, Philadelphia, PA, USA

Abstract

Background: Actuarial and statistical methods have been proposed as alternatives to conventional methods of diagnosing mild cognitive impairment (MCI), with the aim of enhancing diagnostic and prognostic validity, but have not been compared in racially diverse samples.

Objective: We compared the agreement of consensus, actuarial, and statistical MCI diagnostic methods and their relationship to race and prognostic indicators, among diverse older adults.

Methods: Participants (N = 354; M age = 71; 68% White, 29% Black) were diagnosed with MCI or normal cognition (NC) according to clinical consensus, actuarial neuropsychological criteria (Jak/Bondi), and latent class analysis (LCA). We examined associations with race/ethnicity, longitudinal cognitive and functional change, and incident dementia.

Results: MCI rates by consensus, actuarial criteria, and LCA were 44%, 53%, and 41%, respectively. LCA identified three MCI subtypes (memory; memory/language; memory/executive) and two NC classes (low normal; high normal). Diagnostic agreement was substantial, but agreement of the actuarial method with consensus and LCA was weaker than the agreement between consensus and LCA. Among cases classified as MCI by actuarial criteria only, Black participants were over-represented, and outcomes were generally similar to those of NC

[*]Correspondence to: Dawn Mechanic-Hamilton, Penn Memory Center, 3400 Civic Center Boulevard, Philadelphia, PA 19104, USA. Tel.: +215 662 4516; dawn.mechanic@pennmedicine.upenn.edu.

participants. Consensus diagnoses best predicted longitudinal outcomes overall, whereas actuarial diagnoses best predicted longitudinal functional change among Black participants.

Conclusion: Consensus diagnoses optimize specificity in predicting dementia, but among Black older adults, actuarial diagnoses may be more sensitive to early signs of decline. Results highlight the need for cross-cultural validity in MCI diagnosis and should be explored in community- and population-based samples.

Keywords

Actuarial analysis; Alzheimer's disease; cognitive aging; early diagnosis; latent class analysis; memory disorders; mild cognitive impairment; minority health; race factors; statistical models

INTRODUCTION

Cognitive decline in Alzheimer's disease and related disorders (ADRD) is a leading cause of morbidity, creating a large caregiving and financial burden [1]. Early and accurate detection of cognitive change has the potential to allow patients, families, and clinicians to intervene early in the course of the disease and to better manage care plans as there is further decline in cognition and function. Additionally, it is imperative that early detection of cognitive change be accurate across racial and ethnic groups to ensure equitable access to care resources and interventions. Conventional approaches to diagnosis include a complete history, neurologic exam, cognitive testing, care partner report, and the clinical judgment of a provider. The standard practice in memory specialty clinics and many research studies is to use a **consensus** approach to diagnosis, in which this comprehensive information is discussed by a group of experts. This approach categorizes patients based on predefined criteria, typically following previously published standards for defining typical aging, cognitive change consistent with mild cognitive impairment (MCI) [2], or cognitive and functional impairment consistent with ADRDs [3]. While this approach is commonly used, it is not uniformly applied across memory centers.

Although objective assessment of cognition is critical to determine whether reported change is consistent with typical aging versus an early manifestation of an ADRD, incorporation of neuropsychological data into the diagnostic decision-making process varies regarding which domains and tests to include, whether to apply a strict cut-off score, which cut-off score to select, which baseline and demographic variables to account for, and how to define MCI subtypes. These inconsistencies in criteria used to diagnose MCI result in imprecision in the prevalence, neuropathological characteristics, and longitudinal outcomes of individuals with MCI across centers [4–6]. Perhaps most notably, conventional criteria appear susceptible to false positive error [4, 7, 8]. These errors can impede appropriate treatment planning and reduce the ability of clinical trials to identify effective therapies. Thus, clinicians, researchers, and patients alike would benefit from a diagnostic approach that is reliable across sites and valid with regard to representing a disease state and predicting future change.

Consequently, several alternative approaches for reliably identifying normal and abnormal cognitive functioning in aging have arisen. One approach is to use a comprehensive set

of tests that evaluates a wide range of cognitive domains and uses consistent and robust criteria for impairment. This type of **actuarial** approach which prescribes cut-off scores (such as > 1 SD on two tests in a domain) has been described by Jak and Bondi [9–11]. It is designed to be less susceptible to errors in subjective report or clinical judgment and has been shown to reduce diagnostic error in comparison to conventional criteria, thereby strengthening associations with biomarkers and outcomes [9, 11–15]. As it does not require clinical judgment or additional clinical data, it is also designed to be more efficient and cost-effective than consensus approaches. On the other hand, the reliance on cognitive test scores alone might omit important context for interpreting whether an abnormal result truly reflects decline, leading to higher false positive error rates [16]. In particular, the actuarial approach might lead to greater false positive errors in older adults of color, who tend to obtain lower cognitive test scores due to factors such as education discrimination and bias [17].

Another approach to identifying MCI and its subgroups uses **statistical** methods such as cluster analysis [9, 12, 18–29] and mixture modeling (i.e., latent class analysis, LCA; and latent profile analysis, LPA) [30–40]. Cluster analysis and mixture modeling methods have similar goals, but mixture modeling approaches offer advantages over cluster analysis, including probabilistic rather than categorical group assignment, fit statistics to compare competing models, and the ability to handle missing data and both categorical and continuous variables simultaneously [41]. Results from cluster analysis and mixture modeling studies have provided some validation for conventional subtypes by identifying single-domain amnestic, non-amnestic, and multi-domain profiles [23, 30]. In line with findings from conventional subtypes, single-domain amnestic groups have been associated with biomarkers of Alzheimer's disease (AD) and progression from MCI to AD [12, 24, 32, 37, 38, 42], and non-amnestic groups have been associated with non-AD dementias and pathologies such as vascular dementia and frontotemporal dementia [22, 31, 37, 42], while multi-domain MCI has demonstrated the greatest neuropathological burden and risk of progression [9, 20, 24–26, 29, 32, 33, 43]. Moreover, statistical methods have identified MCI subtypes that are not included in conventional classification systems [44], including distinct subgroups within single-domain non-amnestic MCI (e.g., executive, language, and visuospatial) [9, 12, 19, 21, 39, 40] and within multi-domain amnestic MCI (e.g., memory/naming and memory/executive) [25]. Notably, these distinct profiles have unique neuropathological and prognostic characteristics.

Additionally, statistical modeling methods have the ability to reduce both false positive and negative diagnoses. Actuarial approaches and, in many instances, conventional approaches rely on cutoff scores to artificially dichotomize neurocognitive performance as intact or impaired, which is particularly challenging in diverse samples in which normative data highly influence such classifications. Statistical methods, by comparison, can be used with or without normative adjustments, and they account for the full range of performance, potentially improving diagnostic precision. By considering overall performance profiles, they may be better able to determine whether a low score reflects a non-pathological low performance pattern or a relative weakness indicative of pathological decline. Statistical methods have identified large groups of patients who were diagnosed with

MCI by conventional methods yet are comparable to cognitively intact individuals in their neuropsychological performance, neuropathologic profiles, and longitudinal outcomes [9, 12, 20, 25, 26, 29, 30, 33, 36, 45]. Conversely, statistical methods have detected individuals with subclinical dysfunction who would be classified as intact by other methods, but over time have an increased risk of progression to dementia [12, 15, 19, 24, 31, 34, 46].

Altogether, statistical modeling methods provide opportunities to validate conventional diagnostic and subtyping methods, and to identify more appropriate alternatives, which may improve examinations of disease mechanisms and prognosis. While both actuarial and statistical diagnostic approaches offer advantages over conventional diagnostic approaches, it remains unclear which approach offers the greatest diagnostic and prognostic advantage, as few studies have directly compared conventional, actuarial, and statistical approaches [9, 12]. The relative accuracy of these methods across racial and ethnic groups is also uncertain, as most studies comparing these methods have utilized overwhelmingly White samples [9, 12–15].

The aims of this study were to use conventional (i.e., consensus), actuarial, and statistical (i.e., LCA) approaches to 1) identify individuals with normal cognition and MCI subtypes, 2) compare prevalence and diagnostic agreement, and 3) examine longitudinal outcomes and associations with race/ethnicity. This study has several advantages over past approaches that have compared the three diagnostic methods. First, the cohort includes a more racially diverse sample, which will improve cross-cultural validity and generalizability. Secondly, it applies all three diagnostic methods to individuals with MCI and those with normal cognition in order to examine both false negative and false positive diagnostic errors. Thirdly, it utilizes latent class analysis, which has several advantages over cluster analysis.

METHODS

Participants

The present sample was comprised of 354 older adults who completed cognitive testing as part of the Alzheimer's Disease Center Clinical Core at the University of Pennsylvania in Philadelphia, Pennsylvania [47]. Participants were enrolled through a combination of clinician referral, self-referral, and community recruitment. They included individuals presenting with memory concerns and those interested in participating in research as healthy volunteers. Participants completed an annual evaluation including assessment of personal and family history, neurologic and general physical examinations, neuropsychological assessment, and study partner report of everyday cognitive and functional ability. Written informed consent was obtained from each participant prior to completion of study procedures. Procedures were done in accordance with the ethical standards of the University of Pennsylvania School of Medicine Institutional Review Board.

Inclusion criteria for the present study were 1) complete data for all 14 cognitive measures of interest (see below) at one or more visits, the earliest of which was classified as the baseline visit; 2) age 55 or older at baseline; 3) English as primary language; and 4) clinical consensus diagnosis of normal cognition (NC) or MCI at baseline. Participants with any other primary baseline diagnosis were excluded.

Cognitive assessments

All participants completed the Uniform Data Set prescribed by the National Alzheimer's Coordinating Center and additional cognitive tests. The Wechsler Test of Adult Reading (WTAR) [48] was used as a measure of literacy and academic achievement [49, 50]. Cognitive performance across six domains was assessed with the following measures: 1) *learning*: Wechsler Memory Scale–Revised (WMS-R) Logical Memory IA [51], Consortium to Establish a Registry for Alzheimer's Disease (CERAD) Word List Learning [52]; 2) *memory*: WMS-R Logical Memory IIA [51], CERAD Word List Recall and Recognition [52]; 3) *language*: animal fluency [53], Boston Naming Test 30-item version [54]; 4) *visuo-construction*: CERAD Constructional Praxis Test [52], Clock Drawing Test [53]; 5) *attention*: Trail Making Test A [55], Wechsler Adult Intelligence Scale–Revised (WAIS-R) Digit Span Forward [56]; and 6) *executive function*: Trail Making Test B [55], WAIS-R Digit Span Backward [56], WAIS-R Digit Symbol [56]. For all primary analyses, raw scores were converted to z-scores adjusted for age, education, and gender [57]. A second set of analyses was conducted with z-scores adjusted for age, education, gender, and race [57].

Diagnostic procedures

Consensus diagnosis—Clinical diagnosis was determined by the consensus of at least two clinicians using all available clinical information, including demographics, comprehensive neuropsychological data, functional status, medical and psychiatric history, and neuroimaging. MCI was diagnosed in the presence of 1) subjective report of cognitive decline, 2) an impaired score, accounting for demographics and background, on at least one cognitive test, and 3) preservation of everyday functioning [58, 59]. Participants who did not meet these criteria were classified as having NC. Participants with MCI were further classified as having single-domain or multi-domain amnestic MCI (aMCI) or single-domain or multi-domain non-amnestic MCI (naMCI).

Actuarial diagnosis—Actuarial diagnoses were made using the "comprehensive neuropsychological criteria" approach [9, 10]. Participants were classified as having MCI if they had an impaired score, defined as > 1 SD below the normative mean, on two or more measures within a cognitive domain. If this criterion was not met, participants were classified as having NC. Participants with MCI were further classified into the subtypes described above.

Statistical diagnosis—Statistically based diagnoses were made using LCA. LCA identifies, within heterogeneous groups, relatively homogeneous subgroups that demonstrate similar patterns of characteristics (e.g., cognitive scores). Cognitive domain scores, rather than individual test scores, were entered as indicators in LCA to ensure sufficient statistical power. Domain z-scores were calculated by averaging the demographically adjusted z-scores of all tests within the domain. A one-class model was examined, and the number of classes was then increased one class at a time until there was no additional improvement in model fit. Fit statistics included Akaike Information Criterion (AIC), consistent AIC (CAIC), Bayesian Information Criterion (BIC), and sample size-adjusted BIC (aBIC) [60–63], for which smaller values indicate better model fit. The CAIC and BIC were given the most weight in model selection [64]. Log likelihood difference values and likelihood ratio

tests were used to investigate whether the k class model better fit the data than the k-1 class model. Entropy was used as a measure of classification accuracy, with higher values indicating greater accuracy. The optimal model was determined based on a preponderance of evidence, including fit statistics, theoretical interpretation, and empirical support.

Longitudinal outcome measures

Longitudinal outcomes in global cognition, everyday functioning, and incident dementia were examined at baseline and one or more follow-up time points, when available. Global cognition was assessed with the Mini-Mental State Examination (MMSE) [65] or the Montreal Cognitive Assessment (MoCA), with MoCA scores converted to MMSE scores [66]. Everyday functioning was assessed with the Dementia Severity Rating Scale (DSRS), an informant-report questionnaire on which higher scores indicate greater functional difficulty [67]. Follow-up clinical consensus diagnoses were dichotomized as dementia or no dementia.

Statistical analysis

Latent class analysis, described above, was performed in MPlus Version 8.6 [68]. All other statistical procedures were performed in IBM SPSS Statistics Version 28 (Armonk, NY). Differences in years of education and literacy by race/ethnicity were examined using analysis of variance. Differences in race/ethnicity by diagnostic category were examined using Pearson's chi-squared test. Agreement of MCI diagnostic methods was examined using chi-squared tests and kappa statistics.

Longitudinal changes in MMSE and DSRS were each modeled using linear mixed-effects models, with subject-specific random intercepts and with diagnostic classification as a fixed factor. The interaction between time (years) and diagnosis was the effect of interest. Incident dementia was examined in relation to baseline diagnostic classification using Cox proportional hazards regression models. For all analyses, each of the three diagnostic methods was examined in separate models. For each method, one set of models compared MCI and NC (i.e., dichotomous classifications), while another set compared all subtypes of MCI and NC. Models were examined first in the entire sample and then stratified by race.

In secondary analyses, participants whose consensus and actuarial diagnoses disagreed (i.e., discrepant cases) were compared with participants classified as NC by both methods (i.e., concordant NC) and with participants classified as having MCI by both methods (i.e., concordant MCI). Differences among discrepant, concordant NC, and concordant MCI cases were examined with regard to race/ethnicity, LCA classifications, and longitudinal outcomes using chi-squared tests, mixed-effects models, and Cox proportional hazards models.

RESULTS

Sample characteristics

Sample demographics are presented in Table 1. Participants ranged in age from 55–93. Most participants identified as White (68%), and Black participants were well represented (29%), but other racial and ethnic groups were not. Thus, analyses of racial/ethnic differences

were restricted to White and Black participants. The sample was highly educated overall, but Black participants ($M = 14.3$, SD = 3.7) had significantly fewer years of education than White participants ($M = 16.5$, SD = 2.6), with a medium effect size ($p < 0.001$, η^2_p = 0.132). The difference in literacy/academic achievement was more pronounced. Mean WTAR standard scores were 100 (SD = 14) for Black participants and 116 (SD = 10) for White participants, with a large effect size ($p < 0.001$, $\eta^2_p = 0.273$), even when adjusting for years of education ($p < 0.001$, $\eta^2_p = 0.170$).

Consensus and actuarial diagnoses

Rates of MCI by each diagnostic method are presented in Table 2. Forty-four percent of participants were diagnosed with MCI by consensus at baseline. Single-domain aMCI and multi-domain aMCI were the most common subtypes. Subtypes of naMCI were infrequent and were therefore combined into one naMCI group for further analyses. Using actuarial criteria, 53% of participants were classified as having MCI. There was greater representation of naMCI among actuarial versus consensus diagnoses, but multi-domain naMCI was still infrequent and was thus combined with single-domain naMCI for further analyses.

Latent class analysis

Results supported a five-class model (Supplementary Table 1). Values of CAIC, BIC, and log likelihood differences indicated considerable improvement in model fit for the two-class model, steady though less pronounced improvement through the five-class model, and minimal improvement thereafter. The bootstrap likelihood ratio test remained statistically significant in all models examined, including those with unreasonably high numbers of classes. Thus, this test was not used in model selection. While the Vuong-Lo-Mendell-Rubin likelihood ratio test for the five-class model failed to reach statistical significance, this test is very susceptible to error. Classification accuracy (entropy) was acceptable for the five-class solution. Finally, the heterogeneity of cognition among non-demented elders is well-supported and clinically meaningful [69–73]. Thus, we interpreted the five-class solution based on a preponderance of evidence from fit statistics, theory, and empirical support.

LCA identified three groups with MCI, together comprising 41% of the sample, and two groups with NC (Table 2). Neuropsychological performance of the groups is presented in Fig. 1. Cognitive performance in the Low Normal (LN) group consistently fell at or just below normative means. Cognitive performance in the High Normal (HN) group consistently fell at or just above normative means. The three MCI groups included: 1) a Memory Only (M) group with isolated mild to moderate deficits in learning and delayed recall; 2) a Memory/Language (ML) group with profound forgetting, moderately impaired recognition, and mildly impaired language; and 3) a Memory/Executive (ME) group with mild to moderate impairments in learning and delayed recall with normal recognition, severe impairment on Trail Making Test B, and mild to moderate impairments on all other measures of attention, executive function, and language.

Race and diagnosis

Black participants were less likely than White participants to be diagnosed with MCI by consensus [27% versus 51%; χ^2 (1, N = 343) = 16.18, $p < 0.001$] and by LCA [30% versus 46%, χ^2 (1, N = 343) = 6.89, $p = 0.009$]. By contrast, Black participants were equally likely as White participants to be diagnosed with MCI by actuarial criteria [54% versus 52%, χ^2 (1, N = 343) = 0.12, $p = 0.73$]. Thus, while the choice of diagnostic method had a small impact on MCI prevalence among White participants (from 46% to 52%), Black participants were twice as likely to be classified as having MCI using the actuarial approach (54%) than using consensus or LCA (27–30%). Among the five LCA classes, Black participants were more likely to be classified into the LN group, comprising 54% of the LN group but only 17–37% of the other four groups ($p < 0.05$).

Longitudinal outcomes

Follow-up data was available for 230 participants (65% of the sample). Among them, the number of follow-up visits per participant ranged from one to 14 (median = 5), and follow-up time ranged from one to 15 years, with an average of 6.1 years (SD = 3.7). In comparison with participants without longitudinal data, these participants were three years younger at baseline ($p < 0.01$), more likely to be women (68% versus 55%, $p = 0.016$), and less likely to be White (64% versus 76%, $p = 0.022$). They also had better global cognition at baseline (MMSE = 28.7 versus 27.3, $p < 0.001$) and fewer functional difficulties (DSRS = 1.8 versus 4.2, $p < 0.001$), and were less likely to have been diagnosed with MCI (27% versus 75%, $p < 0.001$).

Global cognition

Using dichotomous baseline diagnostic classifications (i.e., MCI versus NC; Fig. 2a), all three diagnostic methods significantly predicted change in global cognition (MMSE; all $p < 0.001$). Effects were stronger for consensus diagnosis (slope difference = 0.55; that is, participants diagnosed with MCI by consensus declined by 0.55 more points per year than those diagnosed with NC by consensus) than for actuarial diagnosis (slope difference = 0.28) and LCA diagnosis (slope difference = 0.34). Across all diagnostic methods, participants with NC demonstrated stable MMSE scores, while participants with MCI demonstrated declines, but this decline was more pronounced in the consensus MCI group than in the actuarial and LCA MCI groups.

Among consensus subtypes (Fig. 2b), longitudinal MMSE change in all MCI subtypes differed from that of the NC group, and declines were steeper among participants with multi-domain aMCI or naMCI versus single-domain aMCI. Among actuarial subtypes (Fig. 2c), longitudinal MMSE change in the multi-domain aMCI group differed significantly from that of all other groups, and change in the single-domain aMCI group differed significantly from that of the NC group. Among LCA classes (Fig. 2d), longitudinal MMSE change in the ME and ML groups differed significantly from that of the other three groups and from each other, with declines steepest in the ME group.

Everyday functioning

Using dichotomous baseline diagnostic classifications (Fig. 3a), all three diagnostic methods significantly predicted change in functional difficulties (DSRS; $p < 0.001$), with effects stronger for consensus diagnosis (slope difference = 0.96) than for actuarial diagnosis (slope difference = 0.64) and LCA diagnosis (slope difference = 0.56). Across all diagnostic methods, participants with NC demonstrated subtle increases in functional difficulties over time, while participants with MCI demonstrated marked increases. These increases were more pronounced in the consensus MCI group than in the actuarial and LCA MCI groups.

Among consensus subtypes (Fig. 3b), longitudinal DSRS change in all MCI subtypes differed significantly from that of the NC group. The multi-domain aMCI and naMCI groups also had greater increases in functional difficulties than the single-domain aMCI group. Among actuarial subtypes (Fig. 3c), longitudinal DSRS change in all MCI subtypes differed significantly from that of the NC group. All MCI subtypes also differed significantly from each other, with the multi-domain aMCI group having the largest increase in functional difficulties, followed by the naMCI subtype. Among LCA classes (Fig. 3d), longitudinal DSRS change in the ME, ML, and LN groups differed significantly from each other and from the M and HN groups, which both demonstrated stable functioning. Increases in functional difficulties were largest in the ME group, followed by the ML group.

Incident dementia

Among the participants for whom follow-up diagnostic information was available, 16% (37/228) developed dementia, with an average time to dementia of 5.6 years (SD = 3.5). Using dichotomous baseline diagnostic classifications, all three diagnostic methods significantly predicted incident dementia (all $p < 0.001$, Fig. 4a). Hazard ratios (HR) were larger for consensus diagnoses (HR = 29.33), indicating a shorter time to dementia, than for actuarial (HR = 13.54) and LCA diagnoses (HR = 9.49). Rates of incident dementia were 50% for persons diagnosed with MCI by consensus versus 42% for persons with LCA-identified MCI and 35% for persons diagnosed with MCI by the actuarial approach. Rates of incident dementia among persons with NC were 4% for the consensus and actuarial methods versus 7% for LCA.

All subtypes of MCI as diagnosed by consensus differed significantly from the NC group in incident dementia, and the multi-domain aMCI subtype demonstrated greater dementia risk than the single-domain aMCI subtype (Fig. 4b). The same was true for actuarial subtypes (Fig. 4c); additionally, dementia risk was greater in the multi-domain aMCI subtype than in the naMCI subtype. Among LCA subtypes (Fig. 4d), the ME and ML LCA classes demonstrated significantly higher dementia risk than the M and LN groups, which in turn had higher dementia risk than the HN group.

Diagnostic agreement and discrepant cases

Diagnostic agreement among the three classification methods is presented in Table 3, with further breakdown by subtype in Supplementary Table 2. Consensus and LCA diagnoses of MCI demonstrated substantial agreement, with concordant diagnoses for 90% of cases. Among participants classified as having MCI by both methods, the ME group

was comprised almost entirely of cases with multi-domain aMCI, while the M and ML groups each contained roughly equal proportions of participants with single-domain aMCI and multi-domain aMCI. Actuarial diagnoses demonstrated lower, but still substantial, agreement with consensus diagnoses and with LCA. Thirteen percent of participants were classified as having MCI by actuarial criteria but NC by consensus (termed the "discrepant group" in subsequent analyses), while only 3% of participants were classified as having NC by actuarial criteria but MCI by consensus. Similarly, 14% of participants were classified as having MCI by actuarial criteria but NC by LCA, while only 1% of participants were classified as having NC by actuarial criteria but MCI by LCA.

Participants classified as having NC by consensus, but MCI by actuarial criteria (i.e., discrepant group), were further compared with participants classified as having MCI by both methods (i.e., concordant MCI) and participants classified as having NC by both methods (i.e., concordant NC). As few participants were classified as MCI by consensus but NC by actuarial criteria, they were excluded from this set of analyses. Non-amnestic MCI was over-represented among discrepant cases, comprising 67% of the discrepant group but only 7% of the concordant MCI group, and the majority of discrepant cases (70%) were classified as LN by LCA (Supplementary Table 2). Finally, Black individuals were over-represented in the discrepant group ($p < 0.001$), as they comprised 66% of this group versus 20% of the concordant MCI group and 31% of the concordant NC group.

The discrepant group demonstrated longitudinal cognitive (Fig. 5a), functional (Fig. 5b), and diagnostic outcomes (Fig. 5c) that were generally similar to that of the concordant NC group and consistently more favorable than that of the concordant MCI group. The discrepant group demonstrated stable global cognition, with a trajectory comparable to the concordant NC group ($p = 0.62$), and significantly better than the concordant MCI group ($p < 0.001$), which declined. The discrepant group exhibited a mild increase in functional difficulties, a trajectory slightly worse than that of the concordant NC group ($p = 0.041$). The difference between the discrepant group and the concordant MCI group was more pronounced ($p < 0.001$), as the concordant MCI group demonstrated marked increases in functional difficulties. Incident dementia risk in the discrepant group (Fig. 5c) was comparable to that of the concordant NC group ($p = 0.42$) and significantly lower than that of the concordant MCI group ($p < 0.001$). Rates of dementia were 35% in the concordant MCI group, 6% in the discrepant group, and 3% in the concordant NC group.

Race and longitudinal outcomes

The relationship between diagnostic method and longitudinal outcomes was further examined in analyses stratified by race. Overall, White participants demonstrated more cognitive and functional decline and a higher incidence of dementia (20%) than Black participants (9%). Results in White participants mirrored results in the overall sample. Specifically, all three diagnostic methods significantly predicted change in global cognition (all $p < 0.001$), with effects stronger for consensus diagnosis (slope difference = 0.59) than for actuarial diagnosis (slope difference = 0.31) and LCA diagnosis (slope difference = 0.36). The same pattern was observed for functional change (all $p < 0.001$; consensus: slope difference = 1.07, actuarial: slope difference = 0.64, LCA: slope difference = 0.67) and for

dementia risk (all $p < 0.001$; consensus: HR = 35.52, actuarial: HR = 13.99, LCA: HR = 8.45).

Among Black participants, the same pattern was observed for dementia risk (consensus: $p = 0.002$, HR = 32.63; actuarial: $p = 0.023$, HR = 12.14; LCA: $p = 0.006$, HR = 20.38). Results for cognitive and functional outcomes differed. Regarding change in global cognition, effect sizes were comparable across all three methods, but estimates varied in precision, resulting in differing significance levels (consensus: $p = 0.062$, slope difference = 0.19; actuarial: $p = 0.005$, slope difference = 0.19; LCA: $p = 0.042$, slope difference = 0.18). Only actuarial diagnosis ($p < 0.001$, slope difference = 0.65) significantly predicted change in everyday functioning, whereas consensus diagnosis ($p = 0.87$, slope difference = 0.03) and LCA diagnosis ($p = 0.50$, slope difference = 0.12) did not.

Race-adjusted scores

To further examine the impact of normative corrections, an additional set of analyses was completed in which cognitive test scores were adjusted for race in addition to age, sex, and education. Actuarial and LCA classifications were revised accordingly. The prevalence of MCI by the actuarial method was unchanged (53%) while the prevalence of LCA-identified MCI was slightly higher (44% versus 41%). LCA results supported a four-class model, with a single NC group. M, ME, and ML groups were comparable to the initial LCA solution. The agreement of the actuarial approach with consensus (0.72 versus 0.68) and with LCA (0.75 versus 0.70) improved slightly, while the agreement of consensus with LCA was unchanged (0.79 versus 0.80). Participants diagnosed with MCI by actuarial criteria but not by consensus (12% of the sample) remained more common than the converse (2%), and Black participants remained over-represented among discrepant cases. Longitudinal analyses followed the same pattern as analyses using age, sex, and education-adjusted norms.

DISCUSSION

The present study investigated the agreement and prognostic utility of three MCI diagnostic methods—consensus, actuarial, and statistical (i.e., LCA)—in a diverse sample of older adults. The consensus method employed widely used criteria for MCI [58, 59] and the consensus of clinicians using cognitive test scores and all available clinical information. Actuarial criteria were based on a previously published approach using only cognitive test scores [9, 10]. Finally, LCA was used to make statistical classifications also based on cognitive performance alone.

Consensus and LCA methods demonstrated substantial agreement, whereas the actuarial approach had weaker, but still substantial, agreement with the other two methods. In line with this agreement, all three methods significantly predicted longitudinal outcomes in the sample overall. However, in race-stratified analyses, findings differed by race and by longitudinal outcome. Consensus diagnoses best predicted incident dementia among both Black and White participants and best predicted cognitive and functional change in White participants. In Black participants, by comparison, actuarial diagnoses best predicted functional change, and effects sizes were comparable for the three methods in predicting

global cognitive change. These results provide validation for the consensus MCI diagnostic approach but also demonstrate the utility of actuarial diagnoses in certain situations.

The disagreement between actuarial classifications and the other two methods was primarily driven by cases diagnosed with MCI by actuarial criteria but classified as normal by consensus, termed the discrepant group. The converse (cases classified as normal cognition by actuarial criteria but MCI by consensus) was rare. Black participants were over-represented in the discrepant group, raising concern that this group might have been classified with MCI due to premorbid performance differences related to sociocultural factors [74], rather than pathological cognitive decline. Furthermore, in most analyses, outcomes in the discrepant group were comparable to those of concordant NC cases and more favorable than those of concordant MCI cases. Thus, these discrepant cases could be interpreted as reflecting false positive diagnoses identified with actuarial criteria.

However, the discrepant group demonstrated subtle functional declines that differed significantly from the concordant NC group. Additionally, as noted, actuarial criteria best predicted functional change among Black participants. Taken together, actuarial criteria appear less specific than consensus diagnoses, but more sensitive to the early signs of future decline, particularly among Black older adults. The pattern of findings suggests that consensus diagnoses may detect MCI later in its course, closer to the onset of dementia, while actuarial criteria may detect earlier changes that long precede the development of dementia.

Several factors may account for the disagreement between actuarial and consensus classification and the reduced specificity of the actuarial criteria. Prior studies have shown that abnormal test scores are quite common even among cognitively healthy older adults, and are more likely to occur in Black older adults and when using more liberal cutoffs and greater numbers of tests [8, 75]. The liberal 1 SD cutoff of the actuarial criteria may lend itself to false positive diagnosis in diverse older adults. Additionally, because both consensus and statistical methods consider weaknesses in the context of overall performance patterns, they are less likely to classify isolated low scores as indicative of impairment. Present findings suggest that these aspects of the actuarial approach may lead to reduced specificity overall, but greater sensitivity for certain outcomes. If greater specificity is desired, a more conservative 1.5 SD cutoff may be warranted [76].

Present findings differ somewhat from several prior studies that demonstrated better prognostic validity and fewer false positive diagnoses among actuarial versus conventional criteria [9, 12, 14, 26]. Differences may be related in part to the operationalization of actuarial criteria in comparison with prior studies. First, we used a larger number of tests and domains than in prior applications of the actuarial criteria, which increased the likelihood of MCI classifications [75]. Specifically, we used 14 tests in six cognitive domains, whereas prior studies used six to 12 tests in three to five domains [9, 13, 14, 77, 78]. Additionally, we included a learning domain in addition to a memory domain, whereas previous studies of the actuarial approach used only a memory domain [10–15]. While this choice limits comparisons to previous findings, this approach ensured that the actuarial and statistical methods utilized all available cognitive test scores in the interest of comprehensiveness and

utilized the same tests and domains as each other in the interest of consistency. This decision ensured that differences between the two methods were attributable to the approaches themselves, rather than to the tests and domains included.

Differing results from prior research may also be related to differences in the operationalization of conventional criteria. First, conventional diagnoses have often been made based on low performance on a single story memory test, which poorly reflects the construct of episodic memory and does not consider the full spectrum of cognitive performance [4]. By comparison, consensus diagnoses in this study used comprehensive neuropsychological data (2 tests per domain) and all other available clinical information, including demographics, medical and psychiatric history, neuroimaging, and care partner report. Incorporation of advanced biomarker data, such as hippocampal volumes, may further explain the enhanced performance of the conventional approach in comparison with prior studies. Moreover, this study used a consensus diagnosis approach at a tertiary care research center. Given the expertise of the study personnel and setting, and given evidence that decision making by heterogeneous groups is typically better than that of individuals [79, 80], expert consensus diagnoses in the present study may offer greater accuracy than diagnoses made by individual clinicians in standard practice settings. These findings suggest that conventional criteria can have good prognostic utility when combined with comprehensive clinical data and when using a consensus approach, rather than when used in isolation.

In diverse and marginalized groups, the selection of culturally appropriate tests and normative data is of critical importance in assessing cognitive decline, particularly when using actuarial or statistical diagnostic approaches that rely solely on cognitive test scores. In the present study, the primary analyses used normative adjustments based on age, sex, and years of education. Additional adjustment for race improved agreement of the actuarial method with consensus and LCA but did not affect the pattern of findings regarding prognostic validity. Notably, both normative methods have limitations, and alternative approaches may be warranted.

Black Americans, on average, tend to obtain lower scores on cognitive tests than White Americans due to a variety of factors stemming from longstanding systemic racism in the United States [81]. Some of these factors influence test performance, but not necessarily everyday cognitive functioning, such as education quality, acculturation, and bias [82–86]. In addition, Black Americans experience greater disadvantage in social and structural determinants of health (SSDoH; e.g., health care access and quality, economic stability, neighborhood infrastructure) [87–89] that increase risk of health problems known to influence cognitive functioning, such as cardiovascular disease, pulmonary disease, and chronic stress [90–93]. To account for racial disparities in cognitive test performance, normative adjustments have often used separate norms for Black and White patients in an effort to avoid pathologizing culturally based performance differences and reduce false positive diagnoses [17]. However, such norms can increase the risk of false negative diagnoses, thereby denying patients needed services. Moreover, these norms have been used to make harmful conclusions about race and biology, and do not consider the many sociocultural and SSDoH factors for which race is a proxy [90].

Therefore, normative adjustments that account for sociocultural factors may be an optimal alternative to reduce both false positive and false negative diagnostic error. Black Americans are disproportionately educated in schools with fewer resources than their White peers, a disparity that was particularly pronounced in the era when this study's participants were educated, an average of 60 years ago [87, 94]. Indeed, while Black participants in this study did have fewer years of education than White participants, the difference in literacy was far more pronounced and persisted when adjusting for years of education, indicating a marked disparity in quality of education. Thus, adjustments based on quality of education (i.e., literacy) may be more appropriate than adjustments based on years of education [83, 95] and may be less susceptible to misdiagnosis of older adults of color than race-based adjustments [17, 90]. Whether the use of education quality-based normative adjustments improves diagnostic concordance and prognostic validity in this sample is a topic of future study.

These findings highlight the importance of collecting information regarding SSDoH to utilize optimal normative adjustments, and to achieve a more complete understanding of the risk factors that impact cognitive aging and decline. Additionally, when appropriate, the use of repeat testing to compare within-participant cognitive stability and change may reduce the need for normative data, as each participant can serve as their own control. Finally, as cognitive measures have been largely developed in predominantly White samples and thus have an inherent degree of measurement bias, another approach to ensuring cross-cultural validity is the use of tests that are developed in diverse groups and demonstrate measurement invariance across groups.

Beyond examining diagnostic accuracy, another aim of the present study was to use a statistical method (LCA) to characterize cognitive heterogeneity among functionally independent older adults with cognitive decline, and to determine whether conventional MCI subtyping methods are sufficient in capturing this heterogeneity. Consistent with prior literature from both conventional [69, 96] and newer approaches [9, 20, 21, 23, 30, 33], LCA identified single-domain amnestic and multi-domain amnestic subtypes, and prognosis was poorest in the latter.

However, in contrast with conventional subtyping schemes, LCA identified two multi-domain amnestic groups that differed in their memory profile and in the number and type of other impaired domains. Specifically, the memory/language group demonstrated an encoding/retention-based memory deficit and prominent language impairment, while the memory/executive group demonstrated a retrieval-based memory deficit, prominent executive dysfunction, and deficits in several other domains. These two groups had similar dementia incidence, but the memory/executive group had significantly steeper cognitive and functional declines, which may be related to its greater breadth of impairment. This finding is consistent with a prior study [25] and suggests that conventional subtyping does not fully represent the heterogeneity of cognitive impairment in MCI. Our results highlight the importance of considering the memory profile and the nature of other affected domains, in addition to the presence of memory and non-memory impairment. Executive and memory impairment together appear to convey a particularly poor prognosis, and the reason for this

finding is a topic of further study. For example, this group may reflect a more advanced prodromal AD state, or the influence of multiple pathologies.

In contrast with conventional subtyping and with prior statistical examinations, LCA did not identify any non-amnestic MCI subtypes in this study. Using consensus and actuarial methods, non-amnestic MCI was present but infrequent in the current sample, and LCA's ability to detect such small groups may be limited. Characteristics of the study site likely contributed to the low rate of non-amnestic MCI. Namely, participants were enrolled at a memory center, resulting in an overrepresentation of AD pathology and amnestic MCI, while patients with non-AD pathologies such as Parkinson's disease and frontotemporal dementia are often seen at other specialty clinics at the same institution.

LCA captured cognitive heterogeneity not only in MCI but also in normal cognition. In line with prior studies [31, 32, 37], we identified two normal cognition groups: high normal and low normal. There are several possibilities regarding the nature of these groups: they may reflect a normal aging group and an increased risk group, a super-aging (reduced risk) group and a normal aging group [31, 32, 97], or two presentations of normal aging with differing baseline levels of performance [37]. The high and low normal groups differed demographically, with Black participants overrepresented in the low normal group. Moreover, using race-adjusted norms eliminated the differentiation between these two groups. The two groups had similar cognitive trajectories regarding global cognition, but functional and diagnostic outcomes were poorer in the low normal group, suggesting that it is indeed at increased risk relative to the high normal group. Thus, its relatively lower baseline cognitive scores may be attributable not just to premorbid differences in test performance, but also to early signs of pathological decline. Further analyses are warranted to examine whether SSDoH and biomarkers differ in the high and low normal groups and contribute to differences in longitudinal outcomes. These findings highlight the importance of considering the heterogeneity of normal cognition and MCI and the contribution of SSDoH, in order to increase accuracy of early detection of decline and identify risk and protective factors that predict clinical progression in a diverse population of patients.

While LCA was outperformed by the consensus and actuarial approaches, it nonetheless significantly predicted longitudinal outcomes and demonstrated substantial agreement with the other two methods. It is noteworthy that the actuarial and LCA approaches demonstrated good predictive validity without the cost, time, and effort of additional clinical information or the consensus of expert clinicians. The actuarial approach is well established in research [e.g., 9, 10, 78] and straightforward to apply clinically, though clinicians would benefit from guidelines regarding the number of tests and domains to include. Further work is needed to implement the LCA method for clinical use. A simple, static approach would utilize group means and standard deviations to determine how well a patient's data adheres to each LCA-identified group. A more dynamic approach would involve the creation of centralized repositories into which patients' scores could be entered and group membership determined by re-running the LCA. Both approaches would benefit from the current movement to harmonize testing protocols across sites [98].

A strength of the present study is the direct comparison of consensus, actuarial, and statistical (i.e., LCA) methods of diagnosing MCI. While these methods have been compared in a pairwise fashion, few studies have directly compared all three [9, 12]. Additional strengths include the use of LCA, which has a number of advantages over cluster analysis [41]; the inclusion of individuals with MCI and healthy cognition, enabling the identification of both false positive and false negative MCI diagnoses; and the long longitudinal follow-up period. The study used three longitudinal outcomes, allowing for a more nuanced understanding of prognostic utility, though it should be noted that consensus diagnosis was influenced by MMSE and DSRS scores and therefore not independent from these outcomes. Similarly, the consensus team was aware of previous consensus diagnoses, which may have biased incident dementia findings in favor of the consensus approach over the actuarial and statistical approaches. Missing longitudinal data reflects a weakness of the study. Participants with missing longitudinal data had poorer baseline cognition and functioning and were more likely to have an MCI diagnosis at baseline. We utilized mixed-effects modeling that accounted for missing information using maximum likelihood estimation. Nonetheless, attrition bias may have influenced longitudinal findings.

The inclusion of Black participants reflects another strength of the study given that prior examinations of actuarial and statistical MCI diagnostic methods have included predominantly White samples. That said, the sample was highly educated, and other ethnic groups, particularly Hispanic older adults, were under-represented. Additionally, Black participants had lower MCI rates and more favorable longitudinal trajectories than White participants, which is in contrast with well-documented findings regarding increased risk of MCI and dementia among Black versus White older adults [99–101]. Clinic-based samples are susceptible to selection bias, and results similar to the present findings have been observed in other such samples due to this bias (e.g., [102]). Additionally, selective survival can result in relatively more favorable longitudinal trajectories in Black older adults [103]. Selection and survival biases likely differentially influenced enrollment and retention of Black and White participants in the present study, and results should be interpreted with this caveat. There is a continued need for methods to enhance recruitment and retention of diverse older adults to better characterize their cognitive and functional trajectories and the validity of MCI diagnostic methods in diverse samples. Future work should explore the present study questions in community- and population-based samples that are more representative of Black older adults.

In conclusion, the present findings demonstrate the strengths and weaknesses of three approaches to differentiating MCI from normal cognitive aging in diverse older adults without dementia. Consensus diagnoses performed best overall, particularly in predicting incident dementia, but among Black older adults, actuarial diagnoses were more sensitive to early signs of decline. Dichotomous LCA classifications demonstrated the weakest associations with longitudinal outcomes, but LCA identified novel subtypes of MCI and normal cognition. Namely, LCA identified two subtypes of multi-domain amnestic MCI with differing outcomes, highlighting the importance of considering specific domains of impairment beyond memory. It also identified high normal and low normal cognition subtypes, with the latter demonstrating increased risk of longitudinal decline. The present study expands on prior studies by including a large proportion of Black older adults;

however, given issues of selection bias, future work should be conducted in samples more representative of older adults of color. A topic of future study is determining whether alternative normative adjustments or other strategies (e.g., repeat testing, cross-culturally valid tests) optimize the accuracy of MCI diagnoses. Such pursuits are necessary for the field to gain a complete understanding of the ways in which lifelong experiences, including racism, health risk factors and status, and sociocultural factors impact brain aging and risk for cognitive and functional decline.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

REFERENCES

[1]. Alzheimer's Association (2021) 2021 Alzheimer's disease facts and figures. Alzheimers Dement 17, 327–406. [PubMed: 33756057]

[2]. Petersen RC (2004) Mild cognitive impairment as a diagnostic entity. J Intern Med 256, 183–194. [PubMed: 15324362]

[3]. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR, Kawas CH, Klunk WE, Koroshetz WJ, Manly JJ, Mayeux R, Mohs RC, Morris JC, Rossor MN, Scheltens P, Carrillo MC, Thies B, Weintraub S, Phelps CH (2011) The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 7, 263–269. [PubMed: 21514250]

[4]. Bondi MW, Smith GE (2014) Mild cognitive impairment: A concept and diagnostic entity in need of input from neuropsychology. J Int Neuropsychol Soc 20, 129–134. [PubMed: 24490866]

[5]. Libon DJ, Drabick DAG, Giovannetti T, Price CC, Bondi MW, Eppig J, Devlin K, Nieves C, Lamar M, Delano-Wood L, Nation DA, Brennan L, Au R, Swenson R (2014) Neuropsychological syndromes associated with Alzheimer's/vascular dementia: A latent class analysis. J Alzheimers Dis 42, 999–1014. [PubMed: 25024329]

[6]. Jak AJ, Bangen KJ, Wierenga CE, Delano-Wood L, Corey-Bloom J, Bondi MW (2009) Contributions of neuropsychology and neuroimaging to understanding clinical subtypes of mild cognitive impairment. Int Rev Neurobiol 84, 81–103. [PubMed: 19501714]

[7]. Brooks BL, Iverson GL, Holdnack JA, Feldman HH (2008) Potential for misclassification of mild cognitive impairment: A study of memory scores on the Wechsler Memory Scale-III in healthy older adults. J Int Neuropsychol Soc 14, 463–478. [PubMed: 18419845]

[8]. Brooks BL, Iverson GL, White T (2007) Substantial risk of "accidental MCI" in healthy older adults: Base rates of low memory scores in neuropsychological assessment. J Int Neuropsychol Soc 13, 490–500. [PubMed: 17445298]

[9]. Bondi MW, Edmonds EC, Jak AJ, Clark LR, Delano-Wood L, McDonald CR, Nation DA, Libon DJ, Au R, Galasko D, Salmon DP (2014) Neuropsychological criteria for mild cognitive impairment improves diagnostic precision, biomarker associations, and progression rates. J Alzheimers Dis 42, 275–289. [PubMed: 24844687]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

[10]. Jak AJ, Bondi MW, Delano-Wood L, Wierenga C, Corey-Bloom J, Salmon DP, Delis DC (2009) Quantification of five neuropsychological approaches to defining mild cognitive impairment. Am J Geriatr Psychiatry 17, 368–375. [PubMed: 19390294]

[11]. Jak AJ, Urban S, McCauley A, Bangen KJ, Delano-Wood L, Corey-Bloom J, Bondi MW (2009) Profile of hippocampal volumes and stroke risk varies by neuropsychological definition of mild cognitive impairment. J Int Neuropsychol Soc 15, 890. [PubMed: 19570306]

[12]. Clark LR, Delano-Wood L, Libon DJ, McDonald CR, Nation DA, Bangen KJ, Jak AJ, Au R, Salmon DP, Bondi MW (2013) Are empirically-derived subtypes of mild cognitive impairment consistent with conventional subtypes? J Int Neuropsychol Soc 19, 635–645. [PubMed: 23552486]

[13]. Jak AJ, Preis SR, Beiser AS, Seshadri S, Wolf PA, Bondi MW, Au R (2016) Neuropsychological criteria for mild cognitive impairment and dementia risk in the Framingham Heart Study. J Int Neuropsychol Soc 22, 937–943. [PubMed: 27029348]

[14]. Wong CG, Thomas KR, Edmonds EC, Weigand AJ, Bangen KJ, Eppig JS, Jak AJ, Devine SA, Delano-Wood L, Libon DJ, Edland SD, Au R, Bondi MW (2018) Neuropsychological criteria for mild cognitive impairment in the Framingham Heart Study's old-old. Dement Geriatr Cogn Disord 46, 253–265. [PubMed: 30391953]

[15]. Edmonds EC, Delano-Wood L, Jak AJ, Galasko DR, Salmon DP, Bondi MW (2016) "Missed" mild cognitive impairment: High false-negative error rate based on conventional diagnostic criteria. J Alzheimers Dis 52, 685–691. [PubMed: 27031477]

[16]. Ganguli M, Snitz BE, Saxton JA, Chang C-CH, Lee C-W, Vander Bilt J, Hughes TF, Loewenstein DA, Unverzagt FW, Petersen RC (2011) Outcomes of mild cognitive impairment by definition: A population study. Arch Neurol 68, 761–767. [PubMed: 21670400]

[17]. Manly JJ (2005) Advantages and disadvantages of separate norms for African Americans. Clin Neuropsychol 19, 270–275. [PubMed: 16019708]

[18]. Blanken AE, Dutt S, Li Y, Nation DA (2019) Disentangling heterogeneity in Alzheimer's disease: Two empirically-derived subtypes. J Alzheimers Dis 70, 227–239. [PubMed: 31177226]

[19]. Machulda MM, Lundt ES, Albertson SM, Kremers WK, Mielke MM, Knopman DS, Bondi MW, Petersen RC (2019) Neuropsychological subtypes of incident mild cognitive impairment in the Mayo Clinic Study of Aging. Alzheimers Dement 15, 878–887. [PubMed: 31128864]

[20]. Thomas KR, Edmonds EC, Delano-Wood L, Bondi MW (2017) Longitudinal trajectories of informant-reported daily functioning in empirically defined subtypes of mild cognitive impairment. J Int Neuropsychol Soc 23, 521–527. [PubMed: 28487004]

[21]. Picón E, Juncos-Rabadán O, Lojo-Seoane C, Campos-Magdaleno M, Mallo SC, Nieto-Vietes A, Pereiro AX, Facal D (2019) Does empirically derived classification of individuals with subjective cognitive complaints predict dementia? Brain Sci 9, 314.

[22]. Delano-Wood L, Bondi MW, Sacco J, Abeles N, Jak AJ, Libon DJ, Bozoki A (2009) Heterogeneity in mild cognitive impairment: Differences in neuropsychological profile and associated white matter lesion pathology. J Int Neuropsychol Soc 15, 906–914. [PubMed: 19891820]

[23]. Libon DJ, Xie SX, Eppig J, Wicas G, Lamar M, Lippa C, Bettcher BM, Price CC, Giovannetti T, Swenson R, Wambach DM (2010) The heterogeneity of mild cognitive impairment: A neuropsychological analysis. J Int Neuropsychol Soc 16, 84–93. [PubMed: 19887015]

[24]. Damian M, Hausner L, Jekel K, Richter M, Froelich L, Almkvist O, Boada M, Bullock R, De Deyn PP, Frisoni GB, Hampel H, Jones RW, Kehoe P, Lenoir H, Minthon L, Olde Rikkert MGM, Rodriguez G, Scheltens P, Soininen H, Spiru L, Touchon J, Tsolaki M, Vellas B, Verhey FRJ, Winblad B, Wahlund LO, Wilcock G, Visser PJ (2013) Single-domain amnestic mild cognitive impairment identified by cluster analysis predicts alzheimer's disease in the European prospective DESCRIPA study. Dement Geriatr Cogn Disord 36, 1–19. [PubMed: 23651945]

[25]. Edmonds EC, Delano-Wood L, Clark LR, Jak AJ, Nation DA, McDonald CR, Libon DJ, Au R, Galasko D, Salmon DP, Bondi MW (2015) Susceptibility of the conventional criteria for mild cognitive impairment to false-positive diagnostic errors. Alzheimers Dement 11, 415–424. [PubMed: 24857234]

[26]. Edmonds EC, Eppig J, Bondi MW, Leyden KM, Goodwin B, Delano-Wood L, McDonald CR (2016) Heterogeneous cortical atrophy patterns in MCI not captured by conventional diagnostic criteria. Neurology 87, 2108–2116. [PubMed: 27760874]

[27]. Edmonds EC, Weigand AJ, Thomas KR, Eppig J, Delano-Wood L, Galasko DR, Salmon DP, Bondi MW (2018) Increasing inaccuracy of self-reported subjective cognitive complaints over 24 months in empirically derived subtypes of mild cognitive impairment. J Int Neuropsychol Soc 24, 842–853. [PubMed: 30278855]

[28]. Edmonds EC, McDonald CR, Marshall A, Thomas KR, Eppig J, Weigand AJ, Delano-Wood L, Galasko DR, Salmon DP, Bondi MW (2019) Early versus late MCI: Improved MCI staging using a neuropsychological approach. Alzheimers Dement 15, 699–708. [PubMed: 30737119]

[29]. Bangen KJ, Clark AL, Werhane M, Edmonds EC, Nation DA, Evangelista N, Libon DJ, Bondi MW (2016) Cortical amyloid burden differences across empirically-derived mild cognitive impairment subtypes and interaction with APOE ε4 genotype. J Alzheimers Dis 52, 849–861. [PubMed: 27031472]

[30]. Hanfelt JJ, Wuu J, Sollinger AB, Greenaway MC, Lah JJ, Levey AI, Goldstein FC (2011) An exploration of subgroups of mild cognitive impairment based on cognitive, neuropsychiatric and functional features: Analysis of data from the National Alzheimer's Coordinating Center. Am J Geriatr Psychiatry 19, 940–950. [PubMed: 22024618]

[31]. Köhler S, Hamel R, Sistermans N, Koene T, Pijnenburg YAL, Van Der Flier WM, Scheltens P, Visser PJ, Aalten P, Verhey FRJ, Ramakers I (2013) Progression To dementia in memory clinic patients without dementia: A latent profile analysis. Neurology 81, 1342–1349. [PubMed: 23997152]

[32]. Zammit AR, Muniz-Terrera G, Katz MJ, Hall CB, Ezzati A, Bennett DA, Lipton RB (2019) Subtypes based on neuropsychological performance predict incident dementia: Findings from the Rush Memory and Aging Project. J Alzheimers Dis 67, 125–135. [PubMed: 30507576]

[33]. McGuinness B, Barrett SL, McIlvenna J, Passmore AP, Shorter GW (2015) Predicting conversion to dementia in a memory clinic: A standard clinical approach compared with an empirically defined clustering method (latent profile analysis) for mild cognitive impairment subtyping. Alzheimers Dement (Amst) 1, 447–454. [PubMed: 27239523]

[34]. Peraita H, Chacon J, Diaz-Mardomingo C, Martinez-Arias R (2015) Profiles of mild cognitive impairment (MCI) in the elderly. Span J Psychol 18, E90. [PubMed: 26585437]

[35]. Blanken AE, Jang JY, Ho JK, Edmonds EC, Han SD, Bangen KJ, Nation DA (2020) Distilling heterogeneity of mild cognitive impairment in the National Alzheimer Coordinating Center database using latent profile analysis. JAMA Netw Open 3, e200413. [PubMed: 32142126]

[36]. Eppig JS, Edmonds EC, Campbell L, Sanderson-Cimino M, Delano-Wood L, Bondi MW (2017) Statistically derived subtypes and associations with cerebrospinal fluid and genetic biomarkers in mild cognitive impairment: A latent profile analysis. J Int Neuropsychol Soc 23, 564–576. [PubMed: 28578726]

[37]. Göthlin M, Eckerström M, Lindwall M, Rolstad S, Eckerström C, Jonsson M, Kettunen P, Svensson J, Wallin A (2020) Latent cognitive profiles differ between incipient Alzheimer's disease and dementia with subcortical vascular lesions in a memory clinic population. J Alzheimers Dis 73, 955–966. [PubMed: 31868665]

[38]. Hanfelt JJ, Peng L, Goldstein FC, Lah JJ (2018) Latent classes of mild cognitive impairment are associated with clinical outcomes and neuropathology: Analysis of data from the National Alzheimer's Coordinating Center. Neurobiol Dis 117, 62–71. [PubMed: 29859866]

[39]. Yu J, Lee TMC (2018) Profiles of cognitive impairments in an older age community sample: A latent class analysis. Neuropsychology 32, 102–109. [PubMed: 28661174]

[40]. Zammit AR, Hall CB, Lipton RB, Katz MJ, Muniz-Terrera G (2018) Identification of heterogeneous cognitive subgroups in community-dwelling older adults: A latent class analysis of the Einstein aging study. J Int Neuropsychol Soc 24, 511–523. [PubMed: 29317003]

[41]. Eshghi A, Haughton D, Legrand P, Skaletsky M, Woolford S (2011) Identifying groups: A comparison of methodologies. J Data Sci 9, 271–291.

[42]. Zammit AR, Hall CB, Bennett DA, Ezzati A, Katz MJ, Muniz-Terrera G, Lipton RB (2019) Neuropsychological latent classes at enrollment and postmortem neuropathology. Alzheimers Dement 15, 1195–1207. [PubMed: 31420203]

[43]. Zammit AR, Hall CB, Katz MJ, Muniz-Terrera G, Ezzati A, Bennett DA, Lipton RB (2018) Class-specific incidence of all-cause dementia and Alzheimer's disease: A latent class approach. J Alzheimers Dis 66, 347–357. [PubMed: 30282367]

[44]. Devlin KN, Giovannetti T (2017) Heterogeneity of neuropsychological impairment in HIV infection: Contributions from mild cognitive impairment. Neuropsychol Rev 27, 101–123. [PubMed: 28536861]

[45]. Edmonds EC, Ard MC, Edland SD, Galasko DR, Salmon DP, Bondi MW (2018) Unmasking the benefits of donepezil via psychometrically precise identification of mild cognitive impairment: A secondary analysis of the ADCS vitamin E and donepezil in MCI study. Alzheimers Dement (N Y) 4, 11–18. [PubMed: 29296659]

[46]. Díaz Mardomingo MDC, Peraita Adrados H (2008) Detección precoz del deterioro cognitivo ligero de la tercera edad. Psicothema 20, 438–444. [PubMed: 18674440]

[47]. Xie SX, Baek Y, Grossman M, Arnold SE, Karlawish J, Siderowf A, Hurtig H, Elman L, McCluskey L, Van Deerlin V, Lee VMY, Trojanowski JQ (2011) Building an integrated neurodegenerative disease database at an academic health center. Alzheimers Dement 7, e84–e93. [PubMed: 21784346]

[48]. Wechsler D(2001)Wechsler Test of Adult Reading: WTAR, The Psychological Corporation, San Antonio, TX.

[49]. Chin AL, Negash S, Xie S, Arnold SE, Hamilton R (2012) Quality, and not just quantity, of education accounts for differences in psychometric performance between African Americans and white non-Hispanics with Alzheimer's disease. J Int Neuropsychol Soc 18, 277–285. [PubMed: 22300593]

[50]. Ryan EL, Baird R, Mindt MR, Byrd D, Monzones J, Bank SM, Manhattan HIV Brain Bank (2005) Neuropsychological impairment in racial/ethnic minorities with HIV infection and low literacy levels: Effects of education and reading level in participant characterization. J Int Neuropsychol Soc 11, 889–898. [PubMed: 16519268]

[51]. Wechsler D (1987) Wechsler Memory Scale—Revised, The Psychological Corporation, San Antonio, TX.

[52]. Morris JC, Heyman A, Mohs RC, Hughes JP, van Belle G, Fillenbaum G, Mellits ED, Clark C (1989) The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. Neurology 39, 1159–1165. [PubMed: 2771064]

[53]. Goodglass H, Kaplan E (1983) The Assessment of Aphasia and Related Disorders, Lea and Febiger, Philadelphia, PA.

[54]. Kaplan E, Goodglass H, Weintraub S (1983) The Boston Naming Test, Lea and Febiger, Philadelphia, PA.

[55]. Reitan RM (1992) Trail Making Test, Reitan Neuropsychology Laboratory, Tucson, AZ.

[56]. Wechsler D(1981) WAIS-R Manual: Wechsler Adult Intelligence Scale-Revised, Psychological Corporation, New York, NY.

[57]. Saad L, Wolk DA, Mechanic-Hamilton D (2017) An update on normative data for neuropsychological performance on memory and language measures in a racially diverse older adult longitudinal cohort. In International Neuropsychological Society 45th Annual Meeting, New Orleans, LA.

[58]. Petersen RC, Morris JC (2005) Mild cognitive impairment as a clinical entity and treatment target. Arch Neurol 62, 1160–1163. [PubMed: 16009779]

[59]. Winblad B, Palmer K, Kivipelto M, Jelic V, Fratiglioni L, Wahlund L-O, Nordberg A, Backman L, Albert M, Almkvist O, Arai H, Basun H, Blennow K, de Leon M, DeCarli C, Erkinjuntti T, Giacobini E, Graff C, Hardy J, Jack C, Jorm A, Ritchie K, van Duijn C, Visser P, Petersen RC (2004) Mild cognitive impairment–beyond controversies, towards a consensus: Report of the International Working Group on Mild Cognitive Impairment. J Intern Med 256, 240–246. [PubMed: 15324367]

[60]. Akaike H (1974) A new look at the statistical model identification. IEEE Trans Automat Contr 19, 716–723.

[61]. Sclove SL (1987) Application of model-selection criteria to some problems in multivariate analysis. Psychometrika 52, 333–343.

[62]. Schwarz G (1978) Estimating the dimension of a model. Ann Stat 6, 461–464.

[63]. Bozdogan H (1987) Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. Psychometrika 52, 345–370.

[64]. Nylund KL, Asparouhov T, Muthén BO (2007) Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. Struct Equ Model 14, 535–569.

[65]. Folstein MF, Folstein SE, McHugh PR (1975) "Minimental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 12, 189–198. [PubMed: 1202204]

[66]. Monsell SE, Dodge HH, Zhou XH, Bu Y, Besser LM, Mock C, Hawes SE, Kukull WA, Weintraub S, Ferris S, Kramer J, Loewenstein D, Lu P, Giordani B, Goldstein F, Marson D, Morris J, Mungas D, Salmon D, Welsh-Bohmer K (2016) Results from the NACC Uniform Data Set Neuropsychological Battery Crosswalk Study. Alzheimer Dis Assoc Disord 30, 134–139. [PubMed: 26485498]

[67]. Clark CM, Ewbank DC (1996) Performance of the dementia severity rating scale: A caregiver questionnaire for rating severity in Alzheimer disease. Alzheimer Dis Assoc Disord 10, 31–39.

[68]. Muthén LK, Muthén BO (2012) Mplus User's Guide, Muthén & Muthén, Los Angeles, CA.

[69]. Knopman DS, Beiser A, Machulda MM, Fields J, Roberts RO, Pankratz VS, Aakre J, Cha RH, Rocca WA, Mielke MM, Boeve BF, Devine S, Ivnik RJ, Au R, Auerbach S, Wolf PA, Seshadri S, Petersen RC (2015) Spectrum of cognition short of dementia: Framingham Heart Study and Mayo Clinic Study of Aging. Neurology 85, 1712–21. [PubMed: 26453643]

[70]. Libon DJ, Xie SX, Eppig J, Wicas G, Lamar M, Lippa C, Bettcher BM, Price CC, Giovannetti T, Swenson R, Wambach DM (2010) The heterogeneity of mild cognitive impairment: A neuropsychological analysis. J Int Neuropsychol Soc 16, 84–93. [PubMed: 19887015]

[71]. Delano-Wood L, Bondi MW, Sacco J, Abeles N, Jak AJ, Libon DJ, Bozoki A (2009) Heterogeneity in mild cognitive impairment: Differences in neuropsychological profile and associated white matter lesion pathology. J Int Neuropsychol Soc 15, 906–914. [PubMed: 19891820]

[72]. Rasquin S, Lodder J, Visser P, Lousberg R, Verhey F (2005) Predictive accuracy of MCI subtypes for Alzheimer's disease and vascular dementia in subjects with mild cognitive impairment: A 2-year follow-up study. Dement Geriatr Cogn Disord 19, 113–119. [PubMed: 15591801]

[73]. Valdois S, Joanette Y, Poissant A, Ska B, Dehaut F (1990) Heterogeneity in the cognitive profile of normal elderly. J Clin Exp Neuropsychol 12, 587–596. [PubMed: 2211979]

[74]. Fujii D (2016) Conducting a Culturally Informed Neuropsychological Evaluation, American Psychological Association, Washington, DC.

[75]. Schretlen DJ, Testas SM, Winicki JM, Pearlson GD, Gordon B (2008) Frequency and bases of abnormal performance by healthy adults on neuropsychological testing. J Int Neuropsychol Soc 14, 436–445. [PubMed: 18419842]

[76]. Robertson K, Larson EB, Crane PK, Cholerton B, Craft S, McCormick WC, McCurry SM, Bowen JD, Baker LD, Trittschuh EH (2019) Using varying diagnostic criteria to examine mild cognitive impairment prevalence and prediet dementia incidence in a community-based sample. J Alzheimers Dis 68, 1439–1451. [PubMed: 30909213]

[77]. Thomas KR, Eppig JS, Weigand AJ, Edmonds EC, Wong CG, Jak AJ, Delano-Wood L, Galasko DR, Salmon DP, Edland SD, Bondi MW (2019) Artificially low mild cognitive impairment to normal reversion rate in the Alzheimer's Disease Neuroimaging Initiative. Alzheimers Dement 15, 561–569. [PubMed: 30610833]

[78]. Thomas KR, Cook SE, Bondi MW, Unverzagt FW, Gross AL, Willis SL, Marsiske M (2020) Application of neuropsychological criteria to classify mild cognitive impairment in the active study. Neuropsychology 34, 862–873. [PubMed: 33197199]

[79]. Kugler T, Kausel EE, Kocher MG (2012) Are groups more rational than individuals? A review of interactive decision making in groups. Wiley Interdiscip Rev Cogn Sci 3, 471–482. [PubMed: 26301530]

[80]. Apfelbaum EP, Phillips KW, Richeson JA (2014) Rethinking the baseline in diversity research: Should we be explaining the effects of homogeneity? Perspect Psychol Sci 9, 235–244. [PubMed: 26173261]

[81]. Nabors NA, Evans JD, Strickland TL (2000) Neuropsychological assessment and intervention with African Americans. In Handbook of Cross-Cultural Neuropsychology, Fletcher-Janzen E, Strickland TL, Reynolds CR, eds. Springer, Boston, MA, pp. 31–42.

[82]. Manly JJ, Miller SW, Heaton RK, Byrd D, Reilly J, Velasquez RJ, Saccuzzo DP, Grant I (1998) The effect of African-American acculturation on neuropsychological test performance in normal and HIV-positive individuals. The HIV Neurobehavioral Research Center (HNRC) Group. J Int Neuropsychol Soc 4, 291–302. [PubMed: 9623004]

[83]. Manly JJ, Jacobs DM, Touradji P, Small SA, Stern Y (2002) Reading level attenuates differences in neuropsychological test performance between African American and White elders. J Int Neuropsychol Soc 8, 341–348. [PubMed: 11939693]

[84]. McKay PF, Doverspike D, Bowen-Hilton D, McKay QD (2003) The effects of demographic variables and stereotype threat on black/white differences in cognitive ability test performance. J Bus Psychol 18, 1–14.

[85]. Loewenstein DA, Argüelles T, Argüelles S, Linn-Fuentes P (1994) Potential cultural bias in the neuropsychological assessment of the older adult. J Clin Exp Neuropsychol 16, 623–629. [PubMed: 7962363]

[86]. Helms JE (1997) The triple quandary of race, culture, and social class in standardized cognitive ability testing. In Contemporary Intellectual Assessment: Theories, Tests, and Issues, Flanagan D, Genshaft J, Harrison P, eds. Guilford Press, New York, pp. 517–532.

[87]. Glymour MM, Manly JJ (2008) Lifecourse social conditions and racial and ethnic patterns of cognitive aging. Neuropsychol Rev 18, 223–254. [PubMed: 18815889]

[88]. Thornton CM, Conway TL, Cain KL, Gavand KA, Saelens BE, Frank LD, Geremia CM, Glanz K, King AC, Sallis JF (2016) Disparities in pedestrian streetscape environments by income and race/ethnicity. SSM Popul Health 2, 206–216. [PubMed: 27314057]

[89]. Brondolo E, Gallo LC, Myers HF (2009) Race, racism and health: Disparities, mechanisms, and interventions. J Behav Med 32, 1–8. [PubMed: 19089605]

[90]. Manly JJ, Echemendia RJ (2007) Race-specific norms: Using the model of hypertension to understand issues of race, culture, and education in neuropsychology. Arch Clin Neuropsychol 22, 319–325. [PubMed: 17350797]

[91]. Orsi JM, Margellos-Anast H, Whitman S (2010) Black-white health disparities in the United States and Chicago: A 15-year progress analysis. Am J Public Health 100, 349–356. [PubMed: 20019299]

[92]. Adler NE, Rehkopf DH (2008) U.S. disparities in health: Descriptions, causes, and mechanisms. Annu Rev Public Health 29, 235–252. [PubMed: 18031225]

[93]. Zahodne LB, Kraal AZ, Zaheed A, Farris P, Sol K (2019) Longitudinal effects of race, ethnicity, and psychosocial disadvantage on systemic inflammation. SSM Popul Health 7, 100391. [PubMed: 31193191]

[94]. Borman GD, Dowling M (2010) Schools and inequality: A multilevel analysis of Coleman's equality of educational opportunity data. Teach Coll Rec 112, 1201–1246.

[95]. Sisco S, Gross AL, Shih RA, Sachs BC, Glymour MM, Bangen KJ, Benitez A, Skinner J, Schneider BC, Manly JJ (2015) The role of early-life educational quality and literacy in explaining racial disparities in cognition in late life. J Gerontol B Psychol Sci Soc Sci 70, 557–567. [PubMed: 24584038]

[96]. Summers MJ, Saunders NLJ (2012) Neuropsychological measures predict decline to Alzheimer's dementia from mild cognitive impairment. Neuropsychology 26, 498–508. [PubMed: 22612573]

[97]. Yu J, Collinson SL, Liew TM, Ng TP, Mahendran R, Kua EH, Feng L (2020) Super-cognition in aging: Cognitive profiles and associated lifestyle factors. Appl Neuropsychol 27, 497–503.

[98]. Langa KM, Ryan LH, McCammon RJ, Jones RN, Manly JJ, Levine DA, Sonnega A, Farron M, Weir DR (2020) The Health and Retirement Study Harmonized Cognitive Assessment Protocol Project: Study design and methods. Neuroepidemiology 54, 64–74. [PubMed: 31563909]

[99]. Chin AL, Negash S, Hamilton R (2011) Diversity and disparity in dementia: The impact of ethnoracial differences in Alzheimer disease. Alzheimer Dis Assoc Disord 25, 187–195. [PubMed: 21399486]

[100]. Kornblith E, Peltz C, Xia F, Barnes D, Byers A, Yaffe K (2020) Race but not sex differences in incidence of dementia diagnosis among older veterans: Epidemiology / Prevalence, incidence, and outcomes of MCI and dementia. Alzheimers Dement 16, e037606.

[101]. Wright CB, Derosa JT, Moon MP, Strobino K, Decarli C, Cheung YK, Assuras S, Levin B, Stern Y, Sun X, Rundek T, Elkind MSV, Sacco RL (2021) Race/ethnic disparities in mild cognitive impairment and dementia: The Northern Manhattan Study. J Alzheimers Dis 80, 1129–1138. [PubMed: 33646162]

[102]. Gleason CE, Norton D, Zuelsdorff M, Benton SF, Wyman MF, Nystrom N, Lambrou N, Salazar H, Koscik RL, Jonaitis E, Carter F, Harris B, Gee A, Chin N, Ketchum F, Johnson SC, Edwards DF, Carlsson CM, Kukull W, Asthana S (2019) Association between enrollment factors and incident cognitive impairment in Blacks and Whites: Data from the Alzheimer's Disease Center. Alzheimers Dement 15, 1533–1545. [PubMed: 31601516]

[103]. Zahodne LB, Manly JJ, Azar M, Brickman AM, Glymour MM (2016) Racial disparities in cognitive performance in mid- and late adulthood: Analyses of two cohort studies. J Am Geriatr Soc 64, 959–964. [PubMed: 27225354]
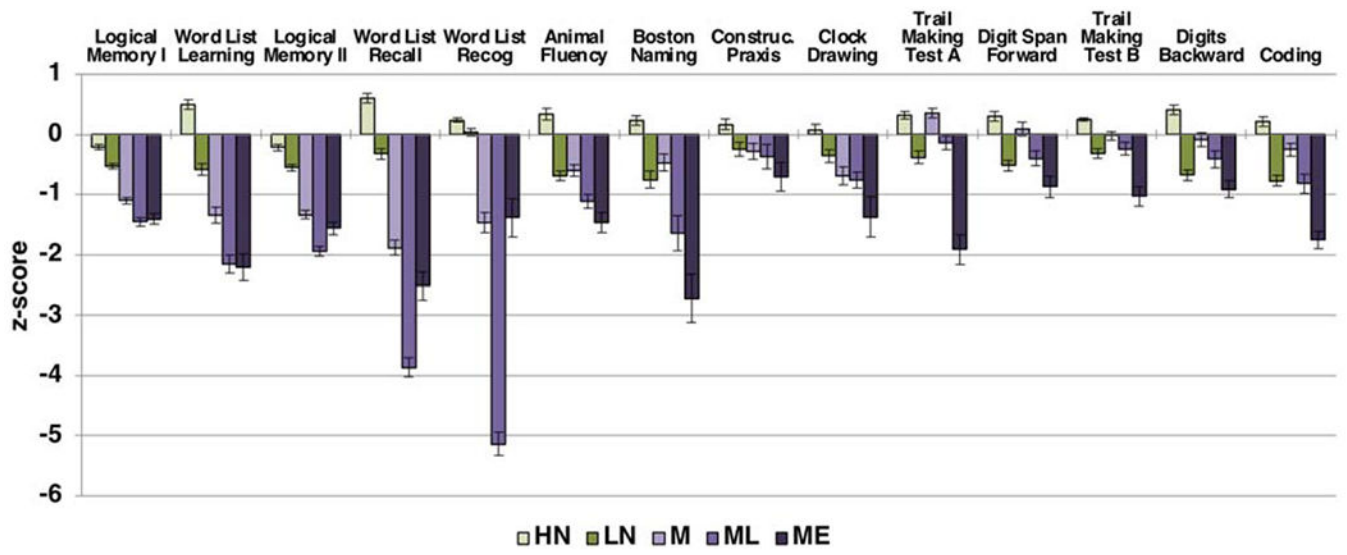
Fig. 1.
Neuropsychological performance of latent classes at baseline ($n = 354$). Error bars reflect standard error.
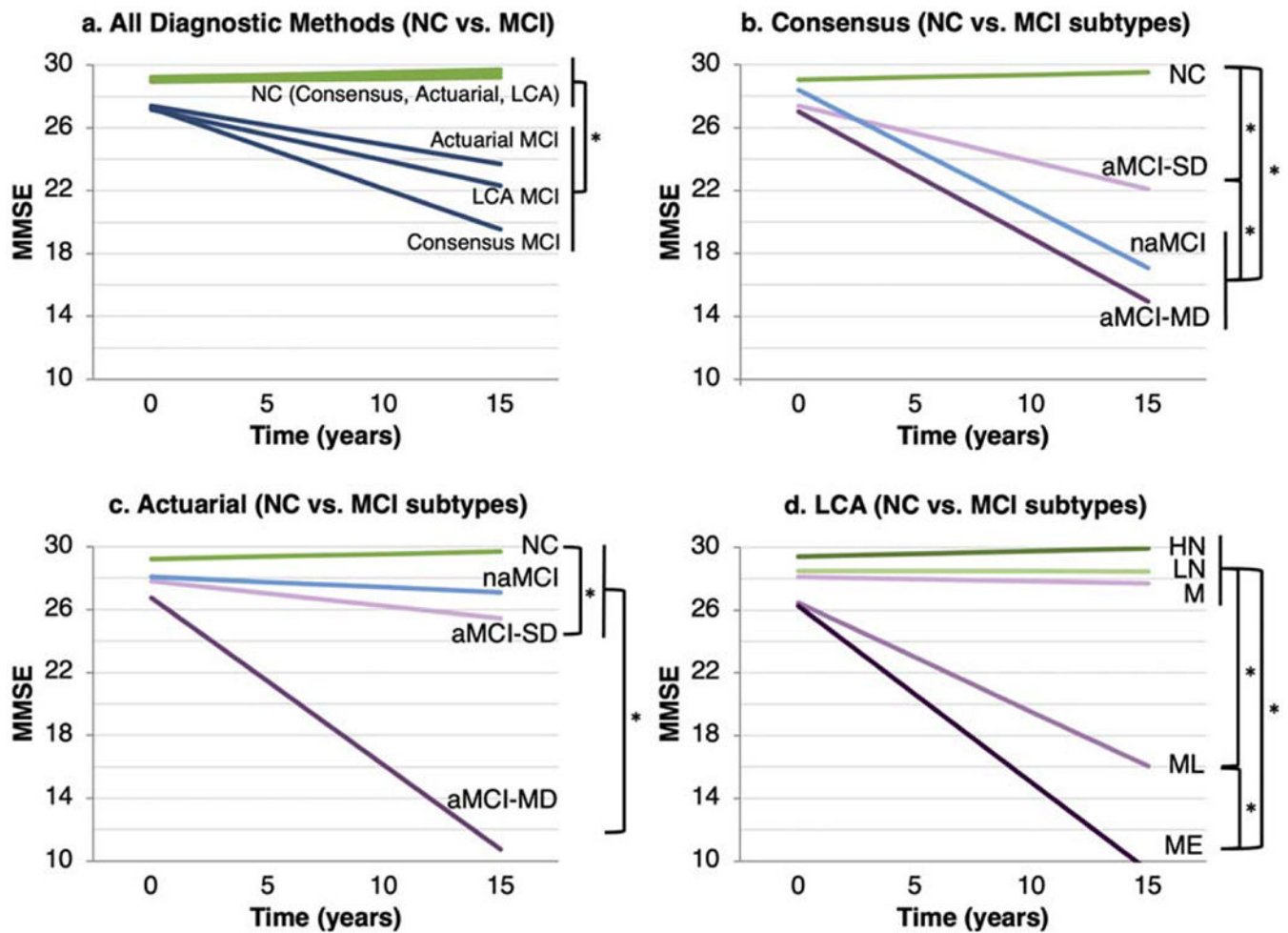
Fig. 2.
Longitudinal change in global cognition by a) baseline dichotomous diagnosis (all diagnostic methods), b) consensus subtype, c) actuarial subtype, and d) LCA subtype. Each bracket/asterisk denotes a significant group difference ($p < 0.05$) in the slope of change in MMSE. For example, in (d), the HN, LN, and M groups differ significantly from the ML and ME groups in the slope of change in MMSE, and the ML and ME groups differ from one another in the slope of change in MMSE. MMSE, Mini-Mental State Examination; LCA, latent class analysis; NC, normal cognition; aMCI, amnestic mild cognitive impairment; naMCI, non-amnestic MCI; SD, single domain; MD, multi-domain; HN, high normal; LN, low normal; M, memory only; ML, memory/language; ME, memory/executive.
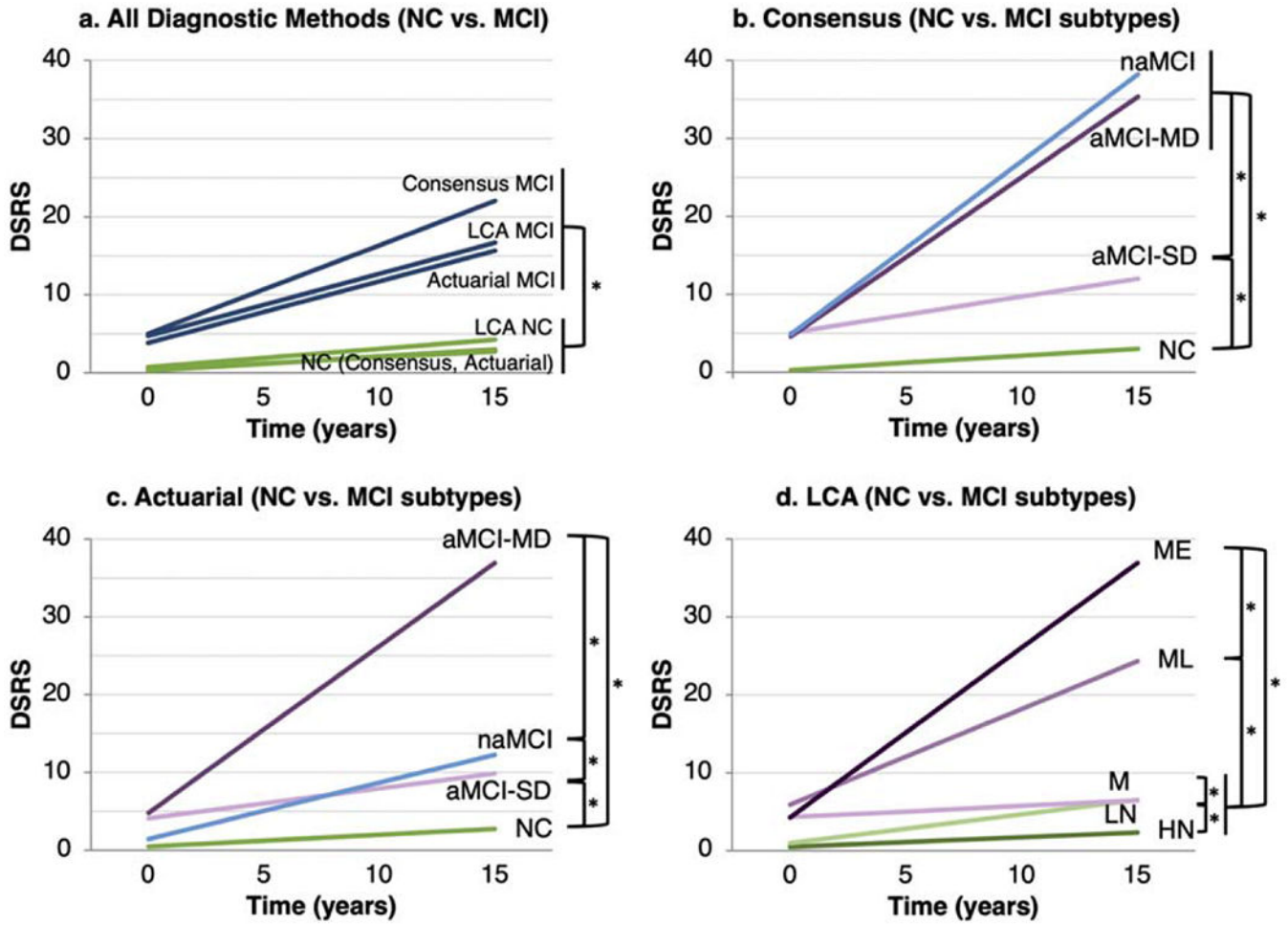
Fig. 3.
Longitudinal change in functional difficulties by a) baseline dichotomous diagnosis (all diagnostic methods), b) consensus subtype, c) actuarial subtype, and d) LCA subtype. Each bracket/asterisk denotes a significant group difference ($p < 0.05$) in the slope of change in DSRS. DSRS, Dementia Severity Rating Scale; LCA, latent class analysis; NC, normal cognition; aMCI, amnestic mild cognitive impairment; naMCI, non-amnestic MCI; SD, single domain; MD, multi-domain; HN, high normal; LN, low normal; M, memory only; ML, memory/language; ME, memory/executive.
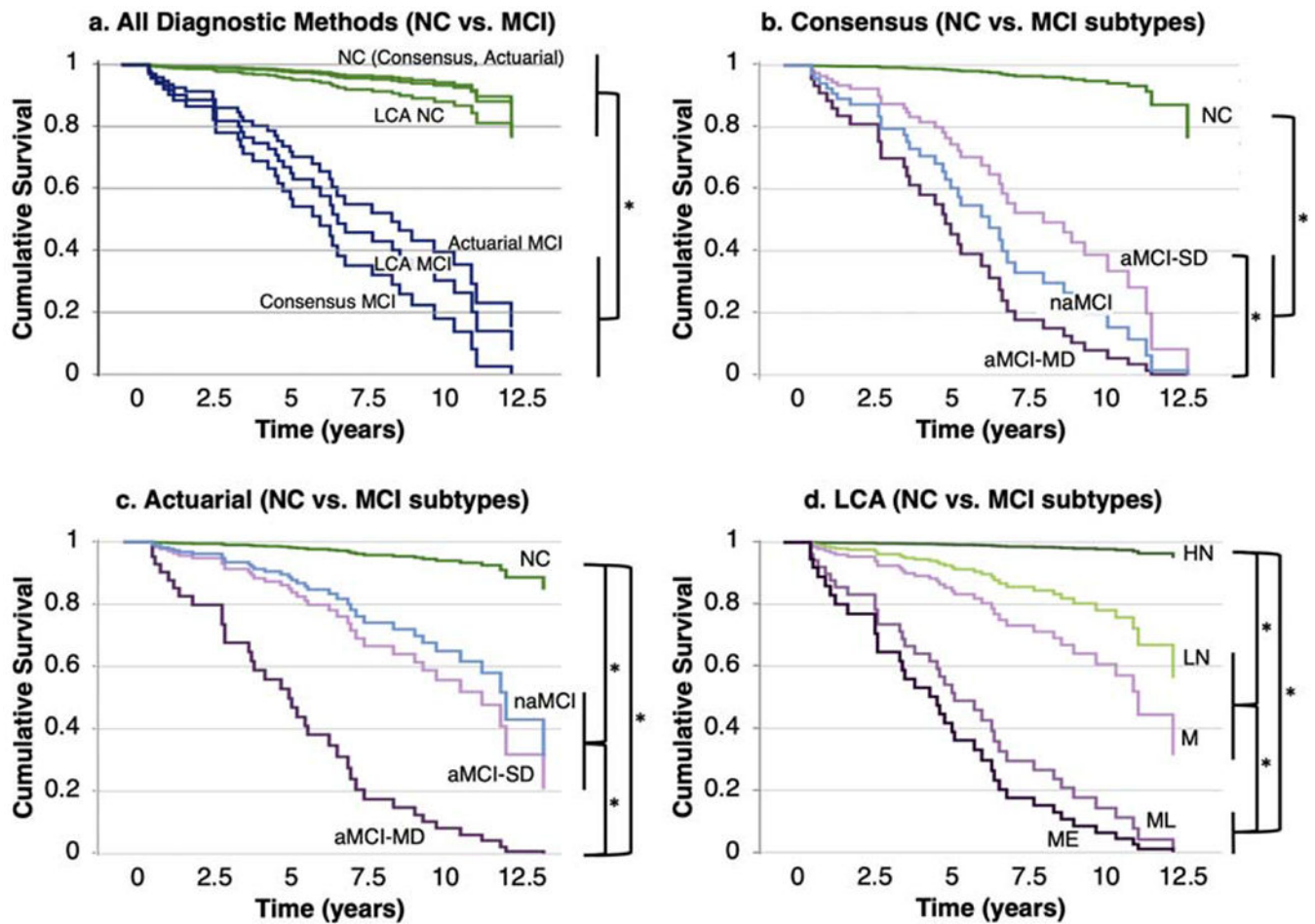
Fig. 4.
Incident dementia grouped by baseline diagnostic category: a) baseline dichotomous
diagnosis (all diagnostic methods), b) consensus subtype, c) actuarial subtype, and d) LCA
subtype. Each bracket/asterisk denotes a significant group difference ($p < 0.05$) in time
to dementia. NC, normal cognition; aMCI, amnestic mild cognitive impairment; naMCI,
non-amnestic MCI; SD, single domain; MD, multi-domain; LCA, latent class analysis;
HN, high normal; LN, low normal; M, memory only; ML, memory/language; ME, memory/
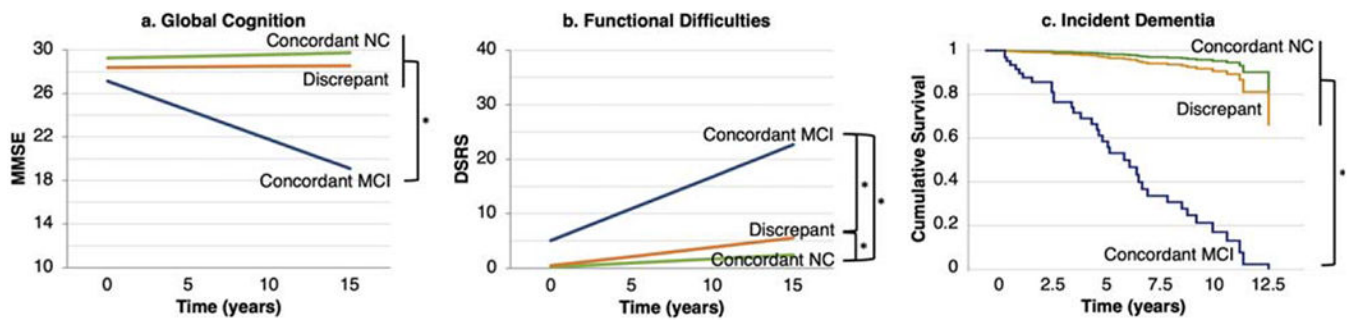executive.

Fig. 5.
Longitudinal outcomes among discrepant and concordant cases. a) Change in global cognition, b) change in functional difficulties, and c) incident dementia among cases diagnosed with mild cognitive impairment (MCI) by the actuarial approach but not by consensus (i.e., discrepant cases), cases diagnosed with MCI by both methods (i.e., concordant MCI), and cases diagnosed with normal cognition (NC) by both methods (i.e., concordant NC). Each bracket/asterisk denotes a significant group difference ($p < 0.05$) in the slope of change in MMSE or DSRS or in time to dementia. MMSE, Mini-Mental State Examination; DSRS, Dementia Severity Rating Scale.

Table 1

Sample Characteristics at Baseline ($N = 354$)

Age (y), M (SD)	70.9 (7.7)
Education (y), M (SD)	15.8 (2.9)
Gender, n (%)	
Women	224 (63%)
Men	130 (37%)
Race, n (%)	
White	241 (68%)
Black	102 (29%)
Asian-American	4 (1%)
Native American	1 (< 1%)
Multi-racial	6 (2%)
Ethnicity, n (%)	
Hispanic or Latino	7 (2%)
Not Hispanic or Latino	347 (98%)

Rates of MCI, Normal Cognition, and Their Subtypes by Diagnostic Method at Baseline

	Consensus	Actuarial		LCA
Normal Cognition	199 (56%)	165 (47%)	Normal Cognition (Total)	210 (59%)
			High Normal	109 (31%)
			Low Normal	101 (29%)
MCI (Total)	155 (44%)	189 (53%)	MCI (Total)	144 (41%)
aMCI, single domain	62 (18%)	79 (22%)	Memory Only	68 (19%)
aMCI, multi-domain	76 (21%)	69 (19%)	Memory/Language	47 (13%)
naMCI, single domain	16 (5%)	35 (10%)	Memory/Executive	29 (8%)
naMCI, multi-domain	1 (< 1%)	6 (2%)		

aMCI, amnestic mild cognitive impairment; naMCI, non-amnestic mild cognitive impairment; LCA, latent class analysis.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3

Agreement of MCI Diagnostic Methods at Baseline

	LCA NC	LCA MCI	% Agreement	κ (SE)	χ^2
Consensus NC	187	12	90%	0.80 (0.03)	226.10[*]
Consensus MCI	23	132			
	Consensus NC	**Consensus MCI**			
Actuarial NC	153	12	84%	0.68 (0.04)	167.39[*]
Actuarial MCI	46	143			
	LCA NC	**LCA MCI**			
Actuarial NC	161	4	85%	0.70 (0.04)	187.41[*]
Actuarial MCI	49	140			

LCA, latent class analysis.

[*]
$p < 0.001$.