Comparison of Conventional and Actuarial Neuropsychological Criteria for Mild Cognitive Impairment in a Clinical Setting

Stephanie Fountain-Zaragoza[1,2], **Sarah Ellen Braun**[3], **Michael David Horner**[2,1], **Andreana Benitez**[4]

[1]Department of Psychiatry and Behavioral Sciences, Medical University of South Carolina, Charleston, SC

[2]Mental Health Service, Ralph H. Johnson Department of Veterans Affairs Medical Center, Charleston, SC

[3]Department of Neurology, School of Medicine, Virginia Commonwealth University, Richmond, VA

[4]Department of Neurology, Medical University of South Carolina, Charleston, SC

Abstract

Introduction: Evidence-based practice in neuropsychology involves the use of validated tests, cutoff scores, and interpretive algorithms to identify clinically significant cognitive deficits. Recently, actuarial neuropsychological criteria (ANP) for identifying mild cognitive impairment were developed, demonstrating improved criterion validity and temporal stability compared to conventional criteria (CNP). However, benefits of the ANP criteria have not been investigated in non-research, clinical settings with varied etiologies, severities, and comorbidities. This study compared the utility of CNP and ANP criteria using data from a memory disorders clinic.

Method: Data from 500 non-demented older adults evaluated in a Veterans Affairs Medical Center memory disorders clinic were retrospectively analyzed. We applied CNP and ANP criteria to the Repeatable Battery for the Assessment of Neuropsychological Status, compared outcomes to consensus clinical diagnoses, and conducted cluster analyses of scores from each group.

Results: The majority (72%) of patients met both the CNP and ANP criteria and both approaches were susceptible to confounding factors such as invalid test data and mood disturbance. However, the CNP approach mis-labeled impairment in more patients with non-cognitive disorders and intact cognition. Comparatively, the ANP approach misdiagnosed patients with depression at a third of the rate and those with no diagnosis at nearly half the rate of CNP. Cluster analyses revealed groups with: 1) minimal impairment, 2) amnestic impairment, and 3) multi-domain impairment. The ANP approach yielded subgroups with more distinct neuropsychological profiles.

Conclusions: We replicated previous findings that the CNP approach is over-inclusive, particularly for those determined to have no cognitive disorder by a consensus team. The ANP approach yielded fewer false positives and better diagnostic specificity than the CNP. Despite clear benefits of the ANP vs. CNP, there was substantial overlap in their performance in this

Corresponding Author: Stephanie Fountain-Zaragoza, 96 Jonathan Lucas Street, Charleston, SC 29425, fountast@musc.edu.

heterogeneous sample. These findings highlight the critical role of clinical interpretation when wielding these empirically-derived tools.

Keywords

diagnostic criteria; cognition; memory disorders; neuropsychology; evidence-based practice; cluster analysis

Introduction

At the crux of evidence-based practice in psychology is providing patient care in accordance with the best available scientific evidence (IOM, 2001; APA Presidential Task Force on EBP, 2006). In neuropsychology, a key element of this is using empirical, objective measures and validated cutoff scores to identify cognitive impairment. With these tools, neuropsychologists employ their expertise to draw inferences about clinical diagnosis and etiology. Thus, efforts toward establishing and iteratively refining the empirical bases for these tests and the diagnostic criteria to which they are applied are crucial for guiding clinical practice (Sweet et al., 2017).

Mild cognitive impairment (MCI) is one diagnosis that has borne a variety of operational and psychometric definitions over time. Historically, the term MCI referred to an amnestic disorder that represented a prodromal phase of Alzheimer's disease (Petersen et al., 2001) but has since broadened to include non-amnestic and multidomain presentations (Petersen, 2004; Petersen & Morris, 2005). Conventional neuropsychological criteria (CNP) operationalize MCI as a single impaired score (> 1.5 SD below age-appropriate norms) on any test in one or more cognitive domains (Petersen & Morris, 2005). A major limitation of this approach is that reliance on a single impaired score increases false positive diagnostic errors, as many neurologically healthy adults obtain one or more impaired scores on testing (Heaton et al., 2004; Heaton et al., 1991; Palmer et al., 1998).

Recent work leverages more comprehensive neuropsychological data to improve the reliability and validity with which MCI is diagnosed, but the application of these new criteria to clinical settings has yet to be examined. Specifically, actuarial neuropsychological criteria (ANP; Jak et al., 2009) were developed that require at least two impaired scores (> 1 SD below age-appropriate norms) per cognitive domain in an effort to improve sensitivity and specificity. This approach has demonstrated improved diagnostic precision at the behavioral, neural, and biomarker level. In longitudinal studies, a greater number of those who met CNP criteria reverted from MCI to normal, while the ANP criteria resulted in greater diagnostic stability over time (Jak et al., 2009; Thomas et al., 2019; Wong et al., 2018). Cluster analyses have revealed that the two criteria identified groups with distinct neuropsychological profiles. In two prior studies, the ANP approach yielded dissociable amnestic, mixed impairment, dysexecutive, language, and visuospatial subtypes, while the CNP approach was limited to amnestic and mixed impairment and a sizeable group who perform within normal limits (Bondi et al., 2014; Clark et al., 2013). There is evidence that a considerable proportion of this "cluster-derived normal" group was mislabeled as MCI, as they had fewer APOE-ε4 carriers, did not differ from cognitively normal older adults in

cortical atrophy or amyloid burden, and had lower rates of progression to dementia (Bangen et al., 2016; Edmonds et al., 2015; Edmonds, Eppig, et al., 2016). Whereas the CNP approach demonstrates this susceptibility to false positives, the ANP approach demonstrates superior identification of individuals at risk for dementia (Bondi et al., 2014; Edmonds et al., 2015; Jak et al., 2016).

Despite these compelling findings, these effects have been observed exclusively in research studies (mostly using data from the Alzheimer's Disease Neuroimaging Initiative), and it is as yet unknown whether the benefits of the ANP criteria generalize to more varied clinical settings in which other test batteries are administered. Retrospective analysis of accumulated data from a clinical practice is one such method of ascertaining the benefits of the ANP criteria. However, inherent to this approach is the problem of circularity that precludes explicit tests of diagnostic accuracy, as the test scores to which the criteria would be applied contributed to the clinical diagnoses. Nonetheless, comparisons of the CNP and ANP criteria in clinical samples are needed to examine how frequently they are met using test batteries common to this setting and how they correspond to more comprehensive diagnoses informed by patient history, collateral informant reports, and professional input from other providers. Importantly, clinical datasets, such as repositories from memory disorders clinics, offer insights into certain realities that are not present in carefully selected research studies, such as variability in disease etiology and performance validity (Barker et al., 2010; Borghesani et al., 2010; Howe, et al., 2007; Martin & Schroeder, 2020; Paulson et al., 2015). As these issues are relevant to many clinical contexts in which the CNP or ANP criteria might be used, it is important to evaluate how these two approaches perform when faced with those challenges.

The primary aim of this study was to compare the rates at which the CNP and ANP criteria for MCI were met in a sample of non-demented patients seen in a VA Medical Center's memory disorders clinic, using a test battery (the Repeatable Battery for the Assessment of Neuropsychological Status [RBANS]; Randolph, 1998) which is commonly used in clinical settings (Rabin et al., 2016), yet has not been subject to investigation in the ANP literature. We also sought to evaluate the correspondence of CNP and ANP groups with clinical diagnoses conferred by an interdisciplinary team of neurologists, neuropsychologists, and geriatric psychiatrists. In accordance with previous studies, we hypothesized that the CNP criteria would be overly inclusive—labeling a larger number of patients as impaired and including more patients who were given no diagnosis by the clinical team—compared to the ANP approach. One unique aspect of this study is that unlike all prior studies of ANP, this study includes performance validity testing, given that patients seen in VA settings have been shown to provide invalid data at higher rates than patients in other clinical or research settings (Martin & Schroeder, 2020). We hypothesized that patients who failed validity testing would be captured by both criteria due to indiscriminate production of low scores. As follow up, we characterized the cognitive profiles of those that met each criterion, using cluster analysis, to explore how the features unique to this VA clinical sample (i.e., varied clinical etiologies, range of severities, comorbidities, and high rate of validity test failure) were reflected in the resulting groups.

Materials and Methods

Participants and Procedures

This study was a retrospective analysis of data collected from patients seen at an outpatient memory disorders clinic at a VA Medical Center in the southeastern United States. All procedures were in compliance with the Institutional Review Board (study #Pro00066285) and approved by the Office of Research and Development at the Ralph H. Johnson VA Medical Center. Informed consent was waived because the study consists solely of retrospective review of data collected in the course of routine clinical care.

Referrals to this clinic came from Primary Care, Neurology, Mental Health, and other VA clinics and typically referenced concerns related to memory problems and dementia; thus, the sample was primarily geriatric (M_{age} = 69.3 years, SD = 8.9, range: 43 – 90 years). Memory disorders clinic visits provided interdisciplinary evaluations involving neuropsychological evaluation (i.e., clinical interview and standardized testing), neurological examination, and psychiatric evaluation. Consensus diagnoses were then assigned by the interdisciplinary team, consisting of a neuropsychologist, behavioral neurologist, and geriatric psychiatrist, upon consideration of all available information. This included patient history, self- and/or informant-reported symptoms, clinical presentation, neurological examination, psychodiagnostic interviewing, lab results, neuroimaging, and neuropsychological test performance on a screening battery that included the RBANS. Consensus diagnoses were given after the totality of this information was discussed by the team in relation to the patient's estimated premorbid functioning. Although patients could be assigned multiple diagnoses (e.g., a cognitive disorder and psychiatric disorder), in this paper we report the primary diagnoses considered, by consensus, to be the leading etiology of the patient's condition.

Criteria from the DSM-IV-TR (American Psychiatric Association, 2000) and later DSM-5 (American Psychiatric Association, 2013) were used to diagnose dementia/major neurocognitive disorder and depressive disorders. Depressive disorder was assigned as a primary diagnosis only when there was no cognitive disorder diagnosed by the consensus team and no existing diagnoses/current treatment for other psychiatric conditions (e.g., PTSD; these patients were assigned to the "other" category). In the spirit of the original MCI concept, the practice of the clinic is to assign an MCI diagnosis only for amnestic presentations for which the suspected etiology is Alzheimer's disease. Amnestic MCI (aMCI) diagnoses were based on impaired RBANS index and subtest scores for memory domains (using Petersen criteria), consideration of other test scores, and information about functional abilities and change from baseline collected in the clinical interviews. Patients who exhibited cognitive impairment but who did not meet criteria for aMCI or dementia were diagnosed with cognitive disorder not otherwise specified (CD-NOS). Of note, no explicit cut-off scores on neuropsychological tests were used for making any of the consensus diagnoses described above. Data were collected, and diagnoses were assigned, before the study hypotheses were formulated.

Patient exclusion/inclusion is detailed in the flow chart in Figure 1. Complete data were available for 824 patients who were consecutively seen for an initial evaluation between

2002 and 2019. Given that we aimed to test criteria for identifying *mild* cognitive impairment, we excluded data from individuals who received a consensus diagnosis of dementia (i.e., Alzheimer's disease, vascular dementia, Lewy body dementia, alcohol-induced dementia, frontotemporal dementia, mixed dementia, or dementia NOS). Excluding patients with dementia was consistent with the development of the CNP (Petersen & Morris, 2005) and ANP (Jak et al., 2009) as well as the approach taken by previous studies comparing the two (Bondi et al., 2014; Edmonds et al., 2015; Edmonds, Eppig, et al., 2016; Jak et al., 2016; Thomas et al., 2019; Wong et al., 2018). We included participants who were assigned CD-NOS, aMCI, depressive disorder (reflecting either major depressive disorder or depression NOS), no diagnosis, and failed validity. The "no diagnosis" group was comprised of patients for whom no clinical diagnosis (psychiatric or cognitive) was suspected. Of note, patients who failed validity testing were often assigned a clinical diagnosis by the interdisciplinary team, but were recoded into a failed validity group for the purposes of this study. Individuals for whom diagnosis was coded as "other" or "diagnosis deferred", but who did not fail validity testing, were excluded given that these designations were nonspecific and did not provide clinically interpretable information for the purposes of this study. In consideration of the cluster analysis performed here which is sensitive to outliers (Almeida et al., 2007; Jin & Han, 2010; Yim & Ramdeen, 2015), data were excluded for participants who had any RBANS subtest score that was an extreme outlier, defined as beyond 3 times the interquartile range. This resulted in a final sample of 500 older adults (M_{age} = 69.3, SD = 8.9 years) who were mostly men (97%) and had completed one year of college on average (M_{edu} = 13.1, SD = 2.9 years). The majority of the sample identified as White (73.8%), followed by African American (24.6%), Hispanic (1.0%), and Other/Not specified (0.6%). Descriptive statistics for RBANS subtests per diagnostic group are presented in Table S1.

Measures

Repeatable Battery of the Assessment of Neuropsychological Status (RBANS; Randolph, 1998).—The RBANS is a brief screening battery comprised of 12 subtests measuring 5 cognitive domains: Immediate Memory (List Learning, Story Learning), Delayed Memory (List Recall, List Recognition, Story Recall, Figure Recall), Attention (Digit Span, Coding), Language (Naming, Verbal Fluency), and Visuospatial/Constructional (Line Orientation, Figure Copy). Normative data from the manual was used to convert raw scores to age-adjusted *z*-scores and to then calculate the 5 corresponding Index scores for each domain and a Total score. For the purpose of this study, age-adjusted *z*-scores for each of the 12 individual subtests were used, as they provide at least two subtest scores per cognitive domain for application of the CNP and ANP criteria.

Test of Memory Malingering (TOMM; Tombaugh, 1997).—The TOMM is the most widely used stand-alone performance validity measure, with roughly 80% of surveyed neuropsychologists reporting frequent administration (Martin et al., 2015). The test includes three trials (Trial 1, Trial 2, and a delayed Retention Trial), each a 50-item, forced-choice picture recognition task. Scores on the Retention trial were not included in validity determinations for this study as this optional portion of the test was infrequently and inconsistently administered. Patients were either administered both Trial 1 and 2 per

standard administration ($n = 404$, 80.8%), or Trial 1 only ($n = 96$, 19.2%), based on accumulating evidence that scores on Trial 1 are sufficient for determining performance validity (Denning, 2012; Horner et al., 2006; Martin et al., 2019). For the purpose of the current study, patients were classified as having failed validity testing based on the following: the standard cutoff score of <45 on Trial 2 (Tombaugh, 1997) was used for those administered both trials; a cutoff of <42 proposed by a recent meta-analysis (Martin et al., 2019) was used for those given Trial 1 only. Thus, the number of patients in the failed validity group represents these categorical determinations.

Statistical Analyses

The full sample of 500 patients was classified regarding cognitive impairment using two approaches that differed in the severity of impairment and the number of impaired scores required per cognitive domain. The CNP criteria require at least one score > 1.5 *SD* below the age-referenced normative mean, whereas the ANP require two scores within a single domain > 1 *SD* below the age-referenced mean. These criteria were applied to the age-adjusted *z*-scores from each of the 12 RBANS subtests, stratifying patients into those meeting neither criterion (CNP-/ANP-), CNP only (CNP+/ANP-), ANP only (CNP-/ANP+), or both (CNP+/ANP+). We tested for differences across groups in age and education using one-way ANOVA with Tukey's *post hoc* tests, and in sex and race/ethnicity using chi-square with Fisher's exact test. Differences in consensus diagnoses across groups (CD-NOS, aMCI, depressive disorder, no diagnosis, failed validity) were determined using chi-square with the Monte Carlo method with 10,000 samples, given that this analysis was too large for a Fisher's exact test. Given the disproportionate cell sizes across these groups, we ensured that all variables met statistical assumptions or analyses were adapted accordingly. Agreement between the ANP and CNP criteria was assessed using Cohen's kappa (κ), which ranges from no agreement (0) to perfect agreement (1). Correspondence with consensus diagnoses was evaluated by assessing the composition of each group (i.e., the number and proportion of each diagnosis per group) and differences between CNP and ANP criteria per diagnosis group were tested using McNemar's test for paired dichotomous data. Results are reported as McNemar's exact binomial test (*p*-values only) when cell values are small or McNemar's test with continuity correction (chi-square and *p*-values) when cell values are sufficiently large.

Neuropsychological data from the groups meeting CNP and ANP were separately submitted to two cluster analyses to investigate the presence of clinical subtypes within those that met each criterion (see Supplementary Material for additional details regarding the statistical analysis and model evaluation methods). We excluded those who failed validity testing to allow us to characterize the neuropsychological profiles of subgroups without the confounding impact of invalid data. Cluster analysis requires all entered variables to be in the same unit of measurement. As such, age-corrected scores for the 12 RBANS subtests were z-standardized within-sample and entered into two sequential cluster analyses. For each, an agglomerative hierarchical cluster analysis using Ward's linkage was conducted as a data-driven method for identifying the cluster structure. The appropriate number of clusters was chosen by visual inspection of the dendrogram plot. Next, k-means cluster analyses were conducted for each potential solution, specifying *a priori* the number of

clusters and their centers determined in the previous step. Discriminant function analyses (DFA) were then used to quantify how well the 13 cognitive variables discriminated the cluster groups using a cross-validation technique. Chi-square tests determined significance, and the percentage of correct classifications was compared across cluster solutions.

We statistically compared the profile of RBANS scores within each cluster while accounting for the overlap between groups (i.e., those who were CNP+/ANP+). To do so, for each of the three clusters we examined differences in mean Z-score for each subtest between those meeting CNP only and those meeting ANP only, excluding the overlapping individuals. We used ANOVA's with Tukey *post hoc* tests for equal variances and Welch's ANOVA's with Games-Howell *post hoc* tests for unequal variances. To correct for multiple comparisons, we report results for main effects of group surviving Bonferroni correction.

Results

Comparing CNP and ANP Criteria

More patients met criteria using the CNP (444 patients, CNP+) than ANP approach (368 patients, ANP+), but overall there was moderate agreement between the two ($\kappa = .42$, $p < .001$). Of the 500 total patients, 9.6% met neither criterion (CNP-/ANP-), 16.8% met CNP but not ANP criteria (CNP+/ANP-), only 1.6% met ANP but not CNP criteria (CNP-/ANP+), and 72.0% met criteria for both (CNP+/ANP+). Descriptive statistics for demographic variables, consensus diagnoses, and test performance per group are presented in Table 1. Comparison of group characteristics revealed a significant main effect of group on age ($F(3,499) = 2.94$, $p = .033$); however, no *post hoc* between-group comparisons were significant. There was a significant between-groups difference in education ($F(3,498) = 8.74$, $p < .001$) such that the CNP+/ANP+ group had fewer years of education on average than both the CNP-/ANP- group ($p < .001$) and the CNP+/ANP- group ($p = .006$). There was no significant association between sex and group (Fisher's exact test = 6.02, p = .082). There was a significant association between race and group (Fisher's exact test = 17.51, $p = .032$) such that compared to expected counts, the CNP-/ANP- had fewer African American patients, the CNP+/ANP- group had more White and fewer African American patients, and the CNP+/ANP+ group had fewer White and more African American patients.

There was a significant association between consensus diagnosis and group ($X^2(12) = 183.66$, Monte Carlo $p < .001$). The CNP-/ANP- group was significantly weighted toward those with non-cognitive diagnoses, comprised solely of patients assigned no diagnosis (35 patients, 72.9%) and depression (13 patients, 27.1%). This suggests good concordance between those who were not labeled as impaired by either approach and consensus diagnoses. However, the CNP+/ANP- also had significantly more patients with non-cognitive diagnoses than expected counts, with the majority of this group having no diagnosis (44 patients, 52.4%) and depression (20 patients, 23.8%), suggesting that relying on CNP criteria alone may yield erroneous labeling of impairment. The much smaller group of patients who were CNP-/ANP+ had a roughly equivalent distribution of those with no diagnosis (3 patients, 37.5%), depression (3 patients, 37.5%), and aMCI (2 patients, 25.0%), but no patients with CD-NOS. The largest group, those meeting both criteria (CNP+/ANP+), demonstrated the best concordance with consensus diagnoses. Specifically, compared to

expected counts, this group had disproportionately more patients with CD-NOS (capturing 90.8% of total CD-NOS cases), aMCI (capturing 92.9% of cases), and failed validity (100% of cases) and fewer with depression (capturing only 54.4% of cases) and no diagnosis (capturing only 32.8% of cases).

We next assessed the performance of the two approaches by examining the diagnostic compositions of patients that met each criterion (Table 2). The two approaches significantly differed in the number of cognitive and non-cognitive diagnoses captured ($X^2(1) = 46.41$, $p < .001$), with CNP labeling impairment in over three quarters (77.8%) of those with non-cognitive diagnoses compared to ANP labeling impairment in roughly half (53.9%). Specifically, the CNP approach identified impairment at greater rates than the ANP among patients who, per consensus diagnosis, were deemed to have a depressive diagnosis but no cognitive disorder (79.8% vs. 58.2%; $p < .001$) and no diagnosis (68.9% vs. 35.3%; $X^2(1) = 35.3$, $p < .001$). Of note, 100% of patients who failed validity testing met both CNP and ANP criteria, highlighting each criterion's lack of specificity in the context of invalid test data. Differences were much less apparent for cognitive diagnoses. There was no significant difference in the proportion of those with aMCI captured by either approach (100% vs. 95.3%; $p = .687$). Although the CNP approach was slightly more inclusive of CD-NOS than the ANP approach, both labeled impairment in the vast majority of cases (100% vs. 90.8%; $p < .001$).

Cluster Analyses

Two cluster analyses of RBANS scores were conducted for patients who passed validity testing, one using data from patients who met CNP criteria ($n = 402$) and the other using data from those who met ANP criteria ($n = 326$). For both, 3-cluster solutions were determined to be the best fit (see Supplementary Material for details), correctly classifying 94.3% of cases for CNP and 92.0% of cases for ANP, compared to 91.5% and 91.7%, respectively, for the 4-cluster solutions. We evaluated the diagnostic compositions (Figure 2) and neuropsychological characteristics (Figure 3 and Table S2) of these clusters. We used the uniform labeling procedures set forth by the American Academy of Clinical Neuropsychology (Guilmette et al., 2020) to characterize test performance.

Overall, the cluster solutions were similar across the ANP and CNP approaches. The first clusters appeared to capture those who were minimally impaired, with all scores falling in the average to low average range. The second clusters captured those with primarily amnestic impairment, and comprised of a mix of patients with aMCI and CD-NOS. This group showed a clear amnestic pattern of scores in the below average to exceptionally low range on most verbal memory tests (i.e., List Learning, Story Learning, List Recall, List Recognition, Story Recall) as well as below average scores on Coding. The third clusters appeared to include those who had deficits in multiple domains, which contained a majority CD-NOS patients. Neuropsychological performance in this group was characterized by below average to exceptionally low scores on all verbal memory subtests as well as Figure Copy, Verbal Fluency, and Coding.

There were several notable discrepancies in the composition and neuropsychological profiles of the resulting clusters between the two approaches. Across all three clusters, we observed

the above-mentioned pattern of greater labeling of impairment in those without a cognitive diagnosis when using CNP criteria (Figure 2). Specifically, each of the CNP clusters contained a larger proportion of those who had been given no diagnosis compared to the ANP clusters (first cluster: $X^2(1) = 34.02$, $p < .001$; second cluster: $X^2(1) = 34.19$, $p < .001$; third cluster: $X^2(1) = 6.63$, $p = .01$). Additionally, the first (i.e., minimal impairment) and third (i.e., multi-domain impairment) CNP clusters contained proportionally fewer patients with aMCI compared to the respective ANP clusters (first cluster: $X^2(1) = 15.83$, $p < .001$; third cluster: $X^2(1) = 7.67$, $p = .006$).

Lastly, we tested for differences in neuropsychological profiles between CNP and ANP for each cluster (Figure 3) by comparing mean RBANS subtest scores for those meeting CNP only (i.e., CNP+/ANP-), ANP only (i.e., CNP-/ANP+), and the overlapping group that met both (i.e., CNP+/ANP+). We focus our reporting on group main effects surviving multiple comparisons correction that showed significant differences between the CNP and ANP approaches. Detailed results are provided in Table S3.

For the first cluster, there were significant group differences in all memory subtests (adj. p's < .001) as well as Coding (adj. $p = .004$), with *post hoc* comparisons revealing that CNP+/ANP- individuals performed significantly better than CNP-/ANP+ individuals (adj. p's < .001). In fact, the CNP first cluster contained those with average scores (mean Z-scores > −0.71) on these subtests, suggesting that this approach labels impairment in those with generally intact cognitive function. For the second cluster, the groups differed significantly on all memory subtests except figure recall (adj. p's < .001), both language subtests (i.e., Naming and Fluency; adj. p's .018), and Digit Span (adj. $p = .040$). Although both groups exhibited memory deficits, *post hoc* comparisons indicated a more pronounced amnestic profile for the CNP-/ANP+ group, with memory scores in the exceptionally low range compared to scores in the low average range for CNP+/ANP- group. For the third cluster, there were significant group differences in all memory subtests except figure recall (adj. p's < .001), both visuospatial subtests (i.e., Figure Copy and Line Orientation, adj. p's < .001), Naming (adj. $p < .001$), and both attention subtests (i.e., Digit Span and Coding, adj. p's .025). Overall, both approaches yielded a pattern of multi-domain impairment, with memory, visuospatial (Figure Copy), language (Fluency), and attention (Coding) scores in the below average range (mean Z-scores > −1.40). However, *post hoc* comparisons revealed that the CNP+/ANP- group demonstrated a more amnestic profile (adj. p's < .001), whereas the CNP-/ANP+ group retained a multi-domain profile with greater impairments (below average scores, i.e., $Z < −1.40$) in certain memory (List Learning, Story Recall), visuospatial (Figure Copy), and attention (Coding) subtests (adj. p's < .001).

Discussion

This study evaluated two criteria for identifying cognitive impairment using the RBANS in a sample of non-demented Veterans seen in a memory disorders clinic. The majority of patients met both criteria and there was no difference in the ability of both the CNP and ANP criteria to identify patients with aMCI. This suggests that, within the context of identifying aMCI, the benefit of using the ANP approach instead of the CNP approach observed in many previous studies using research samples may be less pronounced when

applied to this clinical sample with various severities and etiologies. However, consistent with previous findings, the CNP approach was over-inclusive compared to the ANP approach, falsely labeling impairment in a greater number of patients determined to have no cognitive disorder by a consensus team and who had more intact cognition on average. Notably, all patients who failed validity testing met both criteria, illustrating the relevance of considering performance validity when applying algorithmic criteria in clinical contexts.

Despite some evidence that the CNP and ANP criteria performed similarly in this study, the most salient finding was the greater susceptibility to false positive labeling of impairment for the CNP than ANP criteria. Whereas the CNP criteria mis-labeled impairment in the majority of patients with depression (80%) and no diagnosis (69%), the ANP approach did so at roughly a third of the rate for those with depression and at nearly half the rate for those given no diagnosis. Further, the group meeting CNP criteria exhibited more intact memory performance on the RBANS (many well-within the average range) compared to the ANP criteria. These results mirror a consistent finding from this literature—that the CNP approach captures a subgroup of individuals who perform within normal limits, suggesting that they were falsely labeled as impaired (Bondi et al., 2014; Clark et al., 2013; Edmonds et al., 2015). Thus, our findings generally mirror those of previous studies that have demonstrated lower specificity of the CNP criteria (Bangen et al., 2016; Bondi et al., 2014; Clark et al., 2013; Edmonds et al., 2015; Edmonds, Eppig, et al., 2016; Jak et al., 2016).

Cluster analyses of neuropsychological test scores from patients meeting each criterion illustrated both the general convergence between the two approaches as well as notable differences in resulting neuropsychological profiles. The cluster analyses for both the CNP and ANP groups produced three clusters, each with relatively similar cognitive profiles overall. This likely reflects the substantial overlap between those meeting CNP and ANP (with 72% of the full sample meeting both criteria), and also suggests relative concordance in cognitive subtypes captured by these approaches. The three clusters contained 1) patients with minimal impairments (i.e., average to low average scores), largely comprised of those with no cognitive diagnosis; 2) patients with amnestic impairments (i.e., below average memory scores), predominantly those with aMCI or CD-NOS; and 3) patients with multi-domain impairments (i.e., below average to exceptionally low scores on certain memory, visuospatial, and attention subtests), most of whom were diagnosed with CD-NOS.

However, more detailed investigation of the composition of these clusters revealed several notable differences that illustrate the differential performance of the CNP and ANP criteria. First, the CNP criteria demonstrated more false positive labeling of impairment, including more individuals with no diagnosis in all three cluster-based subgroups and capturing a subset of individuals with intact neuropsychological performance in the first cluster. Second, whereas all three of the CNP-based clusters showed primarily memory impairment (albeit at different severities), the ANP criteria produced subgroups with more distinct neuropsychological profiles. Specifically, the first ANP cluster showed a pattern of subtle memory decrements in the low average range; the second ANP cluster demonstrated a more amnestic pattern of exceptionally low memory scores; and the third ANP cluster exhibited multi-domain (i.e., mixed) impairments in memory, visuospatial abilities, and attention.

These cognitive profiles map onto well-established clinical constructs and mirror evidence from prior studies that the ANP criteria are more likely to produce cluster-derived subgroups with dissociable neuropsychological profiles (Clark et al., 2013).

Although our findings are generally in line with previous research studies illustrating significant benefits of using the ANP over CNP criteria, we also observed some evidence that the two performed similarly in this clinical sample. Specifically, there was substantial overlap in patients selected by both approaches (72%) and no difference in identification of those with aMCI. Similarly, the first paper to introduce the ANP approach concluded that the majority of MCI diagnoses (based on a variety of criteria including CNP, ANP, and others) were consistent across schemes and remained stable over 17-month follow-up (Jak et al., 2009). Other studies found that diagnoses based on both the CNP and ANP criteria were associated with progression to dementia in separate models, although only the ANP remaining significant in a joint model (Jak et al., 2016; Wong et al., 2018). Further, we found that the best concordance with consensus diagnoses occurred when *both* criteria were met, which requires individuals to demonstrate consistent impairment (two tests per cognitive domain per ANP) at a certain level of severity (below −1.5 *SD* per CNP). Together, these findings support the validity of the general concepts behind these algorithmic criteria and suggest that they both yield meaningful diagnostic information.

Whereas previous studies provide strong evidence of enhanced sensitivity and specificity of the ANP criteria, our investigation yielded somewhat milder effects. We suspect that the benefits the ANP approach may be less apparent in this clinical setting for several reasons. Whereas most of the samples in previous studies were selected to represent a single etiology and particular stage of disease (i.e., MCI as a prodrome of Alzheimer's disease), the current study's sample was highly heterogeneous in terms of presenting problem, severity, and etiology. Previous research samples were also screened for many confounding factors including psychopathology and neurological comorbidity. Most of these studies used data from the Alzheimer's Disease Neuroimaging Initiative, in which participants were vetted for substance abuse history, prior head trauma, medication status, general health, and even vascular factors such as brain ischemia (Bondi et al., 2014; Edmonds, Delano-Wood, et al., 2016; Edmonds, Eppig, et al., 2016; Edmonds et al., 2015; Thomas et al., 2019). Contrastingly, the present sample of patients seen in a memory disorders clinic may have had one or many of these factors contributing to their clinical presentation. Thus, the specificity of the ANP criteria may be dampened in this heterogeneous clinical sample due to diagnostic heterogeneity and higher rates of confounding conditions compared to the research samples in which the criteria were developed. However, future studies directly comparing the application of these criteria in both research and clinical settings are needed.

A unique aspect of this study was the evaluation of the extent to which patients who had failed validity testing met CNP and/or ANP criteria. Research samples produce invalid test data at lower rates than clinical samples (McCormick et al., 2013), and the highest rates of performance invalidity are often observed in VA clinical settings (Martin & Schroeder, 2020). As expected, all patients who provided invalid data produced multiple impaired scores, causing them to meet both criteria for impairment. Thus, although both criteria are able to capture patients with clinically meaningful cognitive presentations, they are also

highly susceptible to confounding factors such as invalid test data. We chose to exclude patients who had failed validity testing from cluster analyses to preserve the interpretability of the resulting neuropsychological profiles. Although performance validity is not typically assessed in research studies, our results clearly illustrate the need for including formal validity testing when applying such criteria to real-world clinical settings.

We also observed that both CNP and ANP criteria were met by sizeable numbers of patients without cognitive disorders, including those with depression and those given no diagnosis by the consensus team. This may be due, in part, to the use of data from a single time point. Without reference to psychometric estimates of premorbid functioning, these criteria do not take into account whether scores represent declines from prior functioning. As such, some individuals were mischaracterized as cognitively impaired by these criteria, including individuals from historically disadvantaged backgrounds. Suboptimal but frequently available proxies of this include years/quality of education and race (Jones et al., 2011; Manly, 2005), and indeed the group who met both criteria (CNP+/ANP+) had proportionally fewer years of education and a greater number of African American patients. Thus, our data highlight the importance of considering demographic factors, both clinically and in research, in the determination of clinical diagnoses. Development of neuropsychological criteria should consider the salient issue of measurement of change and account for potential bias endemic to our methods. This may be accomplished by using the best available tests and norms and collecting comprehensive sociocultural and demographic information (APA Guidelines, 2021) to be clearly reported and included in statistical analyses.

The results of this study should be considered with the following limitations in mind. First, given the nature of the data available in this retrospective clinical dataset, there is inherent confounding of the CNP and ANP criteria and consensus diagnoses, since RBANS subtest scores had previously contributed to consensus diagnoses. However, diagnoses were also informed by other test data, clinical record review, the clinical interview, and input from the interdisciplinary team's evaluations. Nonetheless, this circularity precludes us from explicitly quantifying the diagnostic accuracy of the CNP and ANP criteria in this setting. Future prospective studies can use neuropsychological measures that did not contribute to diagnosis in order to directly test classification accuracy. Relatedly, this cross-sectional dataset permitted an assessment of concordance between these criteria and diagnosis only at a single timepoint. Differences in diagnostic accuracy may be more apparent when patients are monitored over time, such as in previous studies that have used progression to dementia as an outcome. Of note, the original ANP criteria, which require two impaired scores in at least one cognitive domain (Jak et al. 2009), have been adapted in some studies to include an additional approach for identifying impairment using the presence of one impaired score in three different domains (Bondi et al. 2014). We chose to use the original criteria in this study since the incremental benefit of adding the additional approach has not been empirically tested, and it is not clear how it should be applied to the RBANS which assesses more than three cognitive domains. Given that varying these parameters (i.e., number of impaired scores, number of test scores comprising each domain, number of domains) may impact study results, future work is needed to determine the optimal approach.

Second, sample characteristics should be considered when interpreting these results. Since all patients were presumably referred to this clinic due to memory concerns, our sample may contain a high proportion of patients with subjective and/or objective cognitive impairments. Therefore, we may be less likely to observe differential performance of the CNP and ANP criteria than previous research samples with a range of participants from cognitively normal to mild cognitive impairment. Alternatively, given the relative diagnostic heterogeneity of this clinical sample, the base rates of aMCI were low (only 16.8%) compared to prior studies that selected specifically for this condition. This may limit our assessment of the performance of the CNP and ANP criteria, as the detection of aMCI specifically is dependent on the frequency of this condition in the sample (Lezak et al., 2004). Our results demonstrate that although both criteria sensitively capture aMCI, this identification is non-specific, which may be expected given the diagnostic composition of this clinical sample.

Diagnostic schemas, such as the CNP and ANP, are necessary elements of evidence-based practice in neuropsychology. Such criteria represent useful rubrics for identifying and characterizing objective impairment and are important for preventing an over-reliance on clinical experience (Bowden, 2017). When applying the CNP and ANP criteria for identifying cognitive impairment to a real-world clinical setting, we observed subtle evidence of benefits to using ANP over CNP. However, in the absence of considering patient context and other clinical variables such as performance validity (Chelune, 2010), both criteria, and particularly the CNP approach, are over-inclusive in their labeling of cognitive impairment. These results highlight that evidence-based neuropsychological practice depends not only on the tools we use, but on how we wield them.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding Sources:

References

Almeida JAS, Barbosa LMS, Pais AACC, & Formosinho SJ (2007). Improving hierarchical cluster analysis: A new method with outlier detection and automatic clustering. Chemometrics and Intelligent Laboratory Systems, 87(2), 208–217. 10.1016/j.chemolab.2007.01.005

APA Presidential Task Force on Evidence-Based Practice. (2006). Evidence-based practice in psychology. The American Psychologist, 61(4), 271. [PubMed: 16719673]

American Psychiatric Association. (2000). Diagnostic and statistical manual of mental disorders (4th ed., text rev.). 10.1176/appi.books.9780890423349.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). 10.1176/appi.books.9780890425596

American Psychological Association, APA Task Force for the Evaluation of Dementia and Age-Related Cognitive Change. (2021). Guidelines for the Evaluation of Dementia and Age-Related Cognitive Change. Retrieved from https://www.apa.org/practice/guidelines/

Barker MD, Horner MD, & Bachman DL (2010). Embedded indices of effort in the repeatable battery for the assessment of neuropsychological status (RBANS) in a geriatric sample. The Clinical Neuropsychologist, 24(6), 1064–1077. [PubMed: 20544558]

Bangen KJ, Clark AL, Werhane M, Edmonds EC, Nation DA, Evangelista N, … Delano-Wood L (2016). Cortical amyloid burden differences across empirically-derived mild cognitive impairment subtypes and interaction with APOE ɛ4 genotype. Journal of Alzheimer's Disease, 52(3), 849–861.

Beatty WW, Ryder KA, Gontkovsky ST, Scott JG, McSwan KL, & Bharucha KJ (2003). Analyzing the subcortical dementia syndrome of Parkinson's disease using the RBANS. Archives of Clinical Neuropsychology, 18(5), 509–520. 10.1016/S0887-6177(02)00148-8 [PubMed: 14591446]

Bondi MW, Edmonds EC, Jak AJ, Clark LR, Delano-Wood L, McDonald CR, … Galasko D (2014). Neuropsychological criteria for mild cognitive impairment improves diagnostic precision, biomarker associations, and progression rates. Journal of Alzheimer's Disease, 42(1), 275–289.

Borghesani PR, DeMers SM, Manchanda V, Pruthi S, Lewis DH, & Borson S (2010). Neuroimaging in the clinical diagnosis of dementia: observations from a memory disorders clinic. Journal of the American Geriatrics Society, 58(8), 1453–1458. [PubMed: 20670380]

Bowden SC (2017). Neuropsychological Assessment in the Age of Evidence-based Practice: Diagnostic and Treatment Evaluations. Oxford University Press.

Chelune GJ (2010). Evidence-based research and practice in clinical neuropsychology. The Clinical Neuropsychologist, 24(3), 454–467. [PubMed: 18821179]

Clark LR, Delano-Wood L, Libon DJ, McDonald CR, Nation DA, Bangen KJ, … Bondi MW (2013). Are empirically-derived subtypes of mild cognitive impairment consistent with conventional subtypes? Journal of the International Neuropsychological Society, 19(6), 635–645. 10.1017/S1355617713000313 [PubMed: 23552486]

Denning JH (2012). The efficiency and accuracy of the Test of Memory Malingering Trial 1, errors on the first 10 items of the Test of Memory Malingering, and five embedded measures in predicting invalid test performance. Archives of Clinical Neuropsychology, 27(4), 417–432. [PubMed: 22543569]

Duff K, Clark HJD, O'Bryant SE, Mold JW, Schiffer RB, & Sutker PB (2008). Utility of the RBANS in detecting cognitive impairment associated with Alzheimer's disease: Sensitivity, specificity, and positive and negative predictive powers. Archives of Clinical Neuropsychology, 23(5), 603–612. [PubMed: 18639437]

Duff K, Hobson VL, Beglinger LJ, & O'Bryant SE (2010). Diagnostic accuracy of the RBANS in mild cognitive impairment: Limitations on assessing milder impairments. Archives of Clinical Neuropsychology, 25(5), 429–441. 10.1093/arclin/acq045 [PubMed: 20570820]

Duff K, Schoenberg MR, Patton D, Mold J, Scott JG, & Adams RL (2004). Predicting change with the RBANS in a community dwelling elderly sample. Journal of the International Neuropsychological Society, 10(6), 828–834. 10.1017/S1355617704106048 [PubMed: 15637773]

Duff K, Schoenberg MR, Patton D, Paulsen JS, Bayless JD, Mold J, … Adams RL (2005). Regression-based formulas for predicting change in RBANS subtests with older adults. Archives of Clinical Neuropsychology, 20(3), 281–290. [PubMed: 15797165]

Edmonds EC, Delano-Wood L, Clark LR, Jak AJ, Nation DA, McDonald CR, … Salmon DP (2015). Susceptibility of the conventional criteria for mild cognitive impairment to false-positive diagnostic errors. Alzheimer's & Dementia, 11(4), 415–424.

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. Journal of Alzheimer's Disease, 52(2), 685–691.

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(20), 2108–2116. 10.1212/WNL.0000000000003326 [PubMed: 27760874]

Folstein MF, Folstein SE, & McHugh PR (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. Journal of psychiatric research, 12(3), 189–198. [PubMed: 1202204]

Guilmette TJ, Sweet JJ, Hebben N, Koltai D, Mahone EM, Spiegler BJ, Stucky K, Westerveld M, & Conference Participants (2020): American Academy of Clinical Neuropsychology

consensus conference statement on uniform labeling of performance test scores, The Clinical Neuropsychologist, DOI: 10.1080/13854046.2020.1722244

Heaton RK, Miller SW, Taylor MJ, & Grant I (2004). Revised comprehensive norms for an expanded Halstead-Reitan Battery: Demographically adjusted neuropsychological norms for African American and Caucasian adults. Psychological Assessment Resources.

Heaton RK, Grant I, & Matthews CG (1991). Comprehensive norms for an expanded Halstead-Reitan battery: Demographic corrections, research findings, and clinical applications; with a supplement for the Wechsler Adult Intelligence Scale-Revised (WAIS-R). Psychological Assessment Resources.

Horner MD, Bedwell JS, & Duong A (2006). Abbreviated form of the test of memory malingering. International Journal of Neuroscience, 116(10), 1181–1186.

Howe LL, Anderson AM, Kaufman DA, Sachs BC, & Loring DW (2007). Characterization of the Medical Symptom Validity Test in evaluation of clinically referred memory disorders clinic patients. Archives of Clinical Neuropsychology, 22(6), 753–761. [PubMed: 17651938]

Institute of Medicine. (2001). Crossing the quality chasm: A new health care system for the 21st century. Washington, DC: National Academy Press.

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. The American Journal of Geriatric Psychiatry, 17(5), 368–375. 10.1097/JGP.0b013e31819431d5 [PubMed: 19390294]

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. Journal of the International Neuropsychological Society, 22(9), 937–943. [PubMed: 27029348]

Jones RN, Manly J, Glymour MM, Rentz DM, Jefferson AL, & Stern Y (2011). Conceptual and measurement challenges in research on cognitive reserve. Journal of the International Neuropsychological Society: JINS, 17(4), 593. [PubMed: 21411036]

Karantzoulis S, Novitski J, Gold M, & Randolph C (2013). The repeatable battery for the assessment of neuropsychological status (RBANS): Utility in detection and characterization of mild cognitive impairment due to Alzheimer's disease. Archives of Clinical Neuropsychology, 28(8), 837–844. 10.1093/arclin/act057 [PubMed: 23867976]

Lezak MD, Howieson DB, Loring DW, & Fischer JS (2004). Neuropsychological assessment. Oxford University Press, USA.

Manly JJ (2005). Advantages and disadvantages of separate norms for African Americans. The Clinical Neuropsychologist, 19(2), 270–275. [PubMed: 16019708]

Martin PK, & Schroeder RW (2020). Base rates of invalid test performance across clinical non-forensic contexts and settings. Archives of Clinical Neuropsychology. 10.1093/arclin/acaa017

Martin PK, Schroeder RW, & Odland AP (2015). Neuropsychologists' validity testing beliefs and practices: A survey of North American professionals. The Clinical Neuropsychologist, 29(6), 741–776. [PubMed: 26390099]

Martin PK, Schroeder RW, Olsen DH, Maloy H, Boettcher A, Ernst N, & Okut H (2019). A systematic review and meta-analysis of the Test of Memory Malingering in adults: Two decades of deception detection. The Clinical Neuropsychologist, 1–32. 10.1080/13854046.2019.1637027

McCormick CL, Yoash-Gantz RE, McDonald SD, Campbell TC, & Tupler LA (2013). Performance on the green word memory test following operation enduring freedom/operation Iraqi freedom-era military service: Test failure is related to evaluation context. Archives of Clinical Neuropsychology, 28(8), 808–823. 10.1093/arclin/act050 [PubMed: 23877970]

McDermott AT, & DeFilippis NA (2010). Are the indices of the RBANS sufficient for differentiating Alzheimer's disease and subcortical vascular dementia? Archives of Clinical Neuropsychology, 25(4), 327–334. 10.1093/arclin/acq028 [PubMed: 20430863]

Palmer BW, Boone KB, Lesser IM, & Wohl MA (1998). Base rates of "impaired" neuropsychological test performance among healthy older adults. Archives of Clinical Neuropsychology, 13(6), 503–511. [PubMed: 14590634]

Paulson D, Horner MD, & Bachman D (2015). A comparison of four embedded validity indices for the RBANS in a memory disorders clinic. Archives of Clinical Neuropsychology, 30(3), 207–216. [PubMed: 25783523]

Petersen RC (2004). Mild cognitive impairment as a diagnostic entity. Journal of Internal Medicine, 256(3), 183–194. 10.1111/j.1365-2796.2004.01388.x [PubMed: 15324362]

Petersen RC, Doody R, Kurz A, Mohs RC, Morris JC, Rabins PV, … Winblad B (2001). Current concepts in mild cognitive impairment. Archives of Neurology, 58(12), 1985–1992. 10.1001/archneur.58.12.1985 [PubMed: 11735772]

Petersen RC, & Morris JC (2005). Mild cognitive impairment as a clinical entity and treatment target. Archives of Neurology, 62(7), 1160–1163. 10.1001/archneur.62.7.1160 [PubMed: 16009779]

Rabin LA, Paolillo E, & Barr WB (2016). Stability in test-usage practices of clinical neuropsychologists in the United States and Canada over a 10-year period: A follow-up survey of INS and NAN members. Archives of Clinical Neuropsychology, 31(3), 206–230. [PubMed: 26984127]

Randolph C, Tierney MC, Mohr E, & Chase TN (1998). The repeatable battery for the assessment of neuropsychological status (RBANS): Preliminary clinical validity. Journal of Clinical and Experimental Neuropsychology, 20(3), 310–319. 10.1076/jcen.20.3.310.823 [PubMed: 9845158]

Sweet JJ, Goldman DJ, & Breting LMG (2017). Evidence-based practice in clinical neuropsychology. In Textbook of Clinical Neuropsychology (pp. 1007–1017). Taylor & Francis.

Thaler NS, Scott JG, Duff K, Mold J, & Adams RL (2013). RBANS cluster profiles in a geriatric community-dwelling sample. The Clinical Neuropsychologist, 27(5), 794–807. 10.1080/13854046.2013.783121 [PubMed: 23548200]

Thomas KR, Edmonds EC, Eppig JS, Wong CG, Weigand AJ, Bangen KJ, … Bondi MW (2019). MCI-to-normal reversion using neuropsychological criteria in the Alzheimer's Disease Neuroimaging Initiative. Alzheimer's & Dementia, 15(10), 1322–1332. 10.1016/j.jalz.2019.06.4948

Tombaugh TN (1997). The Test of Memory Malingering (TOMM): Normative data from cognitively intact and cognitively impaired individuals. Psychological Assessment. 10.1037/1040-3590.9.3.260

Wong CG, Thomas KR, Edmonds EC, Weigand AJ, Bangen KJ, Eppig JS, … Bondi MW (2018). Neuropsychological criteria for mild cognitive impairment in the Framingham heart study's old-old. Dementia and Geriatric Cognitive Disorders, 46(5–6), 253–265. 10.1159/000493541 [PubMed: 30391953]
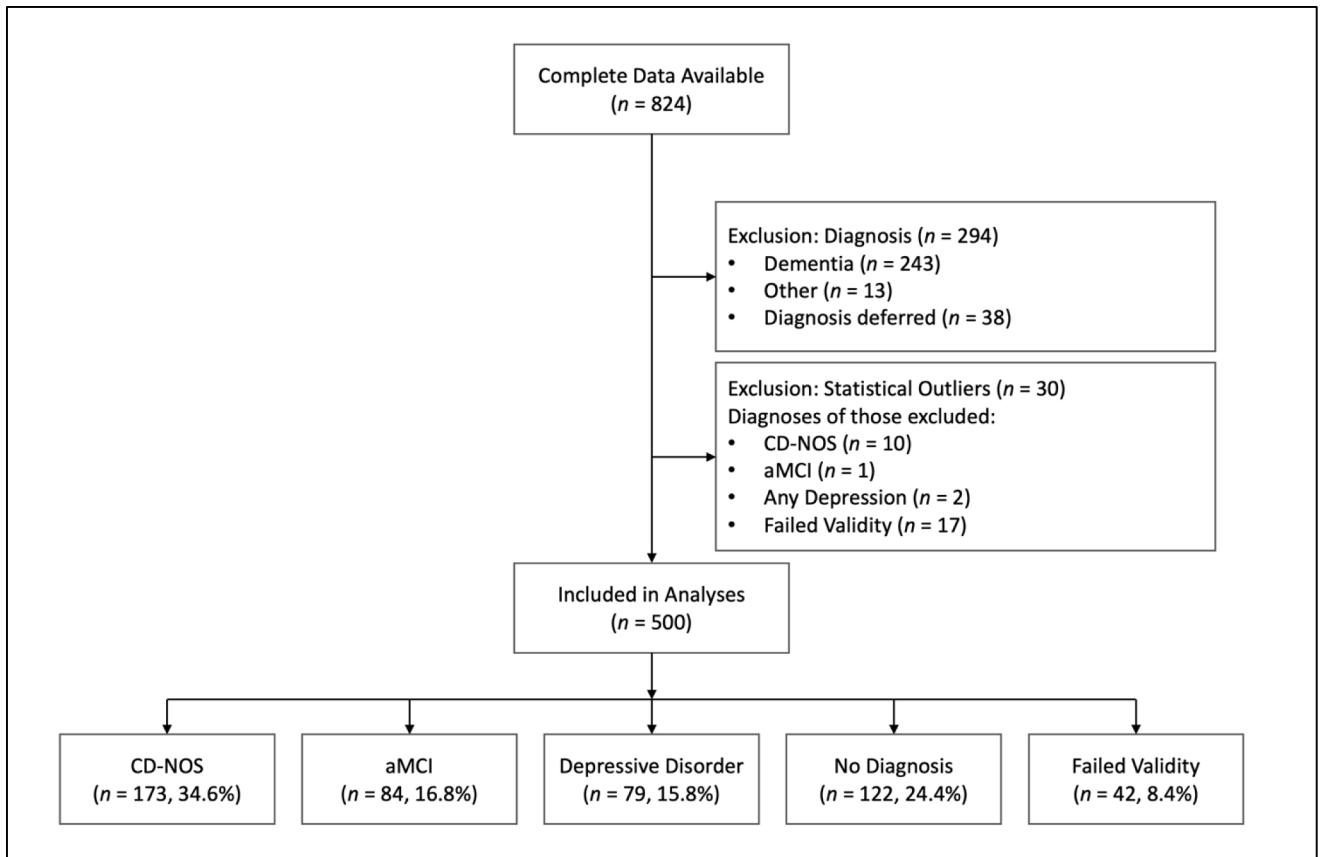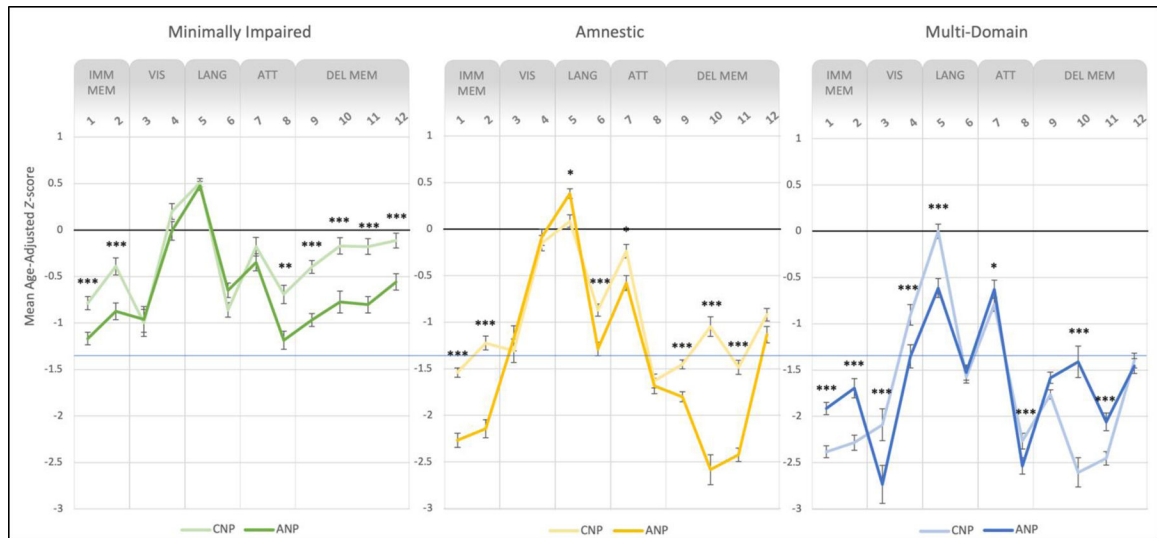
Figure 1.
Flow Chart of Patient Exclusion from Retrospective Database. Note that the Failed Validity group was excluded from cluster analyses.
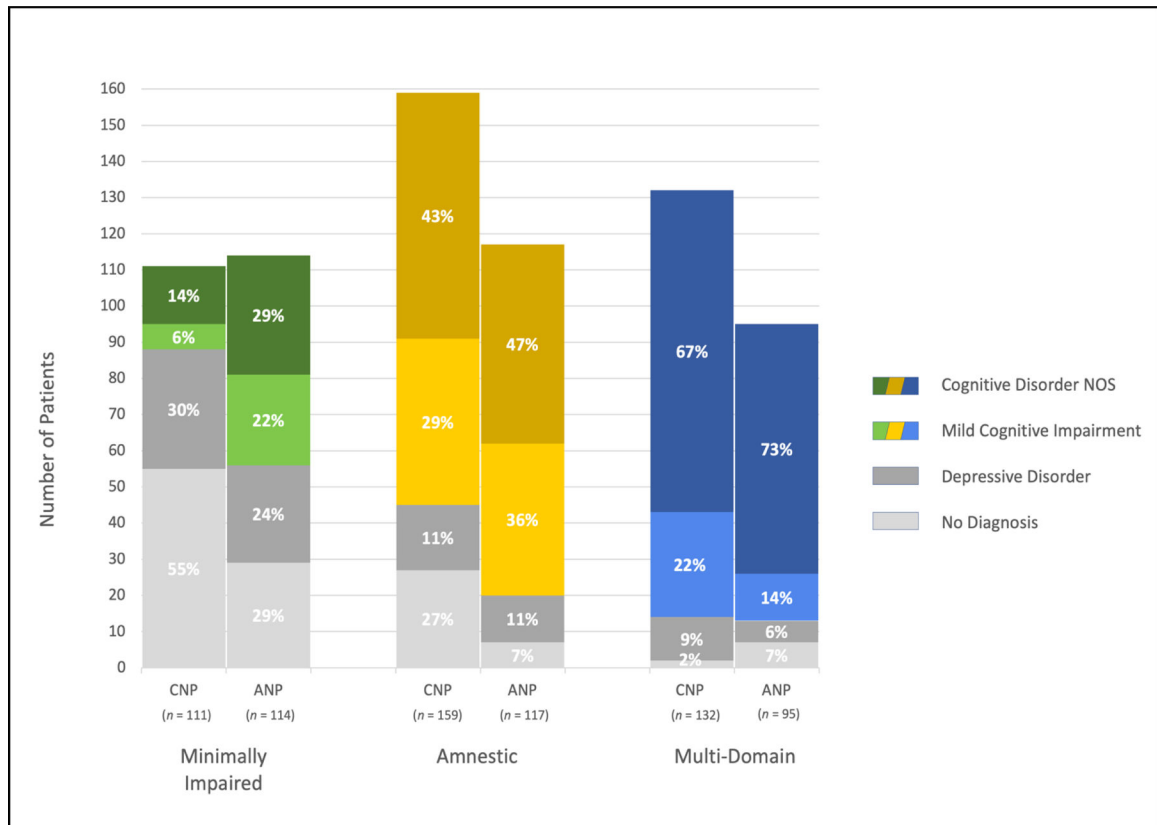
Figure 2.
Diagnostic composition of cluster-derived subgroups. Stacked bars represent the number of patients given each clinical diagnosis (shown in different colors per legend) within each cluster (x-axis): the first with minimal impairment (green/gray), the second with amnestic impairment (yellow/gray), and the third with multi-domain impairment (blue/gray). Bars are presented in pairs, with clusters of those meeting CNP criteria on the left and ANP criteria on the right of each pair. Data labels show the proportion of each cluster comprised of each diagnosis category.

Figure 3.
Neuropsychological profiles of cluster-derived subgroups. Lines represent performance on RBANS subtests for patients meeting CNP criteria (dark colors) and ANP criteria (light colors) for cognitive impairment. The first clusters (left panel, green) had scores in the average to low average range (i.e., minimal impairment), the second clusters (center panel, yellow) had an amnestic pattern of below average scores, and the third clusters (right panel, blue) had below average to exceptionally low scores in multiple domains. Error bars represent one standard error above and below the mean Z-score for each RBANS subtest. Subtests are grouped along the x-axis into the five RBANS Indexes, representing distinct cognitive domains. Asterisks indicate significant group differences in performance within each cluster. Note. $**p$ .01, $***p$ .001, Bonferroni corrected.

Table 1.

Descriptive Characteristics of Groups Meeting Each Criterion.

	CNP−/ANP− (n = 48)	CNP+/ANP− (n = 84)	CNP−/ANP+ (n = 8)	CNP+/ANP+ (n = 360)
Demographics	M, SD or N, %	M, SD or N, %	M, SD or N, %	M, SD or N, %
Age (years)	67.23 ± 9.28	68.54 ± 9.32	62.88 ± 11.94	69.83 ± 8.57
Education (years) [*]	14.6 ± 2.78	13.83 ± 2.67	12.75 ± 1.75	12.69 ± 2.92
Sex (male)	45, 93.8%	79, 94%	8, 100%	353, 98.1%
Race [*]				
Caucasian	41, 85.4%	72, 85.7%	7, 87.5%	249, 69.2%
African American	6, 12.5%	12, 14.3%	1, 12.5%	104, 28.9%
Hispanic/Latino	1, 2.1%	0, 0%	0, 0%	4, 1.1%
Other/not specified	0, 0%	0, 0%	0, 0%	3, 0.8%
Consensus Diagnoses [*]				
Cognitive disorder NOS	0, 0%	16, 19%	0, 0%	157, 43.6%
Mild cognitive impairment	0, 0%	4, 4.8%	2, 25%	78, 21.7%
Depression disorder	13, 27.1%	20, 23.8%	3, 37.5%	43, 11.9%
No diagnosis	35, 72.9%	44, 52.4%	3, 37.5%	40, 11.1%
Failed validity	0, 0%	0, 0%	0, 0%	42, 11.7%
Raw TOMM scores				
TOMM Trial 1 or Trial 2 [a]	49.75 ± 0.56	49.76 ± 0.52	49.25 ± 0.88	48.14 ± 3.98
RBANS subtest age-corrected Z-scores				
List learning immediate recall	−0.31 ± 0.73	−0.84 ± 0.83	−0.6 ± 0.57	−1.86 ± 0.87
Story learning immediate recall	−0.03 ± 0.72	−0.37 ± 0.91	−0.49 ± 0.82	−1.65 ± 1.15
Figure copy	−0.14 ± 0.72	−1.03 ± 1.47	0.13 ± 0.83	−1.63 ± 1.92
Line orientation	0.54 ± 0.72	0.29 ± 0.73	0.82 ± 0.69	−0.55 ± 1.27
Naming	0.51 ± 0.32	0.41 ± 0.55	0.65 ± 0.11	0.09 ± 0.86
Fluency	−0.23 ± 0.77	−0.88 ± 0.76	−0.29 ± 0.89	−1.22 ± 0.89
Digit Span	0.46 ± 1.02	0.02 ± 1.02	−0.37 ± 0.96	−0.59 ± 0.95
Coding	−0.15 ± 0.7	−0.79 ± 1.12	−0.57 ± 0.82	−1.87 ± 1.11
List recall	−0.02 ± 0.77	−0.5 ± 0.78	−0.73 ± 0.39	−1.5 ± 0.75
List recognition	0.06 ± 0.55	−0.13 ± 0.82	−0.82 ± 0.47	−1.82 ± 1.86
Story recall	0.13 ± 0.65	−0.15 ± 0.86	−0.51 ± 0.68	−1.85 ± 1.14
Figure recall	0.33 ± 0.69	−0.11 ± 0.73	−0.2 ± 0.92	−1.13 ± 1.02
RBANS age-corrected index scores				
Immediate Memory	98.17 ± 8.64	90.33 ± 10.42	90.5 ± 10.61	72.12 ± 13.64
Visuospatial	102.73 ± 10.04	94.23 ± 14.31	107.75 ± 9.11	84.41 ± 15.38
Language	98.5 ± 7.57	92.37 ± 6.76	96 ± 8.46	89.16 ± 8.7
Attention	102.67 ± 9.66	93.88 ± 11.76	91.25 ± 15.81	79.37 ± 14.03
Delayed Memory	104.04 ± 7.82	96.8 ± 7.45	91.25 ± 10.29	74.58 ± 16.12
Total	100.96 ± 6.54	90.51 ± 6.8	92.63 ± 7.65	74.33 ± 10

Note. CNP: conventional neuropsychological criteria; ANP: actuarial neuropsychological criteria.

[a]TOMM scores reflect Trial 2 for individuals given both trials and reflect Trial 1 for those given the first trial only.

[*] Groups significantly differed on these variables. Due to the challenge of denoting all significant group differences and directionalities in this table, specific pairwise effects are described in the text.

Table 2.

Diagnostic Composition Per Criterion.

Consensus Diagnoses	Full Sample (n = 500) No.	CNP (n = 444) No.	%	ANP (n = 368) No.	%	Group Difference p-value
Cognitive disorder NOS	173	173	100	157	90.8	< .001***
Mild cognitive impairment	84	84	100	80	95.2	0.69
Depressive disorder	79	63	79.8	46	58.2	< .001***
No diagnosis	122	84	68.9	43	35.3	< .001***
Failed validity	42	42	100	42	100	ns[a]

Note. CNP: conventional neuropsychological criteria; ANP: actuarial neuropsychological criteria. Group differences determined using McNemar's test for related samples of binary data.

[a]Test statistics could not be computed for the failed validity group because the two groups were identical (i.e., contained all cases).