Published in final edited form as:

J Alzheimers Dis. 2011 January ; 25(1): 163–173. doi:10.3233/JAD-2011-101821.

Effects of Varying Diagnostic Criteria on Prevalence of Mild Cognitive Impairment in a Community Based Sample

Emily H. Trittschuh^{a,b}, Paul K. Crane^c, Eric B. Larson^{c,f}, Brenna Cholerton^a, Wayne C. McCormick^c, Susan M. McCurry^e, James D. Bowen^{d,g}, Laura D. Baker^{a,b}, and Suzanne Craft^{a,b}

^aGeriatric Research, Education, and Clinical Center, Veterans Affairs Puget Sound Health Care System, Seattle, Washington, USA

^bDepartment of Psychiatry & Behavioral Science, University of Washington School of Medicine, Seattle, Washington, USA

^cDepartment of Medicine, University of Washington School of Medicine, Seattle, Washington, USA

^dDepartment of Neurology, University of Washington School of Medicine, Seattle, Washington, USA

^eDepartment of Psychosocial & Community Health, University of Washington School of Medicine, Seattle, Washington, USA

^fGroup Health Cooperative Research Institute, Center for Health Studies, Seattle, WA, USA

^gSwedish Neuroscience Institute, Swedish Medical Center, Seattle, WA, USA

Abstract

Mild cognitive impairment (MCI) is proposed to be a prodrome to dementia in some older adults. However, the presentation of MCI in the community can differ substantially from clinic-based samples. The aim of the current study is to demonstrate the effects of different operational definitions of MCI on prevalence estimates in community-dwelling older adults. A consecutive series of 200 participants aged 65 and over from the Adult Changes in Thought (ACT) community-based cohort were approached to undergo comprehensive neuropsychological and medical evaluation; 159 were included in the final analyses. Nondemented subjects were categorized using various diagnostic criteria for MCI. In a novel approach, neuropsychological test scores were evaluated using an individualized benchmark as a point of test comparison, as well as traditional methods that entail comparison to age-based normative data. Diagnostic criteria were further subdivided by severity of impairment (1.0 vs. 1.5 standard deviations [sd] below the benchmark) and extent of impairment (based on a single test or an average of tests within a cognitive domain). MCI prevalence rates in the sample were highly dependent on these diagnostic factors, and varied from 11% to 92% of the sample. Older groups tended to show higher prevalence rates, although this was not the case across all diagnostic schemes. The use of an individualized benchmark, less severe impairment cutoff, and impairment on only a single test all produced higher rates of MCI. Longitudinal follow-up will determine whether varying diagnostic criteria improves sensitivity and specificity of the MCI diagnosis as a predictor for dementia.

Corresponding Author: Brenna Cholerton, PhD, VA Puget Sound Health Care System, GRECC-A-182, 9600 Veterans Dr SW, Tacoma, WA 98493, Telephone: (253) 583-2032, bchol@u.washington.edu, Fax: (253) 589-4073.

The authors have no financial disclosures or conflicts of interest.

Keywords

Age related memory disorders; aging; cognition; Alzheimer's disease; dementia; diagnosis; epidemiology; individual differences; neuropsychological tests; prevalence; normative

Introduction

Mild cognitive impairment (MCI) represents a possible intermediary state between normal cognition and dementia [1, 2]. The general concept of MCI has achieved widespread acceptance as a diagnostic entity; however, the development of a meaningful and agreed upon operational definition has been problematic. Petersen et al [3] initially characterized MCI as amnesic impairment that was presumed to be primarily associated with prodromal Alzheimer's disease (AD). The definition later expanded to include other cognitive domains, with the expectation that the initial pattern of impairment predicted various diagnostic outcomes [4, 5]. Other definitions of MCI arose concurrently [2, 6-8], leading to the primary challenges that continue to plague the study of MCI today: namely, substantial variability in the operationalization of the diagnosis, the methods used to classify MCI, and the study population to which the definition will be applied [1]. Particularly in community settings, reliable identification of cognitive changes that are consistent with underlying neurodegenerative disease is challenging when the multitude of other medical, psychological, and social factors are taken into consideration [9, 10].

Although there is support for MCI as principally a dementia prodrome, with conversion rates of 10-15% per year reported [4], these results are often from self-referred clinic samples [11, 12]. Reports from community samples are more variable; such studies often demonstrate higher initial prevalence rates of MCI, but also report lower subsequent conversion to dementia and greater rates of reversion to normal cognition [12, 13]. This diagnostic instability necessarily raises the question of whether widespread community screening for MCI will be feasible once effective interventions are available. Ignoring the community population, however, could prove disastrous. The number of people aged 65 years and older with AD is expected to increase by more than 50% by the year 2030 [14]. If truly effective disease-modifying treatments were to become available, such treatments would most likely be aimed at those patients in the very early stages of the disease, prior to the onset of significant functional and cognitive decline. An accurate operational definition of MCI that distinguishes it from normal aging is vital to the identification of those most likely to benefit from intervention. This may be particularly important in the community setting, where access to specialized memory clinics could be limited due to psychological, geographical, or socioeconomic barriers. The traditional methods used to diagnose cognitive impairment in specialized memory disorders clinics, however, likely need to be revised to effectively identify affected persons in a community setting.

The detection of cognitive impairment typically involves the use of standardized neuropsychological measures, which are then compared to traditional age-corrected norms [15]. A number of recent publications support taking a more "intra-individual" approach, thus making it theoretically possible to identify cognitive impairment in the individual at the earliest sign of change [16]. Such reports suggest the best indication of early progressive cognitive decline may be subtle changes from the individual's lifelong functioning [17, 18], rather than relying on the average performance within an age group as the standard of comparison. Building on this idea, the current study compares traditional methods of test interpretation (using age-based norms) to an individualized benchmark. Use of a premorbid estimate of cognitive function as a point of reference to compare with individual performance on neuropsychological tests potentially enhances both the sensitivity and

specificity of the diagnosis. For example, high functioning individuals who complain of cognitive decline but perform within the average range will more likely be identified as having MCI if their performance is compared to a premorbid standard than if they were classified using age-normed standardized data. Conversely, lower functioning individuals who may otherwise be classified as having MCI would be excluded from diagnosis if tests scores were consistent with premorbid estimates. This method of test interpretation is likely to be especially useful in large community samples, where detailed clinical histories may not be readily available, and in populations that include a high proportion of the oldest-old, for whom quality normative data are scarce.

The current study investigates the effects of varying diagnostic criterion factors on the range of potential MCI prevalence estimates within a well-characterized, community-based sample of older adults. We report the base rates observed for multiple MCI subtypes, and examine the differences in prevalence when various methods for diagnosis are implemented. In a novel approach, we compare the effects of two methods of cognitive test interpretation (standard normative data vs. the use of an individualized benchmark) on the prevalence of MCI.

Materials and Methods

Subjects

The sample included 200 consecutive participants from the Adult Changes in Thought (ACT) study, approximately 1/10 of those in the total cohort as of January 31, 2007 ($n = 2186$). Detailed study methods for ACT have been published [19, 20]. Briefly, the ACT cohort consists of a population-based sample of adults aged 65 and over without evidence of dementia who were recruited from greater Seattle area clinics serving patients of Group Health Cooperative, an established health care system in Washington state and Northern Idaho. At study entry and at subsequent biennial visits, individuals received standardized batteries of cognitive screening tests and measures of physical function, functional well-being, and general medical information. Those who score below a specified cutoff score of 86 (out of 100 total possible points) on the Cognitive Abilities Screening Instrument (CASI) [21, 22], a global cognitive functioning test similar in scope and measurement properties [23] to the Modified Mini-Mental State Examination (3MS) [24] are referred for a full dementia evaluation. Age-specific rates of dementia and AD from the ACT study are consistent with those found in other studies around the world [19].

For the current study, ACT participants not previously diagnosed with dementia were consecutively approached during a 3-month period for comprehensive cognitive assessment, regardless of their performance on screening measures. There were no additional selection criteria other than a routine study visit within this time period. Of the 200 contacted, 180 gave consent to participate in the current study, 18 refused and two died prior to their visit. An additional 18 participants were determined to meet criteria for dementia as a result of the consensus diagnosis process, and were not included in subsequent analyses. Not surprisingly, these individuals were older and had significantly lower cognitive scores than the remaining 162 participants. Of the remaining subjects, three were missing a preponderance of data in three or more cognitive domains and thus were excluded from further analyses, resulting in a final sample size of 159. The demographic information for those diagnosed with dementia, those who declined participation, and the overall ACT cohort at the time of data collection are included in Table 1. Table 2 provides the self-reported comorbid conditions for the sample. Participants of this substudy, selected consecutively over a fixed period of time, were representative of the overall ACT cohort.

This study was approved by the institutional review boards of the University of Washington, Group Health Research Institute (GHRI), and the VA Puget Sound Healthcare System. ACT participants gave informed consent via processes authorized by the GHRI Institutional Review Board.

Cognitive Examination

At the biennial visit, participants were administered the CASI. Subsequently, participants were given a neuropsychological test battery that assessed five cognitive domains: 1) memory (Logical Memory-delayed and Verbal Paired Associates-delayed from the Wechsler Memory Scale-revised [WMS-R] [25], Word List Recall-delayed and Constructional Praxis-Delayed from the Consortium to Establish a Registry for Alzheimer's Disease [CERAD] neuropsychological battery [26]), 2) language (Short Boston Naming Test from the CERAD, category fluency [27, 28]), 3) visuospatial (Constructional Praxis from the CERAD, Construction subscale from the Dementia Rating Scale [DRS] [29]), 4) attention (Trail Making Test, Part A [30], Attention subscale from the DRS), and 5) executive function (Trail Making Test, Part B, Initiation/Perseveration and Conceptualization subscales from the DRS). In addition, participants underwent a clinical and neurological evaluation and review of medical records. Results from these investigations were assessed at a consensus diagnosis case conference that included study physicians with expertise in dementia diagnosis, at least one psychologist, and study nursing/coordinating personnel. Standardized diagnostic criteria for dementia (DSM IV) [31] and for probable or possible Alzheimer's Disease (NINCDS-ADRDA criteria) [32] were applied.

Individualized Estimate of Premorbid Ability

Study participants were administered a 40-item test of vocabulary knowledge, the Shipley-2 [33]. This test requires the subject to read a target word and then circle the word (from four choices) that represents the best synonym for the target word. The Shipley-2 measures an aspect of cognitive ability that is considered part of crystallized knowledge, and is largely resistant to the decline typically associated with aging or early cognitive disorder [34]. Vocabulary knowledge and similar measures of reading ability are commonly used as estimates of premorbid cognitive ability (IQ). If the Shipley-2 could not be completed ($n=3$), produced a score that was greater than two sd from the mean ($n=1$), or was considered invalid due to language barriers ($n=2$), a substitute score was imputed using demographic information based on the Barona equation for full scale IQ [35].

MCI Diagnosis/Subtyping

MCI was diagnosed using available psychometric data according to the factors described in the following sections. To best describe the full extent of cognitive deficits in this older population, we chose an operationalization of MCI that is similar to that of cognitive impairment, no dementia (CIND), a definition used in many other population-based studies [9, 36]. As a result, subjective cognitive complaints and a complete absence of functional impairment were not required for a diagnosis of MCI. In addition, those with comorbid medical and psychological conditions that may influence cognitive performance were not excluded. Information from the consensus conference was used to determine the presence of significant functional impairments that would lead to a classification of dementia. Four MCI subtypes were generated: 1) MCI single domain, amnesic, 2) MCI multiple domain, amnesic, 3) MCI single domain, nonamnesic, and 4) MCI multiple domain, nonamnesic. The amnesic categories relied upon performance on delayed recall measures [37]. Immediate recall performance can be affected by cognitive domains other than memory, including attention and concentration, and thus was excluded from analyses.

Cognitive test raw scores were transformed into z-scores using published normative data. When possible, age-based normative data from the Mayo Older Adult Normative Studies (MOANS) [38-40] were used given the wide availability and use of these norms in the older population. When these norms were not available for specific tests, alternate age-based norms (supermarket fluency) [27], or age-, education-, and gender-based norms from the Cache County Aging Study were used (CERAD tests) [41]. We then applied a 2x2x2 factorial design for determining prevalence rates of MCI and MCI subtypes. The factors are described below:

Factor 1: Benchmark (Standard versus Individualized)—Standard benchmarks were based on reported means of published normative data for each test. A given test score was considered impaired if it was 1.0 or 1.5 sd below the normative mean (expressed as a z score). To create individualized benchmarks, we determined an adjusted mean based on Shipley-2 vocabulary performance (or a substituted Barona index if the Shipley-2 score was unavailable). The individualized benchmark score was transformed into a z-score and used as the benchmark for determining impairment, rather than the normative mean. For example, if a participant's individualized benchmark score was $z = 0.5$, a z score of -1 standard deviation (sd) below the normative mean would indicate performance 1.5 sd below that expected based on their individualized score.

Factor 2: Severity of impairment (1.0 versus 1.5 sd)—This factor is the amount of deviation from the benchmark (either the individualized benchmark or standard means) required to identify a cognitive deficit/impairment. We examined both 1.0 and 1.5 sd units below the determined benchmark.

Factor 3: Extent of impairment (Any versus Average)—The extent of impairment necessary to meet MCI diagnostic criteria was based on one of two approaches: 1) if a participant's performance on *any* of the tests within a domain was below the relevant threshold, a deficit was identified in that domain; 2) if a participant's *average* performance within a cognitive domain was below the relevant threshold, a deficit in that domain was identified. Missing test data were not counted as deficits.

Analyses

Study population characteristics and demographics were examined to determine whether the sample was representative of the larger community cohort. Chi-square and t-tests were used to determine any differences in mean age, education, or gender. Frequency analyses were conducted to examine prevalence of various MCI subtypes and the effects of the three classification factors described above. All statistical analyses were computed in Stata 10 (Stata Corp., College Station, TX).

Results

Table 3 shows the prevalence rates for all-subtype MCI in the entire sample and by age according to multiple diagnostic criteria. Prevalence of all-subtype MCI ranged from a low of 11.3% for the most conservative criteria set (a 1.5 sd cut-point for the average of test scores within a given cognitive domain as compared to the age-matched population mean), to a high of 91.8% with the most liberal criteria set (a 1.0 sd cut-point for any test score within a cognitive domain as compared to the individualized benchmark). In general, MCI prevalence was higher in the older age group; however this finding was not consistent across all diagnostic criteria. For example, a less stringent diagnostic cutoff (1.0 sd) produced similar prevalence rates between the age groups when only one test within a domain was required for impairment.

Table 4 shows prevalence rates of MCI subtypes according to multiple diagnostic criteria. Single domain, amnesic MCI tended to produce lower prevalence rates, and ranged from 2.5% to 10.7%. Conversely, multiple domain, amnesic MCI had a much wider range, from 1.9% to 61.0%, depending upon the diagnostic criteria applied. When the two groups were combined, amnesic MCI rates in the sample ranged from a low of 4.4% to a high of 64.2%.

Factor 1: Benchmark (Standard versus Individualized)

Across operational definitions, the prevalence of any-subtype MCI was higher when the individualized benchmark was used than when standard normative data were used as the point of test comparison. Similar prevalence rates were found between the standard benchmark, 1.0 sd and the individualized benchmark, 1.5 sd criteria. Review of the distribution of individualized estimates revealed that the sample mean was approximately 0.5 sd higher than the standard normative data (figure 1). Standard scores 1 sd below standard norms are thus lower by 1.5 sd than the mean estimated individualized ability levels. Furthermore, given that the median value for the individualized estimate was .67, for a large proportion of the sample, scores 1 sd below standard norms would represent even more pronounced differences from individualized estimated ability levels.

To establish whether the standard benchmark, 1.0 sd criteria and the individualized, 1.5 sd criteria were identifying the same or different individuals, it was determined that when impairment on any test was required for MCI diagnosis, 84% of individuals were classified consistently across the two criteria sets, while 79% of individuals had concordant classifications when average impairment within cognitive domains was used.

Factor 2: Severity of impairment (1.0 versus 1.5)

As expected, a cutoff score of 1.0 sd below the benchmark produced higher prevalence rates than when a more stringent cutoff was used. Using this more lenient criterion produced prevalence rates that ranged from 26.4% to 91.8%, while the 1.5 sd criterion yielded prevalence rates that ranged from 11.3% to 71.1%.

Factor 3: Extent of impairment (any versus average)

When the MCI definition required impairment on only a single test within a cognitive domain, prevalence rates were much higher (range 45.3-91.8%) than when an average of tests within a domain was used (range 11.3-56.0%). When impairment on any test within a domain was sufficient for diagnosis, the rate of multiple domain, amnesic MCI was consistently higher than all other subtypes, ranging from 13.2% to 61%. Conversely, single domain, nonamnesic MCI yielded the highest prevalence rates (6.3-23.9%) when an average of test scores was required. The prevalence of any amnesic deficit ranged from 4.4 to 22.0% when an average of scores within a domain was used; whereas amnesic deficits ranged from 23.2 to 64.7% when any score within a domain was deemed impaired.

Discussion

Prevalence of MCI as operationalized in this non-demented community sample varied dramatically based on the three factors: choice of benchmark (individualized vs. standard normative comparisons), level of impairment (1.0 vs. 1.5 sd cutoff), and extent of impairment (single test vs. average impairment within a domain). When more stringent diagnostic criteria were used in combination with standard normative data, we obtained a prevalence rate of 11.3%, which is within the range reported in other community-based studies [1, 7, 42-44]. Varying the diagnostic criteria to require less severe deficits and the use of an individualized benchmark tended to produce much higher rates of MCI than generally reported, although not dissimilar to the prevalence of cognitive deficits recently

detailed by the Women's Health Initiative Memory Study [45]. The high prevalence rates produced by the most lenient criteria bring into question the potential predictive validity of these diagnoses. Interestingly, there has been some suggestion that, although different definitions of MCI produce dramatically different prevalence rates, the outcomes are not substantially different; that is, a high proportion appear to go on to develop dementia when compared to those without initial cognitive impairment [36, 46].

Standard cut-points for identifying impaired cognitive function have not yet been established for MCI. We chose to include both a moderate level of impairment (1.5 sd) as well as a milder level of impairment (1.0 sd), which is used in some MCI literature in an attempt to identify cognitive changes at the earliest possible point [47, 48]. As expected, a cut-point of 1.0 sd versus 1.5 sd yielded a considerably higher proportion diagnosed with MCI in this sample, which raises the question of whether the use of a 1.0 sd cutoff increases sensitivity at the cost of specificity. Interestingly, longitudinal data from the LEILA 75+ study suggest that the use of a 1.0 sd cut-point (vs. the 1.5 sd cut-point) produces higher relative predictive power as a result of a large increase in sensitivity with only a minor concurrent decrease in specificity [47]. Traditional screening cut-points inevitably must choose between missed diagnoses (false negatives) and erroneous diagnoses (false positives). Additional longitudinal research will help illuminate choice of appropriate cut-points to balance the trade-off between higher sensitivity or higher specificity in identifying MCI in the population.

When a single impaired score within a domain was used as the standard for diagnosis, the difference was a substantial two- to four-fold increase in MCI cases as compared to requiring an average across all tests within a domain to be below the threshold. It is interesting to note that use of a single impaired test score yielded the highest rates of multiple domain, amnesic MCI, often reported as the most common subtype and potentially the most likely to lead to conversion to dementia [4, 45, 49, 50]. In comparison, use of an average of test scores yielded higher rates of nonamnesic, single domain MCI, which may be less frequently associated with ultimate conversion to AD [47, 49, 50]. The use of an average score is a conservative approach that might increase specificity of the MCI diagnosis; however, this enhancement may come at the cost of lessened sensitivity. The neuropsychological test battery used in this study includes a greater number of tests in the amnesic domain, thus use of the average makes it less likely to be identified as impaired than the other domains. Conversely, requiring only a single test to be impaired may improve sensitivity and allow for the earliest detection of MCI. Importantly, this method may enhance feasibility of MCI screening in large-scale studies, where time constraints limit the number of tests that can be used. However, it may also artificially inflate the prevalence of MCI when multiple tests are given in a single domain, thereby increasing the probability that one of the scores will fall in the impaired range by chance alone and thus reducing specificity. Indeed, Loewenstein et al [51] report that longitudinal outcome for those diagnosed with MCI on the basis of one test results in a much higher rate of reversion to normal.

These results highlight a central problem faced by both researchers and clinicians: namely that slight variations in diagnostic criteria notably alter the prevalence of MCI, which unquestionably leads to variability in sensitivity and specificity. The decision about which criteria to use can thus result in either overdiagnosis, which can create problems due to labeling what may turn out to be "normal" variation, or underdetection, which could delay potential treatments and individual decision-making until after deficits in activities of daily living are more apparent. Given the need to improve both the sensitivity and specificity of the MCI diagnosis in the community, we have explored a unique method in this study, the implementation of an individualized benchmark. Recent reports that focus on the earliest

identification of MCI (and even “pre-MCI”) suggest that an “intra-individual” approach that focuses on individual decline over time, subjective complaints, and detailed reports of functional status may best predict the very earliest stages of impairment [16-18]. Use of an individualized benchmark, which sets the point of comparison according to estimates of lifelong cognitive ability, may provide a surrogate index of decline when longitudinal data and detailed clinical reports are not available. Results from the current study produced similar prevalence rates between the individualized benchmark at a traditionally set level of impairment of 1.5 sd, and the standard benchmark when the impairment criteria were less strict (1.0 sd). For this sample, the distribution of individualized benchmark scores was centered approximately 0.5 sd above the general population mean; thus, using a less stringent standard for impairment compared to a standard benchmark identifies a similar proportion of individuals as using an individualized benchmark at a traditionally set level of impairment. As a result, rates of MCI were lower for this more highly educated group using traditional normative comparisons. Conversely, in a sample with lower levels of education, the rates of MCI would likely be higher using the standard normative approach, a hypothesis that is supported by recent research [52, 53].

Given that the mean level of education attained by this sample is much higher than in many other samples, including the samples from which the normative data were drawn, the generalizability of the prevalence rates reported here may be limited. This point, however, simultaneously illuminates the importance of using some type of individualized benchmark, particularly for those groups whose demographic characteristics fall outside of those represented by the standard norms. In particular, the use of an individualized benchmark may be very useful at higher and lower levels of educational attainment. This study introduces the use of the Shipley-2 vocabulary test as one potential measure of lifelong cognitive abilities; however, its use in a more culturally and educationally diverse population has not been fully assessed and it may not be a suitable measure for those who do not speak English or for whom English is a second language, groups with low levels of literacy, and groups with different educational backgrounds (either outside the U.S. or non-traditional in nature). Careful consideration of alternative measures of premorbid cognitive ability should be given when applying an individualized benchmark to any new population. Given the wide range of cognitive abilities represented in the general population, however, failure to take into account estimated premorbid ability in community-based cohorts can potentially result in reductions in both sensitivity and specificity [54, 55]. As stated above, a current important challenge in the refinement of MCI detection methods and diagnostic criteria is the marked difference between subject characteristics in self-selected memory clinic samples and population-based community samples [12]. Our results show that a substantial proportion of our community-based sample had impaired performance in nonamnesic cognitive domains. Given the range of medical conditions and psychosocial factors that are known to impact nonamnesic domains of cognitive function [56-58], it stands to reason that a community population might have higher rates of nonamnesic MCI. Individuals with nonamnesic subtypes of MCI may be less likely to progress to AD and other forms of dementia [47, 50]. Data from clinic-based samples, which tend to exclude people with significant comorbid conditions and which primarily focus on the identification of MCI as a precursor to Alzheimer's disease, may not generalize to the larger population, in which impaired cognition could potentially result from any number of factors. Our sample, which includes those with heart disease, cerebrovascular disease, diabetes and other medical/psychological illnesses, more closely represents the range of comorbid conditions expressed in the general population. Particularly if truly effective treatments become available, it will be increasingly important to refine methods of diagnosing early dementia in the community at large. An important point is that if a future medication proves efficacious for MCI diagnosed in a clinic-based sample, no assumptions should be made about its likely effect in the community.

Another factor that may have influenced prevalence rates for our sample was the relatively advanced age of the participants. Although there are few studies of MCI in this age group (sample mean age=81) it is not unreasonable to assume a higher MCI prevalence given the reports that consistently show higher rates of dementia with increasing age [59, 60]. In general, our results showed higher prevalence rates with advancing age, although this was not consistent across all diagnostic schemes. In our sample, a higher rate of refusal was noted among a high-risk group (those aged 80-85), which, given the relatively small sample size, may have affected the age-based prevalence rates. Recent research has suggested higher MCI prevalence rates with advancing age but lower overall conversion rates [61], bringing into question the validity and utility of standard diagnostic methods in an old-old population. It is important to note that the stability of the available normative data is reduced at the higher end of the age spectrum due to lower sample sizes. It thus may be particularly valuable to apply reliable individualized benchmarks when identifying the cognitive status of adults in higher age cohorts.

The current study has several limitations. Given the cross-sectional design, it is unclear whether the MCI diagnoses represent prodromal dementia, other factors such as frailty and normal variations in test performance, or some combination of all of these. Multiple health comorbidities in this very old community sample may yield a high prevalence of cognitive decline due to factors other than early dementia. In addition, although our sample was representative of the overall ACT study population, it was relatively small. The test battery employed in the current study was initially developed for the detection of dementia; therefore it includes a greater number of memory assessments, relatively more difficult memory tests as compared to other domains, and screening measures (specifically, the Dementia Rating Scale) that may have limited usefulness in an MCI population. These factors may have impacted our ability to adequately detect nonamnesic MCI in this group. Finally, there are a number of methods for estimating premorbid cognitive ability other than the Shipley-2. Although there is no current agreement as to the best method for such estimates, subsequent research may support the use of instruments that are more sensitive for this age group. Despite these concerns, this study emphasizes a number of important factors and challenges that must be addressed prior to the development of uniform diagnostic standards.

The present results show that varying criteria for MCI greatly affects prevalence rates to a degree that hampers inter-study comparisons and limits their utility in practice. At the very least, studies need to report sufficient details regarding their operationalization(s) of MCI to facilitate detailed understanding of the methods. Use of an individualized benchmark of estimated premorbid ability to assess impairment may enhance early detection of MCI. Future longitudinal studies will determine whether these methods increase diagnostic sensitivity with acceptable specificity. The identification of optimal criteria for early diagnosis will set the stage for other studies examining the relationship of early MCI to biomarker profiles, or assessing the effects of therapeutic interventions.

Acknowledgments

This work was supported by National Institute on Aging grants R01AG024180 and U01AG006781, the Department of Veterans Affairs, and the Group Health Research Institute. The funding sources did not provide scientific input for the study.

References

1. Petersen RC, Roberts RO, Knopman DS, Boeve BF, Geda YE, Ivnik RJ, Smith GE, Jack CR Jr. Mild cognitive impairment: ten years later. *Arch Neurol*. 2009; 66(12):1447–1455. [PubMed: 20008648]

2. Morris JC, Storandt M, Miller JP, McKeel DW, Price JL, Rubin EH, Berg L. Mild cognitive impairment represents early-stage Alzheimer disease. *Arch Neurol.* 2001; 58(3):397–405. [PubMed: 11255443]
3. Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment: clinical characterization and outcome. *Arch Neurol.* 1999; 56(3):303–308. [PubMed: 10190820]
4. Petersen RC, Doody R, Kurz A, Mohs RC, Morris JC, Rabins PV, Ritchie K, Rossor M, Thal L, Winblad B. Current concepts in mild cognitive impairment. *Arch Neurol.* 2001; 58(12):1985–1992. [PubMed: 11735772]
5. Winblad B, Palmer K, Kivipelto M, Jelic V, Fratiglioni L, Wahlund LO, 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. Mild cognitive impairment--beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. *J Intern Med.* 2004; 256(3):240–246. [PubMed: 15324367]
6. Breitner JC, Welsh KA, Gau BA, McDonald WM, Steffens DC, Saunders AM, Magruder KM, Helms MJ, Plassman BL, Folstein MF, et al. Alzheimer's disease in the National Academy of Sciences-National Research Council Registry of Aging Twin Veterans. III. Detection of cases, longitudinal results, and observations on twin concordance. *Arch Neurol.* 1995; 52(8):763–771. [PubMed: 7639628]
7. Graham JE, Rockwood K, Beattie BL, Eastwood R, Gauthier S, Tuokko H, McDowell I. Prevalence and severity of cognitive impairment with and without dementia in an elderly population. *Lancet.* 1997; 349(9068):1793–1796. [PubMed: 9269213]
8. Matthews FE, Stephan BC, Bond J, McKeith I, Brayne C. Operationalization of mild cognitive impairment: a graphical approach. *PLoS Med.* 2007; 4(10):1615–1619. [PubMed: 17973571]
9. Boyle PA, Buchman AS, Wilson RS, Leurgans SE, Bennett DA. Physical frailty is associated with incident mild cognitive impairment in community-based older persons. *J Am Geriatr Soc.* 2010; 58(2):248–255. [PubMed: 20070417]
10. Wilson RS, Schneider JA, Boyle PA, Arnold SE, Tang Y, Bennett DA. Chronic distress and incidence of mild cognitive impairment. *Neurology.* 2007; 68(24):2085–2092. [PubMed: 17562829]
11. Tabert MH, Manly JJ, Liu X, Pelton GH, Rosenblum S, Jacobs M, Zamora D, Goodkind M, Bell K, Stern Y, Devanand DP. Neuropsychological prediction of conversion to Alzheimer disease in patients with mild cognitive impairment. *Arch Gen Psychiatry.* 2006; 63(8):916–924. [PubMed: 16894068]
12. Farias ST, Mungas D, Reed BR, Harvey D, DeCarli C. Progression of mild cognitive impairment to dementia in clinic- vs community-based cohorts. *Arch Neurol.* 2009; 66(9):1151–1157. [PubMed: 19752306]
13. Matthews FE, Stephan BC, McKeith IG, Bond J, Brayne C. Two-year progression from mild cognitive impairment to dementia: to what extent do different definitions agree? *J Am Geriatr Soc.* 2008; 56(8):1424–1433. [PubMed: 18662209]
14. 2009 Alzheimer's disease facts and figures. *Alzheimers Dement.* 5(3):234–270. [PubMed: 19426951]
15. Jak AJ, Bondi MW, Delano-Wood L, Wierenga C, Corey-Bloom J, Salmon DP, Delis DC. Quantification of five neuropsychological approaches to defining mild cognitive impairment. *Am J Geriatr Psychiatry.* 2009; 17(5):368–375. [PubMed: 19390294]
16. Storandt M, Grant EA, Miller JP, Morris JC. Longitudinal course and neuropathologic outcomes in original vs revised MCI and in pre-MCI. *Neurology.* 2006; 67(3):467–473. [PubMed: 16894109]
17. Caselli RJ, Chen K, Lee W, Alexander GE, Reiman EM. Correlating cerebral hypometabolism with future memory decline in subsequent converters to amnesic pre-mild cognitive impairment. *Arch Neurol.* 2008; 65(9):1231–1236. [PubMed: 18779428]
18. Nunes T, Fragata I, Ribeiro F, Palma T, Maroco J, Cannas J, Secca M, Menezes C, Carmo I, Cunha G, Castelo Branco M, Guerreiro M, de Mendonca A. The outcome of elderly patients with

- cognitive complaints but normal neuropsychological tests. *J Alzheimers Dis.* 2010; 19(1):137–145. [PubMed: 20061633]
19. Kukull WA, Higdon R, Bowen JD, McCormick WC, Teri L, Schellenberg GD, van Belle G, Jolley L, Larson EB. Dementia and Alzheimer disease incidence: a prospective cohort study. *Arch Neurol.* 2002; 59(11):1737–1746. [PubMed: 12433261]
 20. Larson EB, Wang L, Bowen JD, McCormick WC, Teri L, Crane P, Kukull W. Exercise is associated with reduced risk for incident dementia among persons 65 years of age and older. *Ann Intern Med.* 2006; 144(2):73–81. [PubMed: 16418406]
 21. Teng EL, Hasegawa K, Homma A, Imai Y, Larson E, Graves A, Sugimoto K, Yamaguchi T, Sasaki H, Chiu D, et al. The Cognitive Abilities Screening Instrument (CASI): a practical test for cross-cultural epidemiological studies of dementia. *Int Psychogeriatr.* 1994; 6(1):45–58. discussion 62. [PubMed: 8054493]
 22. McCurry SM, Edland SD, Teri L, Kukull WA, Bowen JD, McCormick WC, Larson EB. The cognitive abilities screening instrument (CASI): data from a cohort of 2524 cognitively intact elderly. *Int J Geriatr Psychiatry.* 1999; 14(10):882–888. [PubMed: 10521888]
 23. Crane PK, Narasimhalu K, Gibbons LE, Mungas DM, Haneuse S, Larson EB, Kuller L, Hall K, van Belle G. Item response theory facilitated calibrating cognitive tests and reduced bias in estimated rates of decline. *J Clin Epidemiol.* 2008; 61(10):1018–1027. e1019. [PubMed: 18455909]
 24. Teng EL, Chui HC. The Modified Mini-Mental State (3MS) examination. *J Clin Psychiatry.* 1987; 48(8):314–318. [PubMed: 3611032]
 25. Wechsler, D. Wechsler Memory Scale - Revised. San Antonio, TX: The Psychological Corporation; 1987.
 26. Morris JC, Heyman A, Mohs RC, Hughes JP, van Belle G, Fillenbaum G, Mellits ED, Clark C. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology.* 1989; 39(9):1159–1165. [PubMed: 2771064]
 27. Kozora E, Cullum CM. Generative naming in normal aging: Total output and qualitative changes using phonemic and semantic constraints. *The Clinical Neuropsychologist.* 1995; 9(4):313–320.
 28. Lucas JA, Ivnik RJ, Smith GE, Bohac DL, Tangalos EG, Graff-Radford NR, Petersen RC. Mayo's older Americans normative studies: category fluency norms. *J Clin Exp Neuropsychol.* 1998; 20(2):194–200. [PubMed: 9777473]
 29. Mattis, S. Dementia Rating Scale: Manual. Odessa, FL: Psychological Assessment Resources; 1988.
 30. Reitan R. Validity of the Trail Making Test as an indicator of organic brain damage. *Perceptual Motor Skills.* 1958; 8:271–276.
 31. Diagnostic and statistical manual of mental disorders. 4th. Washington, DC: American Psychiatric Association; 1994.
 32. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology.* 1984; 34(7):939–944. [PubMed: 6610841]
 33. Shipley, WC.; Gruber, CP.; Martin, TA.; Klein, AM. Shipley-2. Los Angeles, CA: Western Psychological Services; 2009.
 34. Yuspeh RL, Vanderploeg RD, Kershaw DA. Normative data on a measure of estimated premorbid abilities as part of a dementia evaluation. *Appl Neuropsychol.* 1998; 5(3):149–153. [PubMed: 16318453]
 35. Barona A, Reynolds CR, Chastain R. A demographically based index of premorbid intelligence for the WAIS-R. *Journal of Consulting and Clinical Psychology.* 1984; 52(5):885–887.
 36. Tuokko H, Frerichs R, Graham J, Rockwood K, Kristjansson B, Fisk J, Bergman H, Kozma A, McDowell I. Five-year follow-up of cognitive impairment with no dementia. *Arch Neurol.* 2003; 60(4):577–582. [PubMed: 12707072]
 37. Roberts RO, Geda YE, Knopman DS, Cha RH, Pankratz VS, Boeve BF, Ivnik RJ, Tangalos EG, Petersen RC, Rocca WA. The Mayo Clinic Study of Aging: design and sampling, participation,

- baseline measures and sample characteristics. *Neuroepidemiology*. 2008; 30(1):58–69. [PubMed: 18259084]
38. Lucas JA, Ivnik RJ, Smith GE, Bohac DL, Tangalos EG, Kokmen E, Graff-Radford NR, Petersen RC. Normative data for the Mattis Dementia Rating Scale. *J Clin Exp Neuropsychol*. 1998; 20(4): 536–547. [PubMed: 9892057]
 39. Steinberg BA, Bieliauskas LA, Smith GE, Ivnik RJ. Mayo's Older Americans Normative Studies: Age- and IQ-Adjusted Norms for the Wechsler Memory Scale--Revised. *Clin Neuropsychol*. 2005; 19(3-4):378–463. [PubMed: 16120536]
 40. Steinberg BA, Bieliauskas LA, Smith GE, Ivnik RJ. Mayo's Older Americans Normative Studies: Age- and IQ-Adjusted Norms for the Trail-Making Test, the Stroop Test, and MAE Controlled Oral Word Association Test. *Clin Neuropsychol*. 2005; 19(3-4):329–377. [PubMed: 16120535]
 41. Welsh-Bohmer KA, Ostbye T, Sanders L, Pieper CF, Hayden KM, Tschanz JT, Norton MC. Neuropsychological performance in advanced age: influences of demographic factors and Apolipoprotein E: findings from the Cache County Memory Study. *Clin Neuropsychol*. 2009; 23(1):77–99. [PubMed: 18609337]
 42. Stephan BC, Matthews FE, McKeith IG, Bond J, Brayne C. Early cognitive change in the general population: how do different definitions work? *J Am Geriatr Soc*. 2007; 55(10):1534–1540. [PubMed: 17908056]
 43. Plassman BL, Langa KM, Fisher GG, Heeringa SG, Weir DR, Ofstedal MB, Burke JR, Hurd MD, Potter GG, Rodgers WL, Steffens DC, McArdle JJ, Willis RJ, Wallace RB. Prevalence of cognitive impairment without dementia in the United States. *Ann Intern Med*. 2008; 148(6):427–434. [PubMed: 18347351]
 44. Busse A, Bischof J, Riedel-Heller, Angermeyer MC. MCI: prevalence and incidence according to different diagnostic criteria. *Br J Psychiatry*. 2003; 182:449–454. [PubMed: 12724250]
 45. Rapp SR, Legault C, Henderson VW, Brunner RL, Masaki K, Jones B, Absher J, Thal L. Subtypes of Mild Cognitive Impairment in Older Postmenopausal Women: The Women's Health Initiative Memory Study. *Alzheimer Dis Assoc Disord*. 2010
 46. Fisk JD, Merry HR, Rockwood K. Variations in case definition affect prevalence but not outcomes of mild cognitive impairment. *Neurology*. 2003; 61(9):1179–1184. [PubMed: 14610117]
 47. Busse A, Hensel A, Guhne U, Angermeyer MC, Riedel-Heller SG. Mild cognitive impairment: long-term course of four clinical subtypes. *Neurology*. 2006; 67(12):2176–2185. [PubMed: 17190940]
 48. Sasaki M, Kodama C, Hidaka S, Yamashita F, Kinoshita T, Nemoto K, Ikejima C, Asada T. Prevalence of four subtypes of mild cognitive impairment and APOE in a Japanese community. *Int J Geriatr Psychiatry*. 2009; 24(10):1119–1126. [PubMed: 19449451]
 49. Rasquin SM, Lodder J, Visser PJ, Lousberg R, Verhey FR. 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*. 2005; 19(2-3):113–119. [PubMed: 15591801]
 50. DeCarli C, Mungas D, Harvey D, Reed B, Weiner M, Chui H, Jagust W. Memory impairment, but not cerebrovascular disease, predicts progression of MCI to dementia. *Neurology*. 2004; 63(2): 220–227. [PubMed: 15277612]
 51. Loewenstein DA, Acevedo A, Small BJ, Agron J, Crocco E, Duara R. Stability of different subtypes of mild cognitive impairment among the elderly over a 2- to 3-year follow-up period. *Dement Geriatr Cogn Disord*. 2009; 27(5):418–423. [PubMed: 19365121]
 52. Manly JJ, Bell-McGinty S, Tang MX, Schupf N, Stern Y, Mayeux R. Implementing diagnostic criteria and estimating frequency of mild cognitive impairment in an urban community. *Arch Neurol*. 2005; 62(11):1739–1746. [PubMed: 16286549]
 53. Tervo S, Kivipelto M, Hanninen T, Vanhanen M, Hallikainen M, Mannermaa A, Soininen H. Incidence and risk factors for mild cognitive impairment: a population-based three-year follow-up study of cognitively healthy elderly subjects. *Dement Geriatr Cogn Disord*. 2004; 17(3):196–203. [PubMed: 14739544]
 54. Ram N, Rabbitt P, Stollery B, Nesselrode JR. Cognitive performance inconsistency: intraindividual change and variability. *Psychol Aging*. 2005; 20(4):623–633. [PubMed: 16420137]

55. Christensen H, Mackinnon AJ, Korten AE, Jorm AF, Henderson AS, Jacomb P, Rodgers B. An analysis of diversity in the cognitive performance of elderly community dwellers: individual differences in change scores as a function of age. *Psychol Aging*. 1999; 14(3):365–379. [PubMed: 10509693]
56. Luchsinger JA, Reitz C, Patel B, Tang MX, Manly JJ, Mayeux R. Relation of diabetes to mild cognitive impairment. *Arch Neurol*. 2007; 64(4):570–575. [PubMed: 17420320]
57. Reitz C, Tang MX, Manly J, Mayeux R, Luchsinger JA. Hypertension and the risk of mild cognitive impairment. *Arch Neurol*. 2007; 64(12):1734–1740. [PubMed: 18071036]
58. Rozzini L, Vicini Chilovi B, Conti M, Delrio I, Borroni B, Trabucchi M, Padovani A. Neuropsychiatric symptoms in amnesic and nonamnesic mild cognitive impairment. *Dement Geriatr Cogn Disord*. 2008; 25(1):32–36. [PubMed: 18025827]
59. Ziegler-Graham K, Brookmeyer R, Johnson E, Arrighi HM. Worldwide variation in the doubling time of Alzheimer's disease incidence rates. *Alzheimers Dement*. 2008; 4(5):316–323. [PubMed: 18790458]
60. Fratiglioni L, Viitanen M, von Strauss E, Tontodonati V, Herlitz A, Winblad B. Very old women at highest risk of dementia and Alzheimer's disease: incidence data from the Kungsholmen Project, Stockholm. *Neurology*. 1997; 48(1):132–138. [PubMed: 9008508]
61. Hong TB, Zarit SH, Johansson B. Mild cognitive impairment in the oldest old: a comparison of two approaches. *Aging Ment Health*. 2003; 7(4):271–276. [PubMed: 12888438]

Figure 1. Sample distribution of the individualized benchmark scores

Table 1
Demographic data for study sample and entire ACT cohort

	Included in study (n = 159)		Refused/deceased/missing data (n=23)		Dementia cases (n=18)		ACT Cohort (n = 2186)	
	n	%	n	%	n	%	n	%
Age								
65-69.9	3	1.9	0	0	56	2.6		
70-74.9	33	20.7	2	1	252	11.53		
75-79.9	36	22.6	5	1	477	21.8		
80-84.9	44	27.6	12	4	656	30.0		
85-89.9	33	20.7	3	6	486	22.2		
90+	10	6.3	1	6	259	11.9		
mean (sd)	80.2 (6.5)		80.9 (4.8)	86.7 (5.8)**	81.1(6.3)			
Gender								
Male	65	40.9	10	7	835	38.2		
Female	94	59.1	13	11	1351	61.8		
Race/Ethnicity								
American Indian/Alaska Native	1	.6	1	0	4	0.1		
Asian	1	.6	0	0	81	3.7		
Black or African American	7	4.4	0	1	88	4.0		
White	145	91.2	21	14	1943	88.9		
Other, including mixed	4	2.5	1	2	33	1.5		
Unknown/missing	1	.6	0	1	37	1.7		
Hispanic or Latino	2	1.3	0	1	16	.7		

Study sample (n = 200)				ACT Cohort (n = 2186)
Refused/deceased/missing data (n=23)				
Included in study (n = 159)				
Dementia cases (n=18)				
Education				
0 - 11	13	8.2	2	2
12	21	13.2	4	3
13 - 15	35	22.0	6	6
16+	90 (58) *	56.6 (36.5) *	11	7
mean (sd)	15.4 (3.2)	14.8 (2.9)	13.9 (4.0)	14.8 (3.1)
CASI score				
mean (sd)	94.2 (4.6)	95.1 (3.3)	84.4 (9.0) **	94.0 (14.0)

* represents n and % of the sample with post-graduate education
** differed significantly from nondemented sample (p=.0001)

Table 2
Self-reported comorbid conditions in the sample

	Coronary heart disease [*]	Cerebrovascular disease ^{**}	Diabetes	Hypertension	Parkinson's disease	Depression screen positive [†]
Included n (%)	35 (22.0)	20 (12.6)	20 (12.6)	81 (50.9)	0 (0)	13 (8.2)
Missing n (%)	4 (2.5)	5 (3.1)	4 (2.5)	3 (1.9)	4 (2.5)	7 (4.4)

* Coronary heart disease is defined as history of myocardial infarction, angina, coronary artery bypass graft (CABG), or angioplasty

** Cerebrovascular disease is defined as history of stroke, transient ischemic attack, or carotid endarterectomy

† Based on the Center for Epidemiologic Studies Depression Screen (CES-D)

Table 3
Prevalence of any MCI in full sample and by age group

Benchmark ¹	Severity of impairment ²	Extent of impairment ³	Prevalence					
			Full sample (n=159)			≤80 (n=79)		
			n	%	n	%	n	%
Standard	1.0	Any	120	75.5	60	75.9	60	75.0
		Average	42	26.4	17	21.5	25	31.2
	1.5	Any	72	45.3	31	39.2	41	51.2
		Average	18	11.3	6	7.6	12	15.0
Shipley	1.0	Any	146	91.8	75	94.9	71	88.7
		Average	89	56.0	41	51.9	48	60.0
	1.5	Any	93	71.1	53	67.1	60	75.0
		Average	46	28.9	23	29.1	23	28.7

¹ Benchmark: comparison of cognitive test scores to standard diagnostic criteria or an individualized benchmark(determined by Shipley-2 performance,
² Severity of impairment: the amount of deviation from the mean required to indicate impairment (1.0 or 1.5 sd),
³ Extent of impairment: the diagnosis required either impairment on a single test within a cognitive domain or on an average of all tests within a domain

Table 4
Prevalence of MCI subtypes by multiple diagnostic criteria

MCI subtype	Benchmark ¹	Severity of impairment ²	Extent of impairment ³	Prevalence
Amnesic				n %
Single domain	Standard	1.0	Any	9 5.7
			Average	6 3.8
	Shipley-2	1.5	Any	16 10.1
			Average	4 2.5
Multiple domain	Standard	1.0	Any	6 3.8
			Average	7 4.4
	Shipley-2	1.5	Any	17 10.7
			Average	8 5.0
Nonamnesic	Standard	1.0	Any	62 39.0
			Average	8 5.0
	Shipley-2	1.5	Any	21 13.2
			Average	3 1.9
Single domain	Standard	1.0	Any	97 61.0
			Average	28 17.6
	Shipley-2	1.5	Any	49 30.8
			Average	7 4.4
Multiple domain	Standard	1.0	Any	29 18.2
			Average	21 13.2
	Shipley-2	1.5	Any	26 16.4
			Average	10 6.3
Nonamnesic	Standard	1.0	Any	19 11.9
			Average	38 23.9
	Shipley-2	1.5	Any	27 17.0
			Average	

MCI subtype	Benchmark ¹	Severity of impairment ²	Extent of impairment ³	n	%
Amnesic			Average	25	15.7
			Any	20	12.6
	Standard	1.0	Average	7	4.4
			Any	9	5.7
		1.5	Average	1	0.6
	Multiple domain				
			Any	24	15.1
	Shipley-2	1.0	Average	16	10.1
			Any	20	12.6
		1.5	Average	6	3.8

¹ Benchmark: comparison of cognitive test scores to standard diagnostic criteria or an individualized benchmark(determined by Shipley-2 performance,
² Severity of impairment: the amount of deviation from the mean required to indicate impairment (1.0 or 1.5 sd),
³ Extent of impairment: the diagnosis required either impairment on a single test within a cognitive domain or on an average of all tests within a domain