Published in final edited form as:

Alzheimer Dis Assoc Disord. 2014 ; 28(2): 113–121. doi:10.1097/WAD.000000000000018.

How well do MCI Criteria predict progression to severe cognitive impairment and dementia?

Mary Ganguli, MD, MPH^{1,2,3}, Ching-Wen Lee, PhD¹, Beth Snitz, PhD², Tiffany Hughes, PhD¹, Eric M. McDade, MD², and Chung-Chou H. Chang, PhD^{4,5}

¹Department of Psychiatry University of Pittsburgh School of Medicine, Pittsburgh, PA

²Department of Neurology University of Pittsburgh School of Medicine, Pittsburgh, PA

³Department of Epidemiology University of Pittsburgh Graduate School of Public Health

⁴Department of Medicine University of Pittsburgh School of Medicine, Pittsburgh, PA

⁵Department of Biostatistics University of Pittsburgh Graduate School of Public Health

Abstract

Background—The International Working Group (IWG) criteria for mild cognitive impairment (MCI) have variable utility for predicting progression to dementia, partly depending on setting. We explored an empiric approach to optimize the criteria and cutpoints in a population study.

Methods—In a cohort of adults aged 65+, we identified 1129 individuals with normal or only mildly impaired cognition by cognitive classification, and 1146 individuals without dementia (Clinical Dementia Rating <1). Operationally defining the IWG criterion set, we examined its sensitivity and specificity for the development of severe cognitive impairment and dementia (CDR≥1) over four years. We then disaggregated the criteria and used Classification and Regression Tree (CART) analyses to identify the optimal predictive model.

Results—The operational IWG criteria had 49% sensitivity and 86% specificity for the outcome of severe cognitive impairment, and 40% sensitivity and 84% specificity for the outcome of dementia. CART modeling improved sensitivity to 82% for the cognitive outcome and 76% for the dementia outcome; specificity remained high. Memory scores were the most important predictors for both outcomes. The optimal cutpoints were around 1.0 SD below the age-education mean. The best fit was observed when prediction was modeled separately for each age-education group.

Conclusions—Objective cognitive measurements contributed more to the prediction of dementia than subjective and functional measures. Those with less education only required memory testing, while those with more education required assessment of several cognitive domains. Where only overall norms are available, the appropriate threshold will vary according to the individual's age and education.

INTRODUCTION

The purpose of defining mild cognitive impairment (MCI) is to consistently identify a group of individuals whose cognition, while worse than expected for age, is not sufficiently impaired to be in the dementia range. This descriptive function has been accomplished through various approaches, based on expert consensus.^{1–6} For the definition of MCI to also have diagnostic utility, it should reflect an entity that is relatively homogeneous with regard to underlying etiology or pathology, response to treatment, and/or outcome.^{6,7} The MCI syndrome is heterogeneous with regard to outcome, as it can result from a variety of underlying causes, e.g., progressive conditions such as Alzheimer's disease⁸ or Parkinson's disease,⁹ or other static, transient, or reversible conditions.^{4, 10–12} Further heterogeneity is introduced by demographic variation and comorbidity, particularly outside specialty clinical settings.

To optimize the utility of current MCI criteria, we examined prospectively gathered data from a large, population-based cohort of older adults. We have previously reported one-year outcomes of MCI¹⁰ defined by cognitive classification, an operational version of the International Working Group (IWG) criteria for MCI,¹ and Clinical Dementia Rating (CDR).¹³ Here, we evaluated the sensitivity, specificity, and predictive value of the IWG criterion set for progression to severe cognitive impairment/dementia over four years. We then explored these attributes in a series of models examining each criterion separately, at fixed or flexible thresholds (cutpoints), along with demographic characteristics, associated features, and risk factors that might influence outcomes.

METHODS

Study site and population

Our study cohort named the Monongahela-Youghiogheny Healthy Aging Team (MYHAT⁵) is an age-stratified random population sample drawn from the publicly available voter registration lists for a small-town region of Pennsylvania (USA.) Community outreach, recruitment, and assessment protocols were approved by the University of Pittsburgh IRB for protection of human subjects. Recruitment criteria were (a) age 65 years or older, (b) living within the selected towns, (c) not already in long-term care institutions. Individuals were ineligible if they (d) were too ill to participate, (e) had severe vision or hearing impairments, (f) were decisionally incapacitated. We recruited 2036 individuals over a two-year period, screening out 54 who at study entry exhibited substantial impairment by scoring <21/30 on the age-education-corrected Mini-Mental State Examination (MMSE). This approach involves adjusting the individual participant's score according to age and education.^{14,15} The remaining 1982 individuals underwent a detailed in-home assessment including, but not limited to, the elements below.

Assessments

At baseline and at each annual data collection cycle, we assessed cognitive functioning with a comprehensive test battery tapping several cognitive domains: Attention/ Processing Speed: (Trail Making Test A, Digit Span Forward); Executive Functions (Trail Making Test

B, Fluency for Initial Letters, Clock Drawing Test); Memory and Learning (Logical Memory, Immediate and Delayed; Visual Reproduction, Immediate and Delayed; Object Memory Evaluation); Language (Boston Naming Test, Fluency for Animals); Visuospatial Function (Block Design). We created a composite score for each domain, averaging z-scores across all tests in that domain.¹⁶ We also obtained the following self-reported measures: subjective cognitive concerns (SCC), ability to independently perform instrumental activities of daily living (IADL),⁵ and items comprising the AD8 scale, originally designed for a brief informant interview to detect dementia.¹⁷

Cognitive classification—At each cycle, we classified individuals as cognitively normal if all of their cognitive domain scores fell within 1.0 standard deviation (SD) of the appropriate demographic (age-sex-education) mean, based on our previously published norms;¹⁶ as severely cognitively impaired if two or more domain scores fell 2 or more SDs below the appropriate mean; and as mildly impaired if one or more scores fell 1.0 SD below the mean without meeting criteria for severe cognitive impairment.⁵ We excluded individuals with severe cognitive impairment at baseline.

Clinical Dementia Rating (CDR)—At each cycle, we assessed several aspects of everyday cognitively driven functioning to rate participants on the Clinical Dementia Rating (CDR) scale, disregarding neuropsychological data, as previously reported.^{5,13} We excluded individuals with a baseline summary rating of CDR = 1 (dementia).

Thus, this study included participants who at baseline were either normal or as mildly impaired by both cognitive and CDR classification.

Predictor variables

A. Demographic variables

1. Age: 65–74, 75–84 and 85+ years
2. Education : less than high school graduate (<HS) and high school graduate or more (≥ HS)

B. International Working Group (IWG)^{1,5} criteria for MCI: operational definition

1. Cognitive Impairment: For these analyses we included impairment in one or more domains. We first treated each cognitive domain composite score as a categorical variable dichotomized at a predetermined cutpoint of > 1.0 SD below the age-gender-education-adjusted mean. Subsequently we treated each composite score as a continuous variable, allowing the statistical model to identify the optimal cutpoint.
2. Subjective Cognitive concern: On a 21-item standardized questionnaire of subjective concerns,⁵ we first used a cutpoint of ≥ 2 (50th percentile at baseline) in the initial model. In subsequent models we treated this score as a continuous variable. In later models we also included the AD8 score as a separate variable.

3. Preserved functional independence: We used an IADL scale score <1 (90th %ile at baseline) in the initial model and then treated this score as a continuous variable in subsequent models.
4. Absent dementia: Clinical Dementia Rating of 0 or 0.5; as noted, we excluded participants with baseline summary rating of CDR = 1.
5. Preserved general mental status: Although this variable is not included in the IWG criteria¹, it was one of the previous Mayo criteria². As participants with age-education corrected MMSE scores <21 were already screened out at baseline, we used the uncorrected MMSE with a cutpoint = 21 in the initial analyses. We then treated the uncorrected MMSE as a continuous score in subsequent models, allowing the statistical model to select the optimal cutpoint.

C. Associated Features and Risk Factors—We included the following baseline assessments of risk and associated factors⁶ in the final models to determine whether they improved predictive value.

1. *APOE**4 genotype.
2. History of stroke (self-report).
3. mCES-D score, representing number of depressive symptoms. This is a modified version of the CES-D. It includes all 20 original CES-D items with each item scored as present/absent (1/0) over most of the preceding week, rather than 0–3 as in the original CES-D.¹⁸

Outcome Variables

We defined the outcomes (endpoints) in two ways.

- (a) Progression to cognitive classification of Severe Cognitive Impairment at any followup cycle 2–5, from normal cognition or MCI at baseline
- (b) Progression to Dementia (CDR = 1) at any Cycle 2–5, from CDR <1 at baseline.

Time to event was not considered in these analyses, i.e., all participants were classified as either progressed or not progressed to each of the outcomes at any Cycle 2–5. Like all prospective studies, over time the MYHAT cohort underwent attrition which was associated both with cognitive impairment and with risk factors for cognitive impairment.¹⁹ The statistical method we used here (see below) does not account for non-random attrition. We therefore excluded from these analyses those participants who died or dropped out without reaching either of the outcomes.

STATISTICAL METHODS

Descriptive statistics

We compared the baseline characteristics of participants who did and did not subsequently progress to the outcomes of (a) severe cognitive impairment and (b) CDR = 1, using chi-square or Fisher's Exact tests to compare categorical variables and Wilcoxon rank-sum tests to compare continuous variables.

We used the classification and regression tree (CART) method to identify factors predicting progression to severe cognitive impairment and dementia.¹⁹ CART is a commonly used approach for generating clinical decision rules to differentiate two groups, e.g., those who do and do not develop a given outcome. Based on binary recursive partitioning, CART segregates different values of predictors through a decision tree composed of progressive binary splits at a series of nodes (branching points). The optimal split is selected based on the impurity function, which measures homogeneity within a given group at a given node. In our models, we stopped a tree from growing (branching) further when there were 15 or fewer observations at a node. We controlled tree growth at a threshold of 0.01 on the complexity parameter (i.e. any additional split must increase model fit by at least 1%) and at a threshold of 0.5 on the progression rate (i.e., at the terminal node, at least 50% of participants develop the outcome). We used areas under the receiver operating characteristics curves (AUCs) to evaluate the fit of each model to the observed data. When all models were fit, we compared their AUCs.

For both the cognitive and the CDR outcome, we first examined our 5-component operational definition of IWG-MCI as a whole, calculating its sensitivity, specificity, positive and negative predictive values, and AUC. We then tested a series of 5 CART models (online Table e1). In Model 1, the predictors were the same fixed criteria as in our operational MCI definition, but disaggregated and entered individually, allowing CART to select those criteria that optimized correct classification (prediction of outcomes). In Model 2, we generated separate trees for the two educational groups, and treated the individual criteria flexibly as continuous variables, allowing CART to select not only the predictors but also the cutpoints. In Model 3, we added the AD8 score (continuous variable) as a candidate predictor. In Model 4, we generated separate trees for each of the 6 education (2 levels) × age (3 levels) groups. In Model 5, we added three more candidate predictor variables: *APOE**4, mCESD, and history of stroke.

RESULTS

At baseline (study entry, Cycle 1), the MYHAT cohort (N=1982) was representative of older adults in the targeted communities, with mean (SD) age of 77.6 (7.4) years and median educational level of high school graduate; 61.0% were women and 94.8% were of mixed European descent. Subsequently, 266 participants dropped out after baseline and contributed no followup data, and 570 participants had incomplete followup without reaching either study outcome Table e-2).

Cognitive Classification—At baseline, of 1982 participants, 1190 (60.0%) were cognitively normal, 697(35.2%) were mildly impaired, 54 (2.7%) were severely cognitively impaired, and 41 (2.1%) had insufficient cognitive data for classification. Eliminating the latter two groups, 1887 participants were either cognitively normal or MCI and also either CDR 0 or 0.5 (see below).

We further excluded 254 individuals who had no followup (i.e. had only baseline assessments), and 504 participants who underwent at least one followup assessment but were lost (died or dropped out) before Cycle 5, without reaching the severe cognitive

impairment outcome. Compared to the remaining 1129 individuals, these 504 were significantly older [mean (SD) age 79.1 (7.4) vs. 76.6 (7.1) years], with lower mean (SD) baseline MMSE scores [26.9 (2.4) vs. 27.3 (2.1)], higher mean (SD) baseline IADL score [0.03 (0.23) vs. 0.01 (0.15)], and greater proportions with impaired composite scores (1.0 SD below the overall mean) in all cognitive domains except visuospatial functioning. All P values were <0.05.

Of the 1129 individuals included in the analyses, 120 (10.6%) progressed to severe cognitive impairment at some point over four annual assessments. The proportions of progressors were identical (10.6%) among the 141 individuals with less than high school education, and the 998 participants with high school or greater education. Those who did progress were significantly older than those who did not progress, had significantly lower scores on the MMSE and all five cognitive domains, more subjective cognitive concerns, and higher AD8 scores. They were more likely to carry an *APOE*4* gene and to have higher scores on the depression symptom scale (Table 1).

Clinical Dementia Rating (CDR)—At baseline, 1413 (71.3%) participants were rated as normal (CDR=0), 546 (27.5%) as mildly impaired (CDR=0.5), and 23 (1.2%) as having at least mild dementia (CDR ≥ 1). Excluding the latter, 1959 participants were rated as either CDR=0 or CDR=0.5; further restricting the sample to those also cognitively classified as normal or mildly impaired left 1887 participants. We further excluded 242 individuals who had no followup, and 499 who underwent at least one followup assessment but died or dropped out before Cycle 5 without reaching the dementia endpoint. Compared to the remaining 1146 individuals, the 499 lost to followup were significantly older [mean (SD) age 79.3 (7.4) vs. 76.7 (7.1) years], and had lower mean (SD) baseline scores on the MMSE [26.8 (2.4) vs. 27.3 (2.1), higher baseline IADL scores [0.04 (0.23) vs. 0.01 (0.14)], and higher baseline AD8 scores [0.5 (0.9) vs. 0.4 (0.9)]. All P values were <0.05.

Of the remaining 1146 individuals, 45 (3.9%) progressed to dementia (CDR ≥ 1) over four annual assessments. Progression was observed in 14 (9.4%) of 149 individuals with less than high school education, and 31 (3.1%) of 997 participants with high school or more education. Those who progressed were significantly older than those who did not progress, and less likely to have at least high school education. They had significantly lower scores on the MMSE, all five cognitive domains, more subjective cognitive concerns, and higher scores on AD8 but not the IADL scale (Table 1)

Sensitivity, Specificity, and Predictive Value

In Table 2, we summarize sensitivity, specificity, and predictive value, for both outcomes, of the IWG criterion set as a whole and of each of the CART models 1–5. We present detailed results of the final (best-fitting) models in tabular form (Table 3) and in graphic form (Figures 1 and 2). Detailed results are tabulated for all models are provided in online supplemental Tables e3–e5.

For the outcome of severe cognitive impairment over 4 years (Table 2)—

Treating our operational IWG-MCI criterion set as a whole, with fixed cutpoints on each

criterion (cognitive domain cutoffs 1.0 SD below the demographically appropriate mean), we observed sensitivity and specificity of 49% and 86%.

In Model 1, allowing the CART model to select the predictors at the same fixed cutpoints, sensitivity increased to 50% and specificity increased to 93%. The tree first split (branched) at the 1.0 SD cutpoint below the mean on the memory domain. Of the 120 individuals who progressed to severe cognitive impairment, 60 scored above this memory cutpoint and remained false negatives. For those who scored below the cutpoint, the remaining splits were at the fixed 1.0 SD cutpoints of executive functioning, subjective cognitive concerns, and visuospatial functioning. Attention/speed, language, IADL, and MMSE were not selected.

In Models 2 onwards, we used cognitive domain scores unadjusted for demographics, allowing the CART model to select not only the criteria but also the optimal cutpoints. As more predictor variables were added, the models became progressively more complex, i.e., the trees produced more branches, and the AUC progressively improved until we fit Model 4 (Table 3, Figure 1) with separate trees for the 3 age groups within the two educational groups. Among the 25 individuals aged 65–74 with less than high school education, none progressed to severe cognitive impairment. In the remaining 5 age-education groups, the first split was on memory, with the cutpoints at scores 1.38, 2.06, 0.51, 0.99, and 1.11 SDs below the overall unadjusted mean (Table 3, Figure 1). These scores correspond to 1.04, 1.08, 1.30, 1.05, and 0.52 SDs below the age-education adjusted mean for each group. Among those with lesser education, no further assessments improved prediction. Among those with more education, all domains and the MMSE were selected for different subgroups. The presence of 6 subjective concerns improved prediction only among those aged 75–84 with intermediate memory scores (between 0.99 SD and 1.71 SD below the overall mean).

To summarize the CART models for the cognitive outcome, the IWG- MCI criterion set as a whole had fairly low sensitivity (49%) which improved to 82% when we allowed the CART model to determine the cutpoints and to develop different trees for different age and education groups. Positive predictive value (PPV) improved from 29% to 47% and then decreased to 41%. Specificity and negative predictive value (NPV) were already high at 93% and remained high, as expected for a rare event. The number of “false positives” (i.e. those who met the criteria at baseline but did not progress to the outcome) decreased from 144 (14%) cases for the IWG set to 16 (1.6%) for Model 4. The number of false negatives decreased from 61 (51%) to 48 (40%). The predictor (criteria) variables and optimal cutpoints selected by the models varied across the subgroups. In all models, age groups, and education groups, the first split was in the cognitive domain of memory, but the optimal cutpoint ranged from half a standard deviation below the overall mean (in the youngest and higher educated group) to two standard deviations below the mean (in the oldest and lower educated group). Models 2 and 3 were identical (i.e. the addition of the AD8 did not influence prediction); Models 4 and 5 were also identical, i.e. the addition of the associated features (*APOE*4*, stroke history, and depressive symptoms) did not influence prediction. The final models had significantly higher AUCs than the IWG criterion set and the preceding models (Table 1).

Subjective cognitive concerns

For the cognitive outcome, subjective concerns alone had sensitivity of 54%, specificity of 69%; positive and negative predictive values of 17% and 93%, and AUC = 0.637. For the CDR outcome, subjective concerns performed better, with sensitivity of 60%, specificity of 83%; positive and negative predictive values of 13% and 98%, and AUC = 0.732.

For the outcome of dementia by CDR 1 over 4 years (Table 2)—The IWG-MCI criterion set had even lower sensitivity (40%) for CDR 1 than it did for the cognitive outcome, which was as expected because the predictors were weighted towards cognition. Sensitivity improved to 76% with CART modelling. Positive predictive value (PPV), though low, improved somewhat from 9% to 19% and by modeling the fixed criteria, and declined somewhat to 14% when the criteria were made flexible. Specificity and NPV were already high and remained so. False positives decreased from 175 (16%) cases for the IWG set to 9 (0.8%) for Model 4, while false negatives remained unchanged at 27 (60%) and 28 (62%). The neuropsychological memory domain was the only criterion that entered the final tree, with cutpoints varying by age and education. We note that in Model 2 (Table e-3), 6 subjective cognitive concerns was predictive among those with greater education, but not after introducing age groups in Model 3 (Table e4). As with the cognitive outcome, the three associated /risk features did not influence the model.

In Model 4 (Table 3, Figure 2) among participants with less education and aged 65–74, none progressed to CDR 1; and 5 aged 75–84 who did progressed could not be captured. Of 54 individuals aged 85+, 9 progressed to dementia; the optimal cutpoint on the memory score was 2.06 SD below the unadjusted overall mean, corresponding to 1.08 SD below the mean for this age-education group. For those with higher education aged 75–84 and 85+, the optimal cutpoints were 1.82 and 1.49 SD below the overall mean, corresponding to 1.97 and 0.92 SD below the age-education means. Models 4 and 5 were identical, as for the cognitive outcome. All models 1–4 had significantly higher AUCs than the IWG criterion set, and the fit of each model 2–4 was significantly improved over the preceding one (Table 2).

CONCLUSIONS

In a large, prospective, population-based cohort of older adults, we operationalized the International Working Group criteria¹ for MCI at study baseline.⁵ Including all non-demented participants at baseline, we evaluated the capacity of this criterion set as a whole, and of its individual components, to predict progression to two outcomes: severe cognitive impairment as assessed by neuropsychological measures, and dementia as assessed by the Clinical Dementia Rating scale. We used an empirical statistical method, Classification and Regression Tree, to optimize sensitivity, specificity, and predictive value by selecting the predictors themselves and also the thresholds or cutpoints on the predictors. The criterion set as a whole had modest sensitivity and good specificity for both outcomes. When the criteria and thresholds were treated flexibly, the CART models primarily selected the objective cognitive measures, principally the memory domain as would be expected from previous studies.^{21–23} In the final model for the CDR outcome, only the memory domain provided any value, although subjective concerns did enter the CDR model before age was

considered. For the cognitive outcome, only memory was selected for those with less than high school education; all cognitive domains were selected for those with high school or greater education, with domains and cutpoints depending on age and education. The MMSE with a cutpoint of 24, and subjective cognitive concerns with a cutpoint of 6, also entered the final models. However, IADL and AD8 generated no branching, suggesting that the exclusion criterion related to everyday functioning might not be essential for predictive value, and thus the utility, of the definition of MCI. As expected, greater age and lesser education increased the likelihood of progression. The cutpoints generated by CART were lower for the older and less educated groups, relative to the overall mean. The optimal cutpoints were closer to 1.0 SD than to 1.5 SD below the age-education adjusted mean, consistent with the findings of another population study.²¹ In our final models, specificity remained excellent and sensitivity greatly improved.

None of the three independent factors known to be associated with dementia (stroke history, depression, and *APOE**4 genotype)^{24–25} improved the capacity of the criteria to predict either outcome even though stroke and depression were associated with progression to severe cognitive impairment. Biomarkers of specific underlying causal conditions, such as neurodegenerative or vascular disease, are usually viewed as enhancing specificity of diagnosis in a case of dementia. It remains to be determined whether they can improve prediction of dementia in normal or mildly impaired individuals. One previous study found the effects on sensitivity and specificity of *APOE**4 and the putative AD biomarkers were different in amnesic and non-amnesic MCI.²⁶

The likelihood of progression to dementia depends on the setting and circumstances under which the MCI is assessed; in a specialty clinical setting where the individual is seeking care for mild impairment, MCI is more likely to represent an underlying progressive brain disease than it is in a community setting where MCI is measured by systematic assessment of a population.^{11,12} In the clinical setting, the selection factors that lead patients/families to seek care for cognitive concerns may lead those concerns to have more predictive value than they do in the population setting. Not surprisingly, we found that subjective concerns, examined on their own, had better sensitivity and specificity for the CDR outcome than they did for the cognitive outcome; but for both outcomes, subjective measures had minimal impact once objective cognitive measures were included in the models. Studies of impaired individuals in clinical research settings usually require a reliable informant to provide collateral information. In our population setting, we used solely self-reported data on subjective concerns and IADL collected from dementia-free participants. Further, in operationalizing the IWG criteria for the population setting, we did not use the element of clinical judgment which is standard in specialty clinical settings, although the inter-rater reliability of clinical judgment across studies and settings is not easily established. In contrast to our findings, a previous population study which included informant report and clinical judgment found high sensitivity and low specificity.²² This suggests the testable hypothesis that both informant reports and clinical judgment enhance sensitivity, perhaps at the expense of specificity.

Caveats in interpreting our data include the following. Optimization produced by the CART approach is sample-dependent. Identical results will likely not be found in samples with

different demographic characteristics, different attrition rates, and in which different cognitive tests were administered to tap the same cognitive domains. Particularly for the CDR outcome, which has been reached by only 45 of our study participants to date, we likely lacked sufficient power to build a more predictive tree. Some potential predictors may have failed to predict the study outcomes because they were associated with subsequent attrition; it is untestable whether individuals who had these characteristics might have progressed to the outcomes of interest had they remained in the study. A limitation of the CART method is that it assumes attrition is completely random, i.e., that individuals who drop out have the same risk of progression as those who remain in the study. Trees with different branches might have been generated if we had no attrition, or if nobody or everybody who dropped out would have developed the outcomes by Cycle 5. Further, the CART method treats the outcomes as binary (progressed vs. not progressed) and does not model time to the outcome; this feature could be a limitation, for example, in applications where criteria are used to select participants for intervention trials of specified durations.

In summary, we found that objective cognitive measurements contributed more than subjective and functional measures did to the prediction of dementia in this population setting. Those with less education only required memory testing, while those with more education required assessment of several cognitive domains. The optimal cutpoints were around 1.0 SD below the age-education mean; where only overall norms are not available, the appropriate threshold will vary according to the individual's age and education. This approach may have particular benefits outside specialty settings, where there is greater heterogeneity of demographics and underlying diseases, and may also inform studies of preclinical disease. Overall, our results suggest some potential directions for enhancing the utility of the MCI criteria, and we encourage others to replicate our findings to determine their external validity.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors are grateful to all MYHAT study personnel and 2036 senior citizens who participated in the study.

The work reported here was supported in part by grants # R01AG023651, K07AG44395, K23AG038479, from the National Institute on Aging, NIH, US DHHS.

REFERENCES

1. Winblad B, Palmer K, Kivipelto M, et al. Mild cognitive impairment: beyond controversies, towards a consensus - report of the International Working Group on Mild Cognitive Impairment. *J Int Med*. 2004; 256:240–246.
2. Petersen RC, Smith GE, Waring SC, et al. Mild cognitive impairment: clinical characterization and outcome. *Arch Neurol*. 1999; 56:303–308. [PubMed: 10190820]
3. Morris JC, Storandt M, Miller JP, et al. Mild cognitive impairment represents early-stage Alzheimer disease. *Arch Neurol*. 2001; 58:397–405. [PubMed: 11255443]

4. Saxton JA, Snitz BE, Lopez OL, et al. GEMS Investigators. Functional and cognitive criteria produce different rates of conversion to dementia. *J Neurol Neurosurg Psychiatry*. 2009; 80:737–743. [PubMed: 19279031]
5. Ganguli M, Chang C-CH, Snitz BE, et al. Prevalence of mild cognitive impairment by multiple classifications: the MYHAT Project. *Am J Geriatr Psychiatry*. 2010; 18:674–68. [PubMed: 20220597]
6. Ganguli M. Can the DSM-5 framework enhance the diagnosis of MCI? *Neurology*. 2013; 81:1–6.
7. Albert MS, DeKosky ST, Dickson D, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*. 2011; 7:270–279.
8. Litvan I, Goldman JG, Tröster AI, et al. Diagnostic criteria for mild cognitive impairment in Parkinson's disease: Movement Disorder Society Task Force guidelines. *Mov Disord*. 2012; 27:349–356. [PubMed: 22275317]
9. Ganguli M, Snitz BE, Saxton JA, et al. Outcomes of mild cognitive impairment by definition: a population study. *Arch Neurol*. 2011; 68(6):761–767. [PubMed: 21670400]
10. Mitchell AJ, Shiri-Feshki M. Rate of progression of mild cognitive impairment to dementia--meta-analysis of 41 robust inception cohort studies. *Acta Psychiatr Scand*. Apr; 2009 119(4):252–65. [PubMed: 19236314]
11. Farias ST, Mungas D, Reed BR, et al. Progression of mild cognitive impairment to dementia in clinic- vs community-based cohorts. *Arch Neurol*. 2009; 66:1151–7. [PubMed: 19752306]
12. Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology*. 1994; 46:2412–2414.
13. Folstein MF, Folstein SE, McHugh PR. "Mini-Mental State:" a practical method for grading the cognitive state of patients for the clinician. *J. Psychiatric Research*. 1975; 12:189–198.
14. Mungas D, Marshall SC, Weldon M, et al. Age and education correction of Mini-Mental State Examination for English and Spanish-speaking elderly. *Neurology*. 1996; 46:700–706. [PubMed: 8618670]
15. Ganguli M, Snitz BE, Lee Ching-Wen, et al. Age and education effects and norms on a cognitive test battery from a population-based cohort: the The Monongahela-Youghiogheny Healthy Aging Team (MYHAT). *Aging & Mental Health*. 2010; 14:109–116.
16. Galvin JE, Roe CM, Powlishta KK, et al. The AD8. A brief informant interview to detect dementia. *Neurology*. 2005; 65:559–654. [PubMed: 16116116]
17. Ganguli M, Fu B, Snitz BE, et al. Mild cognitive impairment: incidence and vascular risk factors in a population-based cohort. *Neurology*. 2013; 80:2112–2120. [PubMed: 23658380]
18. Ganguli M, Gilby J, Seaberg E, Belle S. Depressive symptoms and associated factors in a rural elderly population: the MoVIES project. *American Journal of Geriatric Psychiatry*. 1995; 3:144–160.
19. Breiman, Leo; Friedman, JH.; Olshen, RA.; Stone, CJ. *Classification And Regression Trees*. Wadsworth & Brooks/Cole Advanced Books & Software; Monterey, CA: 1984. ISBN 978-0-412-04841-8
20. Busse A, Hensel A, Guhne U, et al. Mild cognitive impairment. Long-term course of four clinical subtypes. *Neurology*. 2006; 67:2176–218. [PubMed: 17190940]
21. Ritchie LJ, Tuokko H. Clinical decision trees for predicting conversion from cognitive impairment no dementia (CIND) to dementia from a longitudinal population-based study. *Archives of Clinical Neuropsychology*. 2011; 26:16–25. [PubMed: 21147863]
22. Restaino M, Matthews FE, Minett T, et al. Predicting risk of 2-year incident dementia using the CAMCOG total and subscale scores. *Age and Ageing*. 2013 ePub DOI 10.1093/ageing/aft082.
23. Elias-Sonnenschein LS, Viechtbauer W, Ramakers IH, et al. Predictive value of APOE-ε4 allele for progression from MCI to AD-type dementia: a meta-analysis. *J Neurol Neurosurg Psychiatry*. Oct; 2011 82(10):1149–56. [PubMed: 21493755]
24. Allan LM, Rowan EN, Firbank MJ, et al. Long term incidence of dementia, predictors of mortality and pathological diagnosis in older stroke survivors. *Brain*. 2011; 134(Pt 12):3716–27. [PubMed: 22171356]

25. Diniz BS, Butters MA, Albert SM, et al. Late-life depression and risk of vascular dementia and Alzheimer's disease: systematic review and meta-analysis of community-based cohort studies. *Br J Psychiatry*. 2013; 202(5):329–35. [PubMed: 23637108]
26. Vos SJ, van Rossum IA, Verhey F, et al. Prediction of Alzheimer disease in subjects with amnestic and nonamnestic MCI. *Neurology*. 2013; 10:1124–32. [PubMed: 23446677]

FIGURE 1.
Final Model (Model 4) for Severe Cognitive Impairment

Figure 2.
Final Model (Model 4) for Dementia (CDR 1)

Table 1**Cohort Characteristics at Baseline**

Baseline	Cognitive classification				CDR classification			
	Cognitively normal or MCI at baseline N=1129	Progressed to severe cognitive impairment n=120	Did not progress n= 1009	p value	CDR <1 at baseline N=1146	Progressed to CDR ≥1 n=45	Did not progress to CDR ≥1 n=1101	p value
Demographics								
Median; Mean (SD) age	77; 76.6 (7.1)	82; 80.9 (7.0)	76; 76.1 (7.0)	<0.001	77; 76.7 (7.1)	85; 83.9 (5.0)	77; 76.4 (7.1)	<0.001
% female	62.7	63.3	62.6	0.881	63.2	57.8	63.4	0.444
% HS educ	87.5	87.5	87.5	0.997	87.0	68.9	87.7	<0.001
MCI criteria								
Median; Mean (SD) MMSE	28; 27.3 (2.1)	25; 25.3 (2.6)	28; 27.6 (1.9)	<0.001	28; 27.3 (2.1)	25; 24.7 (3.2)	28; 27.4 (2.0)	<0.001
Median; Mean (SD) IADL	0; 0.01 (0.15)	0; 0.02 (0.13)	0; 0.01 (0.15)	0.072	0; 0.01 (0.14)	0; 0(0)	0; 0.01 (0.14)	0.652
Median; Mean (SD) subjective concerns	1; 2.1 (2.4)	3; 3.4 (3.2)	1; 2.0 (2.3)	<0.001	1; 2.1 (2.4)	4; 4.7 (3.9)	1; 2.0 (2.3)	<0.001
Median; Mean (SD) AD8	0; 0.4 (0.9)	0; 0.9 (1.3)	0; 0.3 (0.8)	<0.001	0; 0.4 (0.9)	1; 1.6 (1.8)	0; 0.3 (0.8)	<0.001
% worse than 1 SD below the overall unadjusted mean, by domain								
Memory	9.9	50.0	5.2	<0.001	8.7	55.6	6.8	<0.001
Attention	11.4	33.3	8.8	<0.001	10.9	31.1	10.1	<0.001
Executive	9.0	37.5	5.6	<0.001	8.8	40.0	7.5	<0.001
Language	10.0	42.5	6.1	<0.001	9.7	42.2	8.4	<0.001
Visuo-spatial	10.8	34.0	8.2	<0.001	10.7	22.5	10.3	0.031
Associated Features								
% APOE *4	21.4	31.5	20.2	0.007	21.1	30.0	20.8	0.160
Median; Mean (SD) mCES-D	0;0.8 (1.8)	0;1.1 (2.2)	0;0.7 (1.7)	0.023	0;0.8 (1.8)	0;1.2 (1.9)	0;0.7 (1.8)	0.074
% History of stroke	3.7	5.8	3.5	0.200	3.6	6.7	3.5	0.215

Table 2

Classification by IWG Criterion Set and CART Models for Both Outcomes

Outcome: Severe Cognitive Impairment (N=120 progressed, 1009 not progressed)					
	IWG :MCI fixed criteria¹	Classification and Regression Tree (CART) MODELS			
		Model 1²	Model 2³	Model 3⁴	Models 4 & 5⁵
Sensitivity %	49	50	70	70	82
Specificity %	86	93	91	91	86
PPV	29	47	47	47	41
NPV	93	94	96	96	98
AUC [*]	0.674	0.723	0.832	0.832	0.892

Outcome: Dementia byCDR 1 (N=45 cases, 1101 not progressed)					
Sensitivity %	40	49	64	64	76
Specificity %	84	91	86	86	80
PPV	9	19	15	16	14
NPV	97	99	98	98	99
AUC [*]	0.620	0.705	0.780	0.784	0.857

AUC for each model 2,3,4 was significantly higher than for the preceding model (P<0.05)

AUCs for Models 2/3 and 4 were significantly higher than for the IWG criteria (p<-0.0001).

AUCs for Model 2/3 was significantly higher than for Model 1, AUC for Model 4 was significantly higher than for Models 2/3 (p <0.01).

Model 5: same as Model 4 plus *APOE**4, depression symptoms, and stroke history

* AUCs for Models 1,2,3,4 were significantly higher than for the IWG criteria (P <0.05)

* Models 2 and 3 were identical.

¹ MCI Criteria: Operationally defined IWG criterion set as a whole with fixed cutpoints.

² Model 1: MCI criteria with CART modeling to select predictors, fixed cutpoints.

³ Model 2: MCI criteria with CART modeling to select predictors, flexible cutpoints, 2 separate education groups (< High School, High School)

⁴ Model 3: same as Model 2 plus AD8, flexible cutpoints

⁵ Model 4: same as Model 3 plus 3 age groups 65–74, 75–84, 85+ years

Table 3

Model 4: Final CART models for both outcomes

Final tree for outcome of Severe Cognitive Impairment (corresponds to Figure 1)								
Education	Less than high school education, N=141 (15 “cases”)			High school or greater education, N=988 (105 “cases”)				
Age in years	65–74	75–84	85+	65–74	75–84		85+	
N	25	67	49	415	457		116	
# of “cases”	0	8	7	20	52		33	
First node	n/a	Memory 1.38 SD below mean	Memory 2.06 SD below mean	Memory 0.51 SD below mean	Memory 0.99 SD below mean		Memory 1.11 SD below mean	
Subsequent nodes if any				Visuospatial 0.004 SD above mean	Executive 0.62 SD below mean	Memory 1.71 SD below mean	MMSE 24	
				Executive 0.14 SD below mean	Language 1.03 SD below mean	SubjConcern s6		
						Attention 0.32 SD below mean		
						Language 1.10 SD below mean		
Overall properties: Sensitivity 82%, Specificity 86%, PPV41%, NPV98%, and AUC 0.892								

Final Tree for outcome of Dementia (CDR 1; corresponds to Figure 2)								
Education	Less than high school education, N=149 (14 “cases”)			High school or greater education, N=997 (31 “cases”)				
Age in years	65–74	75–84	85+	65–74	75–84		85+	
N	26	69	54	419	459		119	
# of “cases”	0	5	9	1	15		15	
First node	n/a	n/a	Memory 2.06 SD below mean	n/a	Memory 1.82 SD below mean		Memory 1.49 SD below mean	