Optimising Mild Cognitive Impairment for Discriminating Dementia Risk in the General Older Population

Blossom C M Stephan (PhD)\textsuperscript{1*}, George M Savva (PhD)\textsuperscript{1}, Carol Brayne (MD)\textsuperscript{1}, John Bond (BA)\textsuperscript{2}, Ian G McKeith (MD)\textsuperscript{3}, Fiona E Matthews (PhD)\textsuperscript{4}, and the Medical Research Council Cognitive Function and Ageing Study (MRC CFAS)\textsuperscript{5}

\textsuperscript{1} Department of Public Health and Primary Care, Institute of Public Health, Cambridge CB2 0SR, UK
\textsuperscript{2} Institute of Health and Society and Institute for Ageing and Health, Newcastle University
\textsuperscript{3} Institute for Ageing and Health, Newcastle University, NE4 5PL
\textsuperscript{4} MRC Biostatistics Unit, Institute of Public Health, Cambridge CB2 0SR, UK
\textsuperscript{5} MRC CFAS www.cfas.ac.uk

*Address Corresponding Author:
Blossom Stephan
Department of Public Health and Primary Care
Cambridge University
Forvie Site, Robinson Way
Cambridge, CB2 0SR, UK

Funding

Blossom CM Stephan is supported by the ERA-Age Future Leaders of Ageing (FLARE) Postdoctoral Fellowship. This work was also supported by a Medical Research Council (MRC) Project Grant and has local ethics committee and multi-centre ethics committee approval. FM is funded by the Medical Research Council U.1052.00.013. MRC CFAS has been funded by the Medical Research Council and Department of Health. The funding bodies have no role in the study design, data collection, analysis or decision to publish.

Statistical Analysis: The statistical analysis was carried out by Blossom Stephan and George Savva

KEY WORDS Mild Cognitive Impairment, Incident Dementia, Sensitivity and Specificity, Population Screening

COUNTS:
WORD COUNT (exc. Abstract) 4,571
ABSTRACT WORD COUNT 255
NUMBER OF TABLES 3 (Including Appendix)
NUMBER OF FIGURES 3
ABSTRACT

**Background** Criteria for Mild Cognitive Impairment (MCI) predict dementia risk in the clinic. Dementia risk in the population is different and whether there is an optimal MCI derived threshold for discriminating at-risk from not-at-risk cases in the general older population is not known.

**Methods** Data were from the Medical Research Council Cognitive Function and Ageing Study (MRC CFAS). Two risk thresholds were derived from each of seven different concepts of MCI including: Mayo Clinic defined amnestic, non-amnestic, multiple and revised MCI, MCI based on Mini Mental State Examination (MMSE) derived categories, and the definitions of Cognitive Impairment No Dementia (CIND) and Age Related Cognitive Decline (ARCD). Receiver Operating Characteristic (ROC) analysis was used to compare the predictive validity of two-year incident dementia for each risk threshold across the different MCI definitions.

**Findings** MCI derived risk thresholds varied in their ability to predict dementia. MCI thresholds were accurate in identifying individuals not-at-risk of dementia progression (False Negative range, 0-3.4%). No MCI derived threshold accurately identified an at-risk group with a two year progression rate greater than 20%. Criteria for ARCD defined the threshold with the highest sensitivity and specificity for dementia conversion.

**Interpretation** MCI derived thresholds do not reliably identify individuals at-risk of incident dementia at two years when applied in the general population. A large subpopulation of individuals not-at-risk was more reliably identified. What is considered a sufficient level of accuracy for identification of individuals at increased risk of dementia depends on the motivation for screening and on the safety and efficacy of available interventions.
Introduction

The increasing incidence of dementia with the change in the world age demographic is a source of major public health concern. Early and accurate identification of individuals at high risk of dementia has become a research priority, especially with the promise of future preventative strategies to limit the expected rise of chronic neurodegenerative diseases consequent to an increased life span of the populations both in developed and developing countries. In the last decade identification of high risk cases has focused on the concept of Mild Cognitive Impairment (MCI), defined as an intermediate state between normal cognitive ageing and dementia. Numerous definitions for this state have been proposed, generally divided into those terms that capture normal age related cognitive change and those that capture pathological decline(1). The latter are a focus of research attention with the aim of identifying individuals at high risk of future dementia(2). However, it is not clear to what extent criteria for MCI discriminate those who will develop dementia when applied in population based samples.

While useful as an opportunity for early dementia risk screening in clinical samples, MCI as currently conceptualised does not appear to transfer well to the population setting. Studies of dementia incidence in MCI cohorts report up to 40-70% of cases remaining stable or reverting to normal cognitive function over time(3-6), with rates of progression generally depending on operationalisation of MCI criteria including: severity of cognitive impairment (e.g., 1, 1.5 or 2 standard deviations (SD) below the mean), underlying neuropsychological deficit (e.g., amnestic, non-amnestic and multi-domain subtypes) and psychometric test choice(5, 7-12). Additionally, many individuals who develop dementia at follow-up are found to have had a level of impairment outside the MCI range at baseline, and were therefore excluded from an MCI case diagnosis (here called non-classified or NON-C cases)(13). This finding raises the question of whether a risk threshold derived from a combination of the MCI and NON-C groups improves dementia risk prediction in the general population.
Here we compare risk thresholds derived from six commonly used MCI case definitions as well as the Mini Mental State Examination (MMSE)(14) with respect to their ability to discriminate those at risk of incident dementia in the older population. Using data from the Medical Research Council Cognitive Function and Ageing Study (MRC CFAS) we stratified all non-demented individuals into three risk categories, including low, moderate and high based on dementia incidence for the normal, MCI and NON-C groups previously reported for each MCI definition(5, 13). We then compared two-year dementia incidence across different combinations of the three risk categories to assess the current potential use of each as a screening tool for predicting dementia risk.

Methods

Study Design and Subjects
MRC CFAS is a large multi-centre population based prospective cohort study of individuals aged 65 years and older from the UK. Full details of the study design and procedures are published elsewhere and are briefly described here(15) (http://www.cfas.ac.uk). Individuals were randomly selected from the Family Health Service Authority lists in five areas in England and Wales including two rural (Cambridgeshire and Gwynedd) and three urban (Newcastle, Nottingham and Oxford). Baseline interviews were undertaken from 1991-1992. A two-phase screening procedure was used. At baseline screening, 13,004 individuals provided information on physical, behavioural and sociodemographic status in addition to aspects of health including self reported chronic conditions, and cognition using the MMSE. Individuals also completed selected items from the Geriatric Mental State (GMS) Examination(16). The GMS is a standardised psychiatric interview designed to detect dementia, depression, and other psychiatric illness in people aged 65 years and older by use of an algorithmic programme called the AGECAT (Automated Geriatric Examination Computer
Assisted Taxonomy)(16). Following the baseline interview a sub-sample of approximately 20% (n=2,640) were selected based on age, centre and cognitive ability, and weighted toward the cognitively frail to participate in a more detailed assessment interview. This included full mood and organicity sections of the GMS in addition to more detailed cognitive assessment using the Cambridge Cognitive Examination (CAMCOG)(17). Information on medical conditions, health status and functional ability were also collected. Respondents who underwent further assessment were asked to complete one or two yearly follow-ups(15). Data from the initial prevalence screen, first assessment and two year follow-up interviews (Data Version 8.2, December 2006) were used in this analysis, in addition to death notifications from the UK National Health Service Central Register.

**Diagnosis of Dementia, Depression and Anxiety**

The study diagnosis of dementia is based on the GMS AGECAT algorithm, defined as an organicity scale rating of 3 or above. This is comparable to dementia as diagnosed by the DSM-III-R(18-19). Using relevant GMS symptom items depression and anxiety were both defined as an AGECAT symptom level of 3 or above.

**Cognitive Assessment**

General cognitive function was assessed using the MMSE. In those definitions that required “normal general cognitive function” a MMSE score of 21 or less was used to indicate impairment(20-21). Memory and non-memory cognitive performance was evaluated using the subscales of the CAMCOG including: orientation, language, memory (learning, recent, and remote), attention and calculation, praxis, abstract thinking, and perception. Currently there is no consensus on the severity level (e.g., 1SD, 1.5SDs or 2SDs below the mean) for operationalising the MCI criterion of “objective cognitive impairment”(3, 8, 22-23). As such, memory impairment was defined using a cut-off score of 1SD below the mean (estimated using the 16th percentile as subscale scores are not normally distributed), on any of the three
CAMCOG memory sub-tests including learning memory, recent memory and remote memory. Impaired memory could therefore be in a single or multiple memory domains. Normal memory performance was defined as a score above the -1SD cut-off value on all three memory sub-tests. Non-memory impairment was also defined using a cut-off score of 1SD below the mean (16th percentile), on one or more of the following subscales including: orientation, language, attention/calculation, praxis, abstract thinking or perception. Normal non-memory performance was defined as a score above the -1SD cut-off value on all non-memory subscales. In addition, memory and non-memory impairment was defined using a stricter severity level of 1.5SDs below the mean, estimated using the 7th percentile. For Mayo Clinic defined MCI(24-25), severity scores were age standardised using five year age groups.

**Memory Complaint**

Memory complaint could be reported by the individual or their informant. A combined score was created from three questions including: (1) Have you had any difficulty with your memory?; (2) Have you tended to forget things recently?; and, (3) Has he/she had any difficulty with his/her memory? Answers for each question were coded into two categories (yes or no), from which individuals were dichotomized into non-complainers or complainers (positive response to one or more questions).

**Functional Disability: Activities of Daily Living (ADL) and Instrumental Activities of Daily Living (IADL)**

Functional disability was assessed using questions from the Modified Townsend Disability Scale, with an additional three items(26). Using information on a hierarchy of ADL/IADL disability, individuals in CFAS are classified into one of three groups. The first group included those individuals who showed no evidence of impairment in ADL or IADLs on items including washing, cooking hot meals, putting on shoes and socks, completing heavy housework or shopping and carrying heavy bags, and the individual can get around outside.
The second group included individuals with impairments only in IADLs, including those individuals who require regular help on items including heavy housework or shopping and carrying heavy bags. The third group included those individuals with deficits in basic activities of daily living (BADLs) including individuals who require help at least several times per week on items relating to washing, cooking and dressing, or if they are house/chair bound. ADL/IADL impairments are a key determinant of dementia and have been linked to cognitive, mental, physical and sensory causes. However, essentially intact functional ability is required in some, but not all case definitions of MCI(2), and there are currently no guidelines as to what is the most accurate ADL/IADL restriction for a MCI case diagnosis(3). For this analysis those individuals in Group 3 (impaired BADLs) were excluded at baseline. As such, mild IADL deficits were not an exclusion for MCI, similar to the Cardiovascular Health Study Cognition Study(27).

**Exclusion Criteria**

Definitions of MCI typically exclude individuals with psychiatric and vascular co-morbidity to help improve diagnostic accuracy(2). Individuals with anxiety or depression, or self reported history of stroke, heart attack, Parkinson’s disease or severe functional difficulty (defined above) were therefore excluded from MCI mapping. Of the 2,640 individuals at first assessment a total of 818 (31.0% back weighted to the UK population) non-demented individuals had one or more excluding conditions or missing health status information on one or more health variables and were excluded from our analysis. In addition, individuals with a diagnosis of dementia at first assessment were excluded. Of the 2,640 individuals at first assessment, 587 were diagnosed with dementia.

**MCI Classification**

MCI criteria included: Mayo Clinic defined, amnestic MCI (A-MCI: impairment in one or more memory domains)(25, 28), amnestic multi-domain MCI (M-MCI: impairment in
memory and non-memory domains)(29), non-amnestic MCI (N-MCI: impairment in one or more non-memory domains with preserved memory)(30), and revised MCI (R-MCI) that combined the latter three definitions, MCI based on MMSE defined categories, in addition to the definitions of Cognitive Impairment No Dementia (CIND)(31) and Age Related Cognitive Decline (ARCD)(32). The component criteria for each MCI definition vary resulting in different combinations of impaired/non-impaired abilities. Further, while some definitions have specific criteria for implementation (e.g., A-MCI), others lack operational definitions and explicit component criteria and therefore require interpretation as to the exact nature of the deficit (e.g., ARCD)(2).

In CFAS diagnostic criteria for Mayo Clinic defined MCI included: (1) not demented; (2) memory complaint reported by the individual or their informant; (3) normal general cognitive function; (4) no severe functional impairment; and, (5) objective memory and/or non-memory impairment. Mapping of the different Mayo Clinic definitions varied with regard to the last criterion: for A-MCI memory was required to be impaired and non-memory intact, for N-MCI memory was required to be intact and non-memory impaired, and for M-MCI both memory and non-memory test performance was required to be impaired. The definition of R-MCI combines all three criteria. For a diagnosis of CIND the individual was required to have a complaint of memory loss, impaired memory and/or non-memory performance and impaired general cognitive function. The definition of ARCD requires an objectively identified decline in general cognitive function as a consequent to the normal ageing process that is within normal limits for age. Further, individuals may report memory problems (e.g., with names, appointments or difficulty in solving complex problems). In CFAS criteria for an ARCD case diagnosis included impaired general cognitive function that may or may not be accompanied by SMC. Further, for ARCD we also required that memory test performance was not impaired, as memory impairment in combination with impaired general cognitive function was considered to be more severe than age related change. For a diagnosis of MCI based on
the MMSE, scores in the range of 22-26 were used to define the MCI group based on the results of this test alone. For each definition in order to be classified into the normal group an individual had to perform within normal limits on all criteria. As such NCI groups were defined differently depending on the MCI definition. All non-demented persons who were not normal but did not fulfil all MCI diagnostic criteria, such as for example a person who satisfies all A-MCI criteria but is classified as a non-complainer, were coded as non-classified (NON-C) for each definition(13). All individuals with missing MCI criterion data were excluded when each definition was mapped. Population based prevalence estimates and the two year dementia progression rates for the NCI, MCI and NON-C groups (back-weighted to the UK population) for each definition are shown in Table 1.

**Assigning Participants to Dementia Risk Categories**

Three dementia risk categories were defined including low, moderate and high based on the two year dementia progression rate for the NCI, MCI and NON-C groups outlined in Table 1. As shown in Table 1, across all definitions the NCI group had the lowest dementia progression rate and therefore was assigned the LR category. The high risk category included either the MCI or NON-C group depending on which had the highest dementia progression rate, as shaded in Table 1. The remaining group was considered to be at moderate risk. Based on the dementia risk categories each MCI definition therefore defines two thresholds for identifying at-risk individuals. The first compares the LR group to the MR and HR groups combined (Threshold 1: T1). This risk threshold represents a comparison between the NCI group to the risk defined by combining the MCI and NON-C groups. The second is the LR and MR groups combined compared to the HR group (Threshold 2: T2). Here, as defined above the high risk group for each definition is whichever of the NON-C or MCI group had the highest two year dementia progression rate (5, 13). The baseline demographic characteristics for the at-risk and not-at-risk groups defined by each risk threshold at each
cognitive severity level (1SD vs. 1.5SDs below the mean) are available in Supplementary Appendix Table A.

**Analysis**

Diagnostic accuracy for each of the two risk thresholds across the different MCI definitions was compared using the Receiver Operating Characteristic (ROC) method, with two year dementia incidence as the outcome. The ROC curve plots the sensitivity of each classification or the proportion of dementia cases identified, against one minus its specificity, where specificity is the proportion of those not developing dementia correctly identified. A perfect model has both sensitivity and specificity of 100%, and would be plotted in the upper left corner. The prognostic power of the MMSE score was also examined continuously and the ROC curve corresponding to using each MMSE score as a risk threshold for incident dementia is also included.

To compare the groups identified as being at-risk by each risk threshold we plotted the positive predictive value (PPV) for the at-risk group against the number of people classified as being at-risk for all MCI definitions. Further, to compare the groups identified as being not-at-risk by each risk threshold we plotted the size of the subgroup classified as not-at-risk against the negative predictive value (NPV) for all MCI definitions.

Analysis was undertaken using Stata (Version 10: Stata Corporation, College Station, Texas). Population based proportions were estimated from the sample using an inverse probability weight. This weight was applied to each individual and was the reciprocal of the probability of that individual being included in the analysis. The probability was estimated using weighted logistic regression accounting for the over-sampling in the assessment arm of older and more cognitively impaired participants and attrition (death and dropout) between baseline
and follow-up interviews. Binomial confidence intervals for proportions taking into account the probability weights were calculated using a Wilson score interval(33).

Results

There were 137 incident dementia cases over the two year follow-up period (population back-weighted two year incidence rate=4.4%). The ROC results for each MCI definition (at both cognitive severity levels) for each risk threshold are presented in Table 2. As expected, for each definition the severity level of -1.5SDs resulted in an improvement in specificity but with fewer dementia cases detected, i.e., lower sensitivity. All comparisons discussed below are based on the -1SD severity level.

**ROC Curve** A ROC plot comparing the sensitivity and specificity of each risk threshold for the different definitions of MCI is shown in Figure 1. As shown, definitions fall broadly into two classes: (1) those that are sensitive, that capture 80% or more of people who progress to dementia at two years with high false positive rates (including: ARCD [T1, T2], N-MCI [T2], MMSE≤24, CIND [T1] and A-MCI/N-MCI/M-MCI/R-MCI [T1]); and, (2) those that are specific, identifying fewer people who have progress to dementia, but with fewer false positives (including: R-MCI [T2], CIND [T2], M-MCI [T2], MMSE≤21, A-MCI [T2]). Overall, the component criteria for ARCD define the most sensitive and specific thresholds. Similar results for both risk thresholds were found due to the small number of individuals classified into the moderate risk group i.e., satisfy criteria for ARCD (Table 1: n=40, prevalence<1.5%). The MMSE provides classifications that are at least as good as, or better than most other MCI derived risk thresholds.

**Identifying a High Risk Cohort** Figure 2 compares the PPV of each at-risk group with the percentage of persons identified as being at-risk (see Table 2) by the two risk thresholds for
all MCI definitions. As shown, no method was able to identify an at-risk group with a
dementia risk of more than 20%, apart from the very small group with an MMSE≤19.
However this is considerably higher than the 4% risk seen in the general population. Overall,
at-risk case selection was variable. For example, when using M-MCI defined thresholds, less
than 5% of the sample was identified as being at-risk with approximately 20% of these
individuals progressing to dementia at two years. In contrast, the thresholds defined using
criteria for CIND identified approximately 45% of the sample as being at-risk, with dementia
progression of 9%, roughly twice the incidence in the general population. Therefore, when
used to predict at-risk cases, MCI defined risk thresholds are either very inclusive with low
progression rates, or very strict with moderate progression rates. Furthermore, from the
MMSE curve (Figure 2) diagnostic accuracy of each MCI defined threshold was not better
than using the MMSE.

**Identifying a Low Risk Cohort** Each classification defines an at-risk group, and so implicitly
all other cases are defined as being not-at-risk. Figure 3 compares the progression rate in the
not-at-risk group by its size for each risk threshold across the different MCI definitions.
Generally the smaller the not-at-risk group identified, the lower the risk within that group.
Overall, the ARCD defined thresholds identified a large not-at-risk group (approximately
80% of individuals) with a 0.5% chance (dementia progression range 0.0-2.5%) of two year
incident dementia.

**Discussion**
In this study we investigated a possible strategy to maximise identification of individuals at
risk of dementia in a population based cohort. When applied in population based samples
MCI criteria necessarily create three groups including those meeting the criteria and being
defined as MCI, those defined as not impaired (NCI), and those identified as impaired but
without meeting the criteria. Rather than a risk dichotomy based on a positive vs. negative
MCI case diagnosis, we used previously derived incident dementia rates from NCI, MCI and NON-C groups as the basis for three levels of risk stratification and then tested, for previously published MCI definitions, two different risk classification thresholds. These were compared with the discriminatory ability of the MMSE in the same cohort. We found that different definitions of MCI identified subgroups that varied with respect to their dementia risk over a two year period. Overall at-risk case selection was generally poor. In contrast, all MCI criteria accurately defined a not-at-risk threshold that captured the majority of individuals who are unlikely to develop two-year incident dementia.

With regard to at-risk case selection, for most definitions, Threshold 1 where the at-risk group was defined as the combination of the MCI and NON-C groups was better a discriminating individuals at-risk of dementia, with the exception of the definition of N-MCI. For N-MCI discrimination of at-risk from not-at risk cases was better with Threshold 2 where the not-at-risk group was defined as the combination of: (1) people who performed within normal limits on all Mayo Clinic Criteria; and, (2) people who satisfy Mayo Clinical Criteria for N-MCI (including subjective memory complaint, normal general cognitive function, no severe functional impairment, normal memory performance and impaired non-memory performance). Here, the at-risk group was defined as all people who fell outside the N-MCI range (i.e., people with single-amnestic or combined amnestic and non-amnestic impairments). As such, these results support previous findings that suggest that impaired performance in non-memory domains is not as effective at identifying individuals who progress to dementia if memory is not affected(5, 34).

While risk thresholds derived from some MCI definitions identified a small proportion of those at-risk of dementia within two years with a high prevalence of undetected cases (e.g., A-MCI), others were able to identify most people who developed incident dementia through inclusion of a very large proportion of the older population (e.g., CIND). Overall, thresholds
derived from the component criteria of ARCD had the highest sensitivity and specificity. For the definition of ARCD, the moderate risk group was defined as those with age-related cognitive decline, while the high risk group comprised those whose cognitive changes were not simply age related. Both the low vs. moderate/high and low/moderate vs. high thresholds resulted in similar classification accuracy owing to few individuals being identified as at moderate risk (i.e., people who satisfy criteria for ARCD including impaired general cognitive function with normal memory test performance). The highest predictive accuracy was achieved using the low vs. moderate/high risk threshold, with the at risk group in this case including individuals with one of three possible profiles: (1) Normal MMSE/Impaired Memory; (2) Impaired MMSE/Normal Memory (individuals with an ARCD case diagnosis); and, (3) Impaired MMSE/Impaired Memory. Here the not-at risk group included all people with non-impaired MMSE and normal memory test performance. Given that ARCD criteria overlap those of Mayo Clinic defined MCI and CIND the results suggest that additional information used to classify individuals as MCI or CIND, for example, on subjective memory complaint and non-memory test performance does not improve predictive accuracy.

Compared to more detailed MCI definitions the MMSE was generally found to have a higher combined sensitivity/specificity, with the exception of the thresholds derived from the definition of ARCD. MMSE is easy to administer in non-clinical samples and does not require more detailed MCI criteria. However, overall the best threshold was that derived from criteria of ARCD which includes further objective memory testing in addition to the MMSE. Whether the improvement in accuracy when using the MMSE in combination with objective memory testing is large enough to be justifiable, especially in terms of extra data collection costs will depend on the reason for screening as well as the availability of resources.

Using stricter cut-points for cognitive severity level (-1.5SD vs. -1SDs) had the effect of substantially reducing the sensitivity of definitions without sufficient corresponding
improvement in specificity. This is not unexpected in population based samples where the level of ability would be expected to be more variable and generally higher than in clinic based samples. Overall the results suggest that for population dementia risk screening too strict clinical criteria fail to capture all individuals at risk.

With regard to not-at-risk case selection the results suggest that the discriminatory power of the MMSE as well as several different definitions of MCI could be sufficient for identifying a large group of the older population as being at very low risk of dementia incidence. The ARCD derived thresholds were able to identify a sub-group of around 80% of individuals with normal age associated change, alive at two years follow-up and at 0.5% risk of dementia. The remaining 20% had roughly 15% risk of two year dementia incidence and can be considered to be at greater risk. Classification of not-at-risk cases using MCI criteria therefore provides a potentially useful tool to streamline future research into better methods for the remaining groups in who staged methods or timed re-screening may be more appropriate.

There are some limitations to the study. The cognitive assessment was performed with the MMSE and CAMCOG with specific cut-off scores. Different measures and impairment severity levels will lead to varying results. However, there are currently no recommendations regarding which psychometric instrument(s) or severity level maximise screening accuracy and this remains to be tested. In this study, MCI was mapped retrospectively using the application of standardised rules and measures and therefore does not benefit from the flexibility of clinical judgement. However, clinical judgement adds a subjective element to mapping which is not replicable and consistent rules allow for direct comparison across every MCI definition within the same sample. Dementia progression was evaluated within a two year time frame and during longer follow-up more persons would be expected to progress to dementia. However, screening would be expected to be effective over a limited time frame. Shorter time frames of risk would be beneficial for targeted treatment, while longer time
frames of risk might be important to identify those individuals in whom it might be important to defer treatment and also who may have a longer window of opportunity for modification. Medical exclusion criteria were applied in order to be consistent with clinical mapping of MCI(1, 35-39). While currently there are no guidelines on what are the best eligibility criteria to identify a representative MCI sample the exclusion of co-morbid conditions could adversely affect high risk case selection and population representativeness of the sample. However, previous studies have been inconclusive in identifying medical risk factors for incident dementia in MCI(40-41). Further studies examining the two-year risk of dementia across cognitive states in individuals with and without health co-morbidity are therefore a research priority. Lastly, MCI criteria perform best at identifying at-risk cases when individuals with MCI are combined with persons excluded from an MCI cases diagnosis (NON-C cases). However, no objective criteria for identifying the NON-C group exist and risk in this group has not been externally validated. Given the high dementia progression rate in the NON-C group and the finding that at-risk case identification is generally improved when MCI and NON-C groups are combined, the characteristics of the NON-C group, and how best to define MCI so at-risk cases are not excluded from diagnosis must be explored.

Conclusion

If screening for MCI is to be undertaken in the population such criteria would be expected to classify people accurately into one of three groups as defined in this analysis: (1) those at high risk for referral and management/treatment, without criteria being over or under inclusive; (2) those at moderate risk in whom deficits are suspect and where a period of watchful waiting, timed re-screening or active monitoring is more clinically appropriate; and, (3) those not at increased risk (either high functioning or with non-pathological age associated decline) who can be excluded. Criteria that are too strict will cause many cases that would benefit from early diagnosis and further follow-up to be missed, while criteria that are too inclusive will
result in many cases inappropriately identified for intervention (such as enrolment into clinical trials). Currently, clinic based MCI criteria do not perform at the very high levels required for their application as screening tools for those at high risk of incident dementia in the general older population justifying a referral, but they are able to discriminate those at elevated risk in whom watchful waiting or a cheap low-risk intervention would be appropriate from those at very low risk. Different methods for at-risk case identification must be developed which may include the unique characteristics in each definition that clearly distinguish normal age associated change, non progressive and progressive MCI.

Identification and exclusion of not-at-risk groups may present a possible targeted strategy. The number of persons classified as being not-at-risk is substantial, and by using an exclusion approach many individuals in whom preventative treatment would not be beneficial can be correctly excluded from clinical trials. Importantly we have shown that this is possible from a simple screening test such as the MMSE and more complex MCI criteria are not required.
References


impairment among the elderly over a 2- to 3-year follow-up period.


10.1097/JGP.0b013e31819431d5.


26. Bond J, Carstairs V. Services for the elderly: a survey of the characteristics and needs of a population of 5,000,000 old people. Scottish Home and Health Studies No 42. Edinburgh: Scottish Home and Health Department 1982.


List of Tables

T1. Population prevalence and two year dementia progression rate in the NCI, MCI and NON-C groups for each definition by cognitive severity level

T2. Prevalence estimates and results from the ROC analysis by MCI definition and risk threshold

SUPPLEMENTARY APPENDIX A
Table Baseline characteristics of the different groups identified at each threshold across the different MCI definitions
List of Figures

Figure 1 ROC plot comparing the predictive probability of each classification if considered as a diagnostic test for detecting those who have dementia by two years compared with MMSE scores alone.

Figure 2 Performance of each classification when identifying those people at high risk of dementia at two years.

Figure 3 Performance of each classification when identifying a population at very low risk of dementia at two years.
# T1. Population prevalence and two year dementia progression rate in the NCI, MCI and NON-C groups for each definition by cognitive severity level

<table>
<thead>
<tr>
<th>Severity 1SD</th>
<th>Population Prevalence Estimates</th>
<th>Two Year Dementia Progression Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NCI</td>
<td>MCI</td>
</tr>
<tr>
<td></td>
<td>% [n] 95%CI</td>
<td>% [n] 95%CI</td>
</tr>
<tr>
<td>MMSE</td>
<td>34.8 [442] (32.1-37.7)</td>
<td>23.8 [552] (21.6-26.3)</td>
</tr>
<tr>
<td>ARCD</td>
<td>48.8 [737] (46.0-51.6)</td>
<td>1.2 [40] (0.8-1.8)</td>
</tr>
<tr>
<td>CIND</td>
<td>36.8 [506] (34.1-39.7)</td>
<td>11.2 [310] (9.6-13.0)</td>
</tr>
<tr>
<td>N-MCI</td>
<td>34.5 [474] (31.7-37.3)</td>
<td>5.6 [110] (4.4-7.0)</td>
</tr>
<tr>
<td>A-MCI</td>
<td>34.5 [474] (31.7-37.3)</td>
<td>2.0 [41] (1.4-3.0)</td>
</tr>
<tr>
<td>M-MCI</td>
<td>34.5 [474] (31.7-37.3)</td>
<td>2.5 [71] (1.8-3.5)</td>
</tr>
<tr>
<td>R-MCI</td>
<td>34.5 [474] (31.7-37.3)</td>
<td>10.1 [222] (8.5-11.9)</td>
</tr>
<tr>
<td>Severity 1.5SD</td>
<td>ARCD</td>
<td>CIND</td>
</tr>
<tr>
<td></td>
<td>54.5 [870] (51.8-57.2)</td>
<td>1.9 [76] (1.4-2.6)</td>
</tr>
<tr>
<td></td>
<td>51.4 [806] (48.7-54.2)</td>
<td>5.2 [178] (4.3-6.4)</td>
</tr>
<tr>
<td></td>
<td>49.7 [746] (46.9-52.5)</td>
<td>1.9 [56] (1.3-2.7)</td>
</tr>
<tr>
<td></td>
<td>49.7 [746] (46.9-52.5)</td>
<td>1.0 [25] (0.6-1.7)</td>
</tr>
<tr>
<td></td>
<td>49.7 [746] (46.9-52.5)</td>
<td>0.5 [20] (0.3-0.9)</td>
</tr>
<tr>
<td></td>
<td>49.7 [746] (46.9-52.5)</td>
<td>3.3 [101] (2.5-4.3)</td>
</tr>
</tbody>
</table>

**Note:** The shading indicates the group with the highest two year dementia progression rate for the given MCI definition. It is this group which is classified as being at high risk of dementia.
T2. Prevalence estimates and results from the ROC analysis by MCI definition and risk threshold

<table>
<thead>
<tr>
<th>Classification</th>
<th>At-Risk Prevalence Estimates</th>
<th>ROC Analysis Results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%* (95% CI)</td>
<td>Sn (95% CI)</td>
</tr>
<tr>
<td>Severity 1SD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A-MCI [T2]</td>
<td>3.3 (2.2-4.9)</td>
<td>10.8 (2.8-33.5)</td>
</tr>
<tr>
<td>M-MCI [T2]</td>
<td>4.1 (2.9-5.6)</td>
<td>25.5 (10.6-49.5)</td>
</tr>
<tr>
<td>MMSE [T2]</td>
<td>5.7 (4.7-6.9)</td>
<td>24.5 (10.3-47.9)</td>
</tr>
<tr>
<td>CIND [T2]</td>
<td>17.8 (15.4-20.6)</td>
<td>47.4 (27.2-68.5)</td>
</tr>
<tr>
<td>ARCD [T2]</td>
<td>19.4 (16.9-22.2)</td>
<td>86.2 (62.9-95.8)</td>
</tr>
<tr>
<td>ARCD [T1]</td>
<td>21.3 (18.7-24.2)</td>
<td>89.4 (66.7-97.2)</td>
</tr>
<tr>
<td>R-MCI [T2]</td>
<td>28.1 (25.1-31.4)</td>
<td>53.2 (31.0-74.3)</td>
</tr>
<tr>
<td>N-MCI [T2]</td>
<td>35.5 (32.1-39.0)</td>
<td>89.5 (66.8-97.3)</td>
</tr>
<tr>
<td>CIND [T1]</td>
<td>41.4 (37.9-45.0)</td>
<td>83.8 (61.6-94.3)</td>
</tr>
<tr>
<td>MMSE [T1]</td>
<td>44.0 (40.4-47.7)</td>
<td>96.5 (76.8-99.6)</td>
</tr>
<tr>
<td>MAYO [T1]</td>
<td>44.4 (40.8-48.1)</td>
<td>89.9 (67.3-97.5)</td>
</tr>
<tr>
<td>Severity 1.5SDs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A-MCI [T2]</td>
<td>1.6 (0.9-2.7)</td>
<td>10.7 (2.8-33.4)</td>
</tr>
<tr>
<td>M-MCI [T2]</td>
<td>0.8 (0.4-1.4)</td>
<td>18.6 (6.6-42.3)</td>
</tr>
<tr>
<td>CIND [T2]</td>
<td>8.4 (6.8-10.2)</td>
<td>43.4 (23.9-65.1)</td>
</tr>
<tr>
<td>ARCD [T2]</td>
<td>9.1 (7.5-11.0)</td>
<td>43.0 (22.8-68.5)</td>
</tr>
<tr>
<td>ARCD [T1]</td>
<td>12.1 (10.3-14.3)</td>
<td>50.1 (28.4-71.8)</td>
</tr>
<tr>
<td>R-MCI [T2]</td>
<td>5.3 (4.0-7.0)</td>
<td>51.1 (14.3-55.1)</td>
</tr>
<tr>
<td>N-MCI [T2]</td>
<td>16.8 (14.6-19.4)</td>
<td>75.9 (51.9-90.2)</td>
</tr>
<tr>
<td>CIND [T1]</td>
<td>18.1 (15.8-20.7)</td>
<td>73.3 (50.6-88.1)</td>
</tr>
<tr>
<td>MAYO [T1]</td>
<td>19.8 (17.3-22.5)</td>
<td>77.8 (53.8-91.3)</td>
</tr>
</tbody>
</table>

* Percent of the population identified as at-risk by the given threshold (Not at-risk sample is calculated as 100-% defined as being at-risk)

Notes:
1. As the LR group was the same across each Mayo Clinic definition of MCI this result is labelled MAYO
2. The table is ordered by the proportion of the population defined as at-risk by each classification: definitions that identify a very specific subpopulation are at the top and those aiming to capture all individuals at risk of dementia are at the bottom

Key: T1 LR vs. (MR+HR); T2 (LR+MR) vs. HR; Sn Sensitivity; Sp Specificity; PPV Positive Predictive Value; NPV Negative Predictive Value
### Table
Baseline characteristics of the different groups identified at each threshold across the different MCI definitions

<table>
<thead>
<tr>
<th>Severity</th>
<th>ISD</th>
<th>Not at-Risk</th>
<th>At-Risk</th>
<th>Not at-Risk</th>
<th>At-Risk</th>
<th>Not at-Risk</th>
<th>At-Risk</th>
<th>Not at-Risk</th>
<th>At-Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-MCI</td>
<td>T1</td>
<td>72.8 (68-77)</td>
<td>73.2 (68-77)</td>
<td>203:271</td>
<td>253:425</td>
<td>10.7 (9-12)</td>
<td>9.6 (9-10)</td>
<td>27.5 (26-29)</td>
<td>25.7 (24-28)</td>
</tr>
<tr>
<td>A-MCI</td>
<td>T2</td>
<td>72.9 (68-77)</td>
<td>75.3 (68-80)</td>
<td>439:672</td>
<td>17:24</td>
<td>10.2 (9-11)</td>
<td>10.2 (9-11)</td>
<td>26.7 (25-29)</td>
<td>26.4 (25-28)</td>
</tr>
<tr>
<td>M-MCI</td>
<td>T1</td>
<td>72.8 (68-77)</td>
<td>73.2 (68-77)</td>
<td>203:271</td>
<td>253:425</td>
<td>10.7 (9-12)</td>
<td>9.6 (9-10)</td>
<td>27.5 (26-29)</td>
<td>25.7 (24-28)</td>
</tr>
<tr>
<td>M-MCI</td>
<td>T2</td>
<td>73.0 (69-77)</td>
<td>72.6 (69-77)</td>
<td>430:651</td>
<td>26:45</td>
<td>10.3 (9-11)</td>
<td>9.0 (9-9)</td>
<td>26.8 (25-29)</td>
<td>25.2 (23-27)</td>
</tr>
<tr>
<td>N-MCI</td>
<td>T1</td>
<td>72.8 (68-77)</td>
<td>73.2 (68-77)</td>
<td>203:271</td>
<td>253:425</td>
<td>10.7 (9-12)</td>
<td>9.6 (9-10)</td>
<td>27.5 (26-29)</td>
<td>25.7 (24-28)</td>
</tr>
<tr>
<td>N-MCI</td>
<td>T2</td>
<td>72.9 (68-77)</td>
<td>73.1 (68-77)</td>
<td>250:334</td>
<td>206:36</td>
<td>10.6 (9-11)</td>
<td>9.6 (9-10)</td>
<td>27.3 (26-29)</td>
<td>25.6 (24-28)</td>
</tr>
<tr>
<td>R-MCI</td>
<td>T1</td>
<td>72.8 (68-77)</td>
<td>73.2 (68-77)</td>
<td>203:271</td>
<td>253:425</td>
<td>10.7 (9-12)</td>
<td>9.6 (9-10)</td>
<td>27.5 (26-29)</td>
<td>25.7 (24-28)</td>
</tr>
<tr>
<td>R-MCI</td>
<td>T2</td>
<td>73.0 (68-77)</td>
<td>72.9 (68-77)</td>
<td>293:403</td>
<td>163:29</td>
<td>10.5 (9-11)</td>
<td>9.6 (9-10)</td>
<td>27.2 (26-29)</td>
<td>25.3 (23-27)</td>
</tr>
<tr>
<td>CIND</td>
<td>T1</td>
<td>72.4 (68-76)</td>
<td>76.4 (71-81)</td>
<td>353:535</td>
<td>118:192</td>
<td>10.5 (9-11)</td>
<td>9.9 (9-10)</td>
<td>27.6 (27-29)</td>
<td>25.4 (24-28)</td>
</tr>
<tr>
<td>CIND</td>
<td>T2</td>
<td>72.5 (68-76)</td>
<td>76.4 (71-81)</td>
<td>341:465</td>
<td>130:262</td>
<td>10.3 (9-11)</td>
<td>9.6 (9-10)</td>
<td>27.2 (27-29)</td>
<td>25.5 (24-28)</td>
</tr>
<tr>
<td>ARCD</td>
<td>T1</td>
<td>72.4 (68-76)</td>
<td>76.9 (71-82)</td>
<td>367:503</td>
<td>89:193</td>
<td>10.4 (9-11)</td>
<td>9.7 (9-10)</td>
<td>27.1 (26-29)</td>
<td>23.7 (22-26)</td>
</tr>
<tr>
<td>ARCD</td>
<td>T2</td>
<td>72.6 (68-76)</td>
<td>76.9 (71-81)</td>
<td>393:553</td>
<td>63:143</td>
<td>10.3 (9-11)</td>
<td>9.9 (9-10)</td>
<td>27.0 (26-29)</td>
<td>23.9 (22-27)</td>
</tr>
<tr>
<td>MMSE</td>
<td>T1</td>
<td>72.1 (68-76)</td>
<td>74.2 (70-79)</td>
<td>197:245</td>
<td>263:466</td>
<td>10.6 (9-11)</td>
<td>9.7 (9-10)</td>
<td>27.7 (27-29)</td>
<td>25.5 (24-28)</td>
</tr>
<tr>
<td>MMSE</td>
<td>T2</td>
<td>72.8 (68-76)</td>
<td>77.3 (71-83)</td>
<td>402:592</td>
<td>58:119</td>
<td>10.3 (9-11)</td>
<td>8.9 (9-9)</td>
<td>27.0 (26-29)</td>
<td>21.8 (20-24)</td>
</tr>
</tbody>
</table>

**Key:** T1 LR vs, (MR+HR); T2 (LR+MR) vs. HR; M:F Male:Female; p25 25th Percentile; p75 75th Percentile
Figure 1 ROC plot comparing the predictive probability of each classification if considered as a diagnostic test for detecting those who have dementia by two years compared with MMSE scores alone.
Figure 2 Performance of each classification when identifying those people at high risk of dementia at two years
Figure 3 Performance of each classification when identifying a population at very low risk of dementia at two years