PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Diagnosing Mild Cognitive Impairment (MCI) in Clinical Trials: A Systematic Review
AUTHORS	Stephan, Blossom; Minett, Thais; Pagett, Emma; Siervo, Mario; Brayne, Carol; McKeith, Ian

VERSION 1 - REVIEW

REVIEWER	Petersen, Ronald C Mayo Clinic, Neurology no competing interests
REVIEW RETURNED	10-Sep-2012

RESULTS & CONCLUSIONS	This manuscript reviews the implementation of criteria for mild cognitive impairment (MCI) in clinical trials. The authors reviewed the extant literature and applied criteria for the application of the Petersen criteria for MCI and concluded that the criteria were inconsistently applied across numerous studies. They found that there was great variability in the implementations of the five criteria for defining MCI and indicated that there is a strong need for standardization of the criteria. As such, this is a relevant review done by scholars in the field. The authors are known experts in the area of cognition and study design and have applied a rigorous technique to the topic. The conclusions are quite appropriate and are important. The review, however, tends to imply that this problem is unique to the construct of MCI. In particular, they have taken the diagnosis of amnestic MCI (aMCI) and indicated that it is inconsistently applied across studies. While this is absolutely true, the same argument could be made for any other entity that is defined clinically. Therefore, if one were to do a review of the literature for dementia, Alzheimer's disease, frontotemporal lobar dementia, vascular cognitive impairment, dementia with Lewy bodies or most other cognitive disorders, the conclusions would be exactly the same. Even for a commonly studied entity such as the dementia of Alzheimer's disease, there is no standard application of instruments or methodology. While certain scales have been adopted for many Alzheimer's disease clinical trials, there is no standardization and wide variability. As such, again, while the conclusions are accurate, the authors need to acknowledge that this is not peculiar to MCI. In addition, they note in the discussion that there have been recent efforts at standardization of the criteria, especially using the application of biomarkers. This will no doubt stratify subjects with an underlying Alzheimer's disease pathophysiology. This will likely help in standarizing the characterization of subjects who are on the road to Alzheimer's disease. In addition DSM 5 is moving in that direction.

	The authors conclude, however, that aMCI is an unstable construct, but this must be interpreted in a specific context. That is, they imply that aMCI is due to a degenerative disease, likely Alzheimer's disease, in most contexts. While this is the most frequently studied subtype of MCI, the construct itself does not necessarily mean that subjects with MCI will progress. In fact, several review papers on MCI indicate that the construct, broadly defined, could be of variable etiologies, and as such, its progression or lack thereof is determined by the underlying etiology. One might expect some MCI subjects to return to normal and this does not mean that the construct is unstable. For example, if a person had aMCI due to depression, and the depression were treated successfully, the aMCI may resolve as one would expect. This does not indicate that the construct is unstable; rather, this is the expected outcome. The authors need to be clear that, when they are talking about aMCI of presumed Alzheimer's disease etiology, progression is expected. Most studies evaluating aMCI due to presumed Alzheimer's disease do note a progression over time. It is also noteworthy to acknowledge that, when the construct of aMCI is defined using specific criteria, e.g., the Mini-Mental State Exam, Recall on the Logical Memory Subtest of the Wechsler Memory Scale-Revised and the Clinical Dementia Rating, the results can be quite uniform. For example, the RCT performed by the Alzheimer's Disease Cooperative Study involving vitamin E and donepezil used these instruments to characterize subjects with aMCI. When those exact implementation criteria were adopted by the Alzheimer's Disease Neuroimaging Initiative, virtually identical results with respect to progression (approximately 16% per year) to dementia were found. This indicates that the criteria can be implemented in multi-center settings and produce consistent results. The reversion rate in these studies was extremely low. In summary, this is a scholarly and useful review of the literature. It does have a negative tone toward the construct of MCI, and while the conclusions are accurate, i.e., recommending standardization of instruments used to define the criteria, they are not unique to MCI, and this should be acknowledged. The recent definition of MCI due to Alzheimer's disease and DSM 5 will likely standardize the situation. The same inconsistencies of application of criteria can be found with prodromal Alzheimer's disease, as has been proposed. The criteria for prodromal Alzheimer's disease do recommend a single memory test, the Free and Cued Selective Reminding Test, but this has not been empirically validated. In the several papers regarding prodromal Alzheimer's disease, the subjects have had MCI at the outset, and the Free and Cued Selective Reminding Test was applied to that subpopulation. When the Free and Cued Selective Reminding Test alone was applied to a non-demented, population-based sample of subjects, as in the 3 City Study from Bordeaux, the instrument did not perform well. As such, the criteria proposed for prodromal Alzheimer's disease are similarly liable to variable applications of the criteria. Therefore, it must be recognized that prodromal Alzheimer' disease begins with MCI as the clinical construct and as such is no more precise than MCI itself. All this is not to say that there should not be an effort toward standardization of the criteria, but it should be acknowledged that this problem is not unique to MCI but more broadly applies to any clinical diagnosis.

REVIEWER	Katie Palmer Researcher, Department of Psychology, Stockholm University, Sweden. No competing interests.
REVIEW RETURNED	17-Sep-2012

GENERAL COMMENTS	This is a systematic review investigating how the criteria for MCI have been applied across RCTs. The authors focus on one definition of MCI (Petersen et al's 1999 criteria), and report findings from 22 RCTs using participants defined as MCI according to this criteria. They report that, even though all studies quote the same referenced definition, the operationalization of the criteria vary greatly. This is an excellent review, which focuses on an important point that is often discussed in articles and conference on this subject. This review describes in the detail the differences in operationalizing the criteria, and helps to quantify the extent of the differences. It systematically addresses the issue in a thoughtful way. Overall, this is becoming an increasingly growing problem. If researchers and pharmaceutical companies are trying to target MCI as a therapeutic target population, then needs to be more consistency in how this 'syndrome' is diagnosed. Overall, this is a well-conducted review. I have only a few suggestions for minor revisions: 1. I think it will be informative to add a column in Supplementary Table 1 stating the main outcome measure of the trial (dementia progression, improvement in cognitive functioning, delay of institutionalization? etc). This would add another interesting dimension to the discussion. 2. The Discussion is somewhat weighted on the issue of MCI representing 'prodromal dementia', but it should be recognized and discussed further that a large proportion of the studies did not have dementia progression as a primary outcome. Does the aim of the RTC affect what type of criteria are used? 3. The review includes only RCTS that have stated that they use the Petersen 1999 MCI criteria. I think it was a good methodological choice to focus just on one definition, as it would be impossible otherwise to provide a coherent analysis of the differences between studies. As the authors discuss, many trials also use other definitions of prodromal AD and other MCI-similar concepts and syndromes. So this review highlights only the tip of the problem. The real problem of inconsistency will greatly exceed this once you take into account also the other numerous MCI-similar concepts that are used in RCTS. The authors have somewhat discussed this, but it would strengthen their report and conclusions to discuss this further. 4. It could be interesting to add a short discussion on how this compares with the situation for another disorder such as dementia. i.e. are there any systematic reviews comparing differences in the operationalization of dementia criteria in AD, and if so, are the results similar? 5. 'Supplementary Table 2' has been labeled as 'Supplementary

	Table 1'. 6. The characteristics of the included studies are listed in Supplementary Table 1. There is a huge wealth of information here, and at first it was difficult to follow the 'first 9 columns / last 8 columns' format of the table. However, I think all this information is informative to the reader, and should be kept in the table, but I would suggest simply splitting it into 2 separate tables: 1) General characteristics of the included studies and 2) Application of MCI criteria used in the included studies. 7. Would be good to add some footnotes for the abbreviations in Supplementary Table 1'. (e.g. what is NS? Not stated? If so how is this different from 'Unknown'.

VERSION 1 – AUTHOR RESPONSE

Reviewer: Ronald C Petersen

The review, however, tends to imply that this problem is unique to the construct of MCI. In particular, they have taken the diagnosis of amnestic MCI (aMCI) and indicated that it is inconsistently applied across studies. While this is absolutely true, the same argument could be made for any other entity that is defined clinically. Therefore, if one were to do a review of the literature for dementia, Alzheimer's disease, frontotemporal lobar dementia, vascular cognitive impairment, dementia with Lewy bodies or most other cognitive disorders, the conclusions would be exactly the same. Even for a commonly studied entity such as the dementia of Alzheimer's disease, there is no standard application of instruments or methodology. While certain scales have been adopted for many Alzheimer's disease clinical trials, there is no standardization and wide variability. As such, again, while the conclusions are accurate, the authors need to acknowledge that this is not peculiar to MCI. In addition, they note in the discussion that there have been recent efforts at standardization of the criteria, especially using the application of biomarkers. This will no doubt stratify subjects with an underlying Alzheimer's disease pathophysiology. This will likely help in standarizing the characterization of subjects who are on the road to Alzheimer's disease. In addition DSM 5 is moving in that direction.

As highlighted by the reviewer lack of uniformity in clinical diagnosis is not only a problem within the field of MCI, but is also a problem within the field of dementia generally. We have included a discussion of this throughout the manuscript:

Abstract: Conclusion (page 3) TEXT added "Lack of uniformity in clinical diagnosis however is not exclusively a problem for MCI but also for other clinical states such as dementia including Alzheimer's disease and vascular dementia. Defining a uniform approach to MCI classification, or indeed for any classification concept within the field of dementia, should be a priority if further trials are to be undertaken in the older aged population based on these concepts."
Key Messages: Point 3 (page 4) TEXT added "Lack of specific methods for clinical diagnosis is not a problem unique to the field of MCI. Across studies there continues to be inconsistency in the instruments and methodology used to diagnose Alzheimer's disease and Vascular Dementia, including its prodromal stage, Vascular Cognitive Impairment no Dementia (VCIND). Revision of diagnostic criteria including standardisation of methods and instruments for operationalisation of each dementia subtype and for the different disease stages (e.g., prodromal, preclinical and clinical) should be a research priority."
Discussion: Paragraph 2 (page 14/15) TEXT added ". Indeed, within the field of dementia there is a lack of consistency in operationalisation protocols not only for aMCI, but its associated disorders (e.g., Cognitive Impairment no Dementia53), dementia and its sub-types (such as Alzheimer's Disease and

vascular dementia), pre-MCI54 and other preclinical states such as VCIND55. Different diagnostic criteria for MCI affect prevalence56 and progression57. Similarly for dementia different criteria have been found to affect prevalence58, 59. Inconsistency in case classification can have impactions for research and trial recruitment and outcomes."

The authors conclude, however, that aMCI is an unstable construct, but this must be interpreted in a specific context. That is, they imply that aMCI is due to a degenerative disease, likely Alzheimer's disease, in most contexts. While this is the most frequently studied subtype of MCI, the construct itself does not necessarily mean that subjects with MCI will progress. In fact, several review papers on MCI indicate that the construct, broadly defined, could be of variable etiologies, and as such, its progression or lack thereof is determined by the underlying etiology. One might expect some MCI subjects to return to normal and this does not mean that the construct is unstable. For example, if a person had aMCI due to depression, and the depression was treated successfully, the aMCI may resolve as one would expect. This does not indicate that the construct is unstable; rather, this is the expected outcome. The authors need to be clear that, when they are talking about aMCI of presumed Alzheimer's disease etiology, progression is expected. Most studies evaluating aMCI due to presumed Alzheimer's disease do note a progression over time.

We have updated the discussion on whether aMCI is a progressive vs. non-progressive condition. We have highlight that MCI can have different underlying causes and that this can affect whether the condition reflects a high-risk dementia state.

TEXT added (page 16) "Indeed, symptoms of MCI are not always a consequence of Alzheimer's pathology, but rather can have multiple aetiologies such as depression or vascular disease each with different outcomes (e.g., dementia progression, improvement with treatment for the underlying health symptoms)62, 63. Better methods are needed to determine the underlying cause of disease in this patient group to accurately identify those individuals whose MCI is associated with Alzheimer's Disease."

It is also noteworthy to acknowledge that, when the construct of aMCI is defined using specific criteria, e.g., the Mini-Mental State Exam, Recall on the Logical Memory Subtest of the Wechsler Memory Scale-Revised and the Clinical Dementia Rating, the results can be quite uniform. For example, the RCT performed by the Alzheimer's Disease Cooperative Study involving vitamin E and donepezil used these instruments to characterize subjects with aMCI. When those exact implementation criteria were adopted by the Alzheimer's Disease Neuroimaging Initiative, virtually identical results with respect to progression (approximately 16% per year) to dementia were found. This indicates that the criteria can be implemented in multi-center settings and produce consistent results. The reversion rate in these studies was extremely low.

We have highlighted the potential use of the Alzheimer's Disease Cooperative Study methodology to map MCI, particularly in light of the fact that when using this operationalisation strategy dementia progression has been found to be consistent across studies including the multicentre ADNI study. We also highlight the need for further research (across cohorts [population-based vs. clinical] and countries).

TEXT ADDED (Page 16/17) "Better methods are needed to determine the underlying cause of disease in this patient group to accurately identify those individuals whose MCI is associated with Alzheimer's Disease. One possibility could be defining aMCI as in the Alzheimer's Disease Cooperative Study trial9 (based on a subjective memory complaint, MMSE score, impaired performance on the Logical Memory II Subscale, no functional impairment and a CDR score of 0.5) as implementation of this methodology has been found to result in a consistent rate of dementia progression (approximately 16%/year) across studies, including the multicentre Alzheimer's Disease

Neuroimaging Initiative64. Further research is needed to test this method of operationalisation across cohorts (clinical and population based; across countries) and calculate prevalence and longitudinal course in order to determine generalisability of these findings. ”

In summary, this is a scholarly and useful review of the literature. It does have a negative tone toward the construct of MCI, and while the conclusions are accurate, i.e., recommending standardization of instruments used to define the criteria, they are not unique to MCI, and this should be acknowledged. The recent definition of MCI due to Alzheimer's disease and DSM 5 will likely standardize the situation. The same inconsistencies of application of criteria can be found with prodromal Alzheimer's disease, as has been proposed. The criteria for prodromal Alzheimer's disease do recommend a single memory test, the Free and Cued Selective Reminding Test, but this has not been empirically validated. In the several papers regarding prodromal Alzheimer's disease, the subjects have had MCI at the outset, and the Free and Cued Selective Reminding Test was applied to that subpopulation. When the Free and Cued Selective Reminding Test alone was applied to a non-demented, population-based sample of subjects, as in the 3 City Study from Bordeaux, the instrument did not perform well. As such, the criteria proposed for prodromal Alzheimer's disease are similarly liable to variable applications of the criteria. Therefore, it must be recognized that prodromal Alzheimer' disease begins with MCI as the clinical construct and as such is no more precise than MCI itself. All this is not to say that there should not be an effort toward standardization of the criteria, but it should be acknowledged that this problem is not unique to MCI but more broadly applies to any clinical diagnosis.

We have edited the section discussing the new “lexicon” of AD. As suggested we highlight that any new definition of prodromal disease or pre-MCI, similarly to aMCI, would need standardised criteria and an operational protocol. Further, we argue that it would need to be validated across settings and in different populations (e.g., the oldest-old).

TEXT ADDED (Page 17) “For example, MCI as defined by Petersen criteria may no longer be considered at-risk, but as already AD, and encompassed in the new term “prodromal AD”; an early symptomatic stage pre-dementia where a patient shows evidence of memory impairment and positive ratings on pathophysiological and topographical markers of AD65. Clinical trial research may therefore shift some focus to asymptomatic at-risk states (e.g., pre-MCI) where individuals are biomarker positive for AD but are otherwise healthy. However, like aMCI efforts are needed to standardise criteria and develop an operational protocol for any new stage of disease (e.g., prodromal AD and pre-MCI) and undertake validation across settings including oldest-old age groups and populations (vs. clinical samples).”

Reviewer: Katie Palmer

This is a systematic review investigating how the criteria for MCI have been applied across RCTs. The authors focus on one definition of MCI (Petersen et al's 1999 criteria), and report findings from 22 RCTs using participants defined as MCI according to this criteria. They report that, even though all studies quote the same referenced definition, the operationalization of the criteria vary greatly. This is an excellent review, which focuses on an important point that is often discussed in articles and conference on this subject. This review describes in the detail the differences in operationalizing the criteria, and helps to quantify the extent of the differences. It systematically addresses the issue in a thoughtful way. Overall, this is becoming an increasingly growing problem. If researchers and pharmaceutical companies are trying to target MCI as a therapeutic target population, then needs to be more consistency in how this 'syndrome' is diagnosed. Overall, this is a well-conducted review. I have only a few suggestions for minor revisions:

1. I think it will be informative to add a column in Supplementary Table 1 stating the main outcome

measure of the trial (dementia progression, improvement in cognitive functioning, delay of institutionalization? etc). This would add another interesting dimension to the discussion.

As requested we have added an extra column to Supplementary Table 1a called “Outcomes tested” and highlight in the text that across trials the outcomes varied considerably. We chose not to focus on the outcomes of the trials in this paper as we felt that this would detract from the main focus being variability in MCI classification.

TEXT added (Page 9) “Outcomes varied extensively across studies and included assessment of cognitive function (in all studies either as a primary or secondary outcome, with no neuropsychological assessment applied consistently) in addition to non-cognitive measures (e.g., vascular health such as blood pressure, quality of life, depression, cerebrospinal fluid (CSF) biomarkers of Alzheimer’s Disease pathology and neuroimaging).”

2. The Discussion is somewhat weighted on the issue of MCI representing ‘prodromal dementia’, but it should be recognized and discussed further that a large proportion of the studies did not have dementia progression as a primary outcome. Does the aim of the RTC affect what type of criteria are used?

We chose not to focus on the outcomes tested as the aim was to look at operationalisation of aMCI criteria. We have added an additional column to Table 1a (“Outcomes Tested”) which highlights that all trials assessed cognition. Further, this column highlights that the outcomes tested were just as variable as MCI operationalisation. No two trials used the same methodology which makes it difficult to establish any relationship between the aim of the study and the type of criteria used.
3. The review includes only RCTS that have stated that they use the Petersen 1999 MCI criteria. I think it was a good methodological choice to focus just on one definition, as it would be impossible otherwise to provide a coherent analysis of the differences between studies. As the authors discuss, many trials also use other definitions of prodromal AD and other MCI-similar concepts and syndromes. So this review highlights only the tip of the problem. The real problem of inconsistency will greatly exceed this once you take into account also the other numerous MCI-similar concepts that are used in RCTS. The authors have somewhat discussed this, but it would strengthen their report and conclusions to discuss this further.

We have added a new section to the Discussion highlighting that the problem of inconsistency in diagnosis is not unique to Petersen et al defined aMCI, but also exists for other concepts that define preclinical dementia (such as Cognitive Impairment no Dementia [CIND] and pre-MCI) and also dementia and its subtypes (including Alzheimer’s Disease and vascular dementia).

Discussion: Paragraph 2 (page 14/15) TEXT added “Indeed, within the field of dementia there is a lack of consistency in operationalisation protocols not only for aMCI, but its associated disorders (e.g., Cognitive Impairment no Dementia53), dementia and its sub-types (such as Alzheimer’s Disease and vascular dementia), pre-MCI54 and other preclinical states such as VCIND55. Different diagnostic criteria for MCI affect prevalence56 and progression57. Similarly for dementia different criteria have been found to affect prevalence58, 59. Inconsistency in case classification can have impactions for research and trial recruitment and outcomes.”

4. It could be interesting to add a short discussion on how this compares with the situation for another disorder such as dementia. i.e. are there any systematic reviews comparing differences in the operationalization of dementia criteria in AD, and if so, are the results similar?

There are results that have shown how dementia prevalence can vary substantially (as large as a factor of 10) using different diagnostic criteria (References: Erkinjuntt et al., N Engl J Med, 1997;

Wancata et al., Am J Geriatr Psychiatry, 2007) and we have added this to the discussion. We did not find any systematic review that has specifically focused on classification criteria for dementia or its preclinical states in randomised controlled trials.

TEXT ADDED (Page 15) “Different diagnostic criteria for MCI affect prevalence56 and progression57. Similarly for dementia different criteria have been found to affect prevalence58, 59. Inconsistency in case classification can have impactions for research and trial recruitment and outcomes.”

5. ‘Supplementary Table 2’ has been labelled as ‘Supplementary Table 1’.

This has been corrected.

6. The characteristics of the included studies are listed in Supplementary Table 1. There is a huge wealth of information here, and at first it was difficult to follow the ‘first 9 columns / last 8 columns’ format of the table. However, I think all this information is informative to the reader, and should be kept in the table, but I would suggest simply splitting it into 2 separate tables: 1) General characteristics of the included studies and 2) Application of MCI criteria used in the included studies.

We have split the table as suggested. The first table (Table 1a) includes the general characteristics of each trial in addition to the new “Outcomes tested” column and the second (Table 1b) includes how MCI was mapped in each trial.

Reference to the tables was also updated in the text:

TEXT added (page 8/9) “Supplementary Table 1a summarises the general characteristics, demographics and outcomes tested in each included article. Supplementary Table 1b summarises the operationalisation protocol used for identifying aMCI cases in each trial.”

7. Would be good to add some footnotes for the abbreviations in Supplementary Table 1’. (e.g. what is NS? Not stated? If so how is this different from ‘Unknown’.

We have updated Supplementary Table 1 to consistently use “unknown” where information is missing. We had also added a key for the acronyms.

VERSION 2 – REVIEW

REVIEWER	Ronald C. Petersen Mayo Clinic College of Medicine Rochester, MN USA
REVIEW RETURNED	29-Nov-2012

GENERAL COMMENTS	This is a revision of the manuscript previously submitted, assessing methodologies in clinical trials for mild cognitive impairment (MCI). As such, it is significantly improved and more of a balanced presentation than the original manuscript. The authors have done a scholarly job in reviewing the literature and reporting their findings in a systematic fashion. As such, the data speak for themselves. The fundamental modification in the manuscript regards the acknowledgement that the issues raised and conclusions drawn by this exercise are not peculiar to MCI. They have mentioned in the abstract and throughout the discussion that the type of variability found in standardizing criteria for MCI could equally be applied to dementia due to AD, vascular cognitive impairment and, I would

	suggest that they also expand this to include frontotemporal lobar degeneration and dementia with Lewy bodies. In the latter two instances, they could make an additional comment about how difficult it would be to standardize the criteria for FTLD, including various types of asphasia, from nonfluent to semantic to logopenic, and also the challenges with characterizing the behavioral features of FTLD. Similarly, DLB has its own set of challenges in defining fluctuations, dream enactment behavior, etc. The point being that, while MCI is a focal point of this paper, one would not want the reader to come away with the conclusion that this is problematic for the construct of MCI alone. With respect to emphasis, the paper is actually critiquing the construct of RCTs in MCI and not the clinical construct of MCI itself. The data indicate that there is a great deal of variability in the manner in which the criteria for MCI are implemented, but really should not be misinterpreted to criticize the clinical construct of MCI itself. That is an important distinction to note in the Discussion section. Again, at the risk of being redundant, the problems in operationalizing the five criteria attributable to MCI could be replicated with any other clinical diagnosis of a cognitive disorder. As an aside, DSM-5 is being completed, and the same issues are being encountered with a variety of neurocognitive disorders. In that sense, it is not realistic to specify precise instruments and cutoff scores for any of these clinical entities at the present time. This type of statement should be acknowledged in the paper. From a terminology perspective, the authors should be explicit as to their use of some terms. For example, in the initial bullet point in the Article Summary section, they refer to “preclinical dementia.” This is an accurate statement, but one might infer that they are really talking about preclinical Alzheimer’s disease. In the second bullet, they describe “preclinical cases” using the amnestic MCI (aMCI) diagnostic criteria. Technically, that is incorrect because aMCI individuals are “clinical” not “preclinical.” I believe they mean pre-dementia rather than preclinical. With the recent National Institute on Aging – Alzheimer’s Association definitions of the preclinical, MCI and dementia states of Alzheimer’s disease, the terminology can get confusing. In addition, some colleagues have used the term prodromal Alzheimer’s disease, and that overlaps with both MCI due to AD and the aMCI discussion here. This is mentioned in the Discussion section, and it should be highlighted for the reader as to the similarities and distinctions. Essentially, prodromal Alzheimer’s disease refers to a stage of MCI due to AD. The use of biomarkers may, in fact, help sort out some of these issues. The authors have included a statement regarding the MCI trial done by the Alzheimer’s Disease Cooperative Study and the current project underway by the Alzheimer’s Disease Neuroimaging Initiative. It is noteworthy and perhaps could be emphasized more that, when these criteria are applied in a strict fashion, the performance of the criteria in these clinical trials is almost identical. In that sense, this would give the reader some sense of hope that the field is in not total disarray. In fact, there probably is becoming more consistency in the field regarding MCI than not. In summary, this version of the manuscript is significantly improved. The tone of being critical of MCI, specifically, has been softened. I think some additional elaboration on a couple of the topics mentioned above might be useful in putting the discussion in its

	appropriate context.

VERSION 2 – AUTHOR RESPONSE

Reviewer: Ronald C. Petersen
Mayo Clinic College of Medicine
Rochester, MN USA

COMMENT: The fundamental modification in the manuscript regards the acknowledgement that the issues raised and conclusions drawn by this exercise are not peculiar to MCI. They have mentioned in the abstract and throughout the discussion that the type of variability found in standardizing criteria for MCI could equally be applied to dementia due to AD, vascular cognitive impairment and, I would suggest that they also expand this to include frontotemporal lobar degeneration and dementia with Lewy bodies.

RESPONSE: We had added the examples of Lewy Body dementia and frontotemporal dementia to the Abstract (Page 3) and Discussion (Page 15).

COMMENT: In the latter two instances, they could make an additional comment about how difficult it would be to standardize the criteria for FTLD, including various types of asphasia, from nonfluent to semantic to logopenic, and also the challenges with characterizing the behavioral features of FTLD. Similarly, DLB has its own set of challenges in defining fluctuations, dream enactment behavior, etc. The point being that, while MCI is a focal point of this paper, one would not want the reader to come away with the conclusion that this is problematic for the construct of MCI alone.

RESPONSE: We have updated the discussion to highlight that for some dementias and their subtypes it may be difficult to have one operationalisation protocol and have included the examples above. We have also included the example of vascular dementia/VCIND where the type and location of pathology may result in variability in symptom profile.

Text Added (Page 15, Paragraph 1) For some dementias and their related conditions it may however be difficult and unrealistic to have one set of operational criteria, precise assessment instruments or cut-off values. For example, a single set of criteria may not be possible for defining symptom fluctuations (e.g., as seen in Lewy Body dementia), capturing variability in symptom profiles (e.g., the different type of aphasic deficit presented in frontotemporal dementia) or reflecting differences in neuropathological profiles (e.g., for vascular dementia and VCIND the type and location of vascular damage may result in variable symptom profiles).

COMMENT: With respect to emphasis, the paper is actually critiquing the construct of RCTs in MCI and not the clinical construct of MCI itself. The data indicate that there is a great deal of variability in the manner in which the criteria for MCI are implemented, but really should not be misinterpreted to criticize the clinical construct of MCI itself. That is an important distinction to note in the Discussion section.

RESPONSE: We have updated the first paragraph of the Discussion so that it makes clear that the problem discussed is the lack of consistency in how MCI is operationalised in clinical trials and stress that future work is needed to develop a consistent recruitment protocol for MCI clinical trials.

Text Added (Page 14, Paragraph 2) A priority for clinical trial research is to agree a uniform set of criteria to operationalize MCI. The recruitment protocols identified in this review could provide the basis for future work to determine best practice (e.g., in terms of testing classification accuracy of the different methods used,) in order to inform the development of a consistent recruitment methodology for MCI clinical trials.

COMMENT: Again, at the risk of being redundant, the problems in operationalizing the five criteria attributable to MCI could be replicated with any other clinical diagnosis of a cognitive disorder. As an aside, DSM-5 is being completed, and the same issues are being encountered with a variety of neurocognitive disorders. In that sense, it is not realistic to specify precise instruments and cutoff scores for any of these clinical entities at the present time. This type of statement should be acknowledged in the paper.

RESPONSE: We have edited the Discussion (Page 15, Paragraph 1) to highlight that for some dementias and their related conditions it may not be possible to have a single operationalisation protocol with specific assessment instruments and cut-off values. For aMCI agreement could be reached for example using the methods from the Alzheimer's Disease Cooperative Study and ADNI as mentioned (Page 17, Paragraph 1).

COMMENT: From a terminology perspective, the authors should be explicit as to their use of some terms. For example, in the initial bullet point in the Article Summary section, they refer to "preclinical dementia." This is an accurate statement, but one might infer that they are really talking about preclinical Alzheimer's disease. In the second bullet, they describe "preclinical cases" using the amnestic MCI (aMCI) diagnostic criteria. Technically, that is incorrect because aMCI individuals are "clinical" not "preclinical." I believe they mean pre-dementia rather than preclinical. With the recent National Institute on Aging – Alzheimer's Association definitions of the preclinical, MCI and dementia states of Alzheimer's disease, the terminology can get confusing. In addition, some colleagues have used the term prodromal Alzheimer's disease, and that overlaps with both MCI due to AD and the aMCI discussion here. This is mentioned in the Discussion section, and it should be highlighted for the reader as to the similarities and distinctions. Essentially, prodromal Alzheimer's disease refers to a stage of MCI due to AD. The use of biomarkers may, in fact, help sort out some of these issues.

RESPONSE: The terminology has been corrected in the "Article Focus", "Key Messages" and "Strengths and Limitations Sections". We now refer to MCI as "pre-dementia".

COMMENT: The authors have included a statement regarding the MCI trial done by the Alzheimer's Disease Cooperative Study and the current project underway by the Alzheimer's Disease Neuroimaging Initiative. It is noteworthy and perhaps could be emphasized more that, when these criteria are applied in a strict fashion, the performance of the criteria in these clinical trials is almost identical. In that sense, this would give the reader some sense of hope that the field is in not total disarray. In fact, there probably is becoming more consistency in the field regarding MCI than not.

RESPONSE: We have added to the discussion that if further testing across cohorts supports generalizability of the findings from the Alzheimer's Disease Co-operative Study and ADNI (e.g., in terms of consistency in prevalence and dementia progression rates) then such criteria could be recommended for use in all future aMCI clinical trials.

Text added (Page 17, Paragraph 1) Such results could have important implications in terms of identifying a standard protocol for all future aMCI clinical trials.