The status of Mild Cognitive Impairment (MCI) as a therapeutic target for regulatory approval is unclear. On the one hand, many expert physicians feel confident that people who present to memory clinics with memory problems, who have objective data, reviewed clinically, that they are neither normal nor demented can be diagnosed with MCI, and found to be at an increased risk of dementia.1,2 For these physicians, the real question - and they see it as a pressing one, of both clinical and public policy importance - is whether this increased risk of progression can be lessened. Others are less sure about each of these points. They question whether MCI is a valid entity, noting that the outcomes of MCI depend greatly on where it is detected, that the results of objective tests are inconsistent within3,4 and across studies (i.e. inconsistent in their interpretation, cut-points, correlations and outcomes).5,7 This paper considers the question of whether MCI is a well enough established entity that it can be a valid target of therapy within a regulatory context.

From the Department of Medicine (Geriatric Medicine & Neurology) (KR), Dalhousie University, Halifax, NS; Department of Medicine (Neurology) (HC), McGill University, Montreal, QC; Department of Medicine (Neurology) (HHF), University of British Columbia, Vancouver, BC, Canada.


Reprint requests to: Kenneth Rockwood, Centre for Health Care of the Elderly, 1421-5955 Veterans’ Memorial Lane, Halifax, Nova Scotia, B3H 2E1, Canada.
VALIDATION OF THE CONSTRUCT OF MILD COGNITIVE IMPAIRMENT

The validity of a construct can be evaluated in several ways, but a three-part approach distinguishes between content, construct and criterion validity.8 Content validity refers to whether an idea is valid on its face. Although at least one careful study, which employed neuropathological comparison, has called it into question,9 both broad camps (of advocates and skeptics) agree that many patients with dementia pass through a transitional stage. Still, there is considerable nuance, even in areas of apparent agreement. For example, the advocates of the MCI as high-risk state now recognize subgroups10 - an amnestic MCI type which progresses to dementia, and other types (e.g. frontotemporal MCI11) that progress to other dementias. In this way, some sense can be made of the considerable heterogeneity that is noted in people with MCI.12-16 Clearly, content validity is a weak form of validation, but can be a useful starting point. Here it illustrates a central feature of MCI as now constituted, which is its heterogeneity. It is the resolution of heterogeneity which will be important to how MCI is viewed as a target for therapy.

The persistence of heterogeneity of who is included in various MCI definitions, and what outcomes they experience is at the heart of the MCI controversy, but variable outcomes are not unique to MCI. People with MCI come from a larger pool of people who have cognitive states that, while impaired, do not meet criteria for dementia. For example, the alcohol amnestic syndrome, post-stroke cognitive impairment, congenital cognitive abnormalities, chronic schizophrenia, depression, post-infectious syndromes and traumatic brain injury were each designated as belonging to the group “Cognitive Impairment, No Dementia” (CIND). The CIND group was the single largest entity in the population, according to the Canadian Study of Health and Aging,17 and people with many CIND syndromes showed improvement over five years of follow-up.18

The setting of MCI in a wider context of CIND is key. The various sets of criteria can be read as a progressive refinement from the most heterogenous category of CIND to a purer ‘AMCI’ as confounders (alcohol amnestic syndrome, depression, etc.) are eliminated. [Table] From a construct standpoint, there are two problems – what is the nature of what is left over? Is it AMCI or very early Alzheimer’s disease? In addition, what is the nature of the conditions that have been excluded? Are they actually confounders, or are they – as might be argued about depression that presents chiefly with cognitive features – variants of an underlying process such as Alzheimer’s disease?

Much of the evidence for and against the validity of MCI comes from studies of construct validity, in which MCI is evaluated in its correlation with standard measures. A host of studies point to patients with MCI having scores on a variety of tests that are “in between” patients with no cognitive impairment and those with dementia.19-24 Some of this is inevitable from the way that MCI is defined. In consequence, it is important to understand the extent to which the memory impairment (in the case of amnestic MCI) correlates with impairment in other domains. This is important because if patients with AMCI have scores “in between” NCI and dementia in memory, but not in other cognitive domains, then it would be easy to accept that AMCI is an isolated memory problem that might herald subsequent dementia. But the common observation that AMCI patients also have “‘in between’ scores on language, attention and concentration, and function can further be taken as evidence that these patients have very mild dementia. The reasoning is that they have global cognitive impairment (memory and at least one other domain) likely with some functional impairment, which thereby would meet the definition of dementia. In short, one important objection to MCI is that it really is early Alzheimer’s disease.7 The counter-argument from clinicians hinges on two points. First, the definition of dementia stresses that deficits must be “significant”, and the “in-between” deficits do not meet this clinically crucial but conceptually ephemeral level of impairment. Secondly, in practice, clinicians must hesitate in concluding or communicating that patients have what is usually an untreatable condition (i.e., dementia), when that individual may be reclassified within three years as being normal! By its very nature MCI in most settings is heterogeneous12-14 and any MCI cohort contains individuals who will not progress to dementia16,19 and therefore it would be wrong to tell patients that MCI is really just early Alzheimer’s Disease. How to distinguish clinically between people who will progress (and thus are legitimate targets of therapy) and those who will not is not yet established.

For people who subscribe to the view that what we really need is better diagnosis of very early dementia, then how might investigations proceed? Some feel that the state could be quickly clarified by better testing - for example, by better neuropsychological batteries,19,25 and specifically, better tests of executive function26 or of neuroimaging,27 making the diagnosis less dependent on clinical judgment. Others, however, are less persuaded that more “objective” testing (either neuropsychological3 or neuroimaging28) would substantially improve diagnosis. There is also disagreement about the same data. For example, even amongst people who meet all of the MCI criteria, the time for progression to dementia varies by many years between patients. Mild Cognitive Impairment advocates see this as confirming the high risk state, making it distinct from early dementia. Mild Cognitive Impairment skeptics call it unacceptable heterogeneity, likely reflecting very early dementia. On the other hand, proponents of (A)MCI as a high risk state might argue that this is a misuse of the term dementia – akin to saying that someone with a rectal temperature of 37.6 (in a setting where infection is suspected) has very early fever. (This too can be countered by pointing out the need to know not just a single observation, but its course.) In such a circumstance, a question that merits consideration is what to call the earliest clinical expression of Alzheimer’s disease.

Rates of progression are of particular importance, because predictive validity is an aspect of criterion validation, which is held as the highest form of validity. In a systematic review of 19 longitudinal studies, Bruscoli and Lovestone29 concluded that heterogeneity reflected not just how patients were recruited for MCI studies (on this, there is widespread agreement12,30) but also how the criteria were employed. Specifically, they noted that more stringent measurement of deficits resulted in better prediction of conversion, raising the possibility that highly specified MCI represents very early dementia. Similar disagreements about interpretation exist in considering biological markers, such as CSF beta-amyloid and tau. For example, of 52 patients with MCI, those with elevated levels of tau had a higher
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Petersen et al., 1999</td>
<td>Y</td>
<td>Not indicated</td>
<td>N</td>
<td>Not indicated</td>
<td>WAIS-R</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>WMS-R</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Auditory Verbal Learning Test</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Wide-Range Achievement Test II</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MMSEI</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Dementia Rating Scale (DRS)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Free and Cued Selective Reminding Test</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Boston Naming Test</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Controlled Oral Word Association Test</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Category Fluency Procedures</td>
<td></td>
</tr>
<tr>
<td>Morris et al., 2001</td>
<td>Not indicated</td>
<td>Y</td>
<td>Not indicated</td>
<td>Not indicated</td>
<td>CDR</td>
<td>=0.5</td>
</tr>
<tr>
<td>Bennett et al., 2002</td>
<td>N</td>
<td>Y</td>
<td>Not indicated</td>
<td>Not indicated</td>
<td>Extensive neuropsychological battery</td>
<td>Population specific cut-off scores used</td>
</tr>
<tr>
<td>Dafny et al., 2002</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Not indicated</td>
<td>CERAD word list recall</td>
<td>&lt;0.5</td>
</tr>
<tr>
<td>Hümmer et al., 2002</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>MMSE&lt;20</td>
<td>CERD Selective Reminding Test (total immediate recall during 6 trials + delayed recall)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Visual reproduction Test (WMS) (immediate and delayed recall)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Logical Memory Test (WMS) (immediate and delayed recall)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Boston Naming Test (abbreviated)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Verbal Fluency Test</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Trail making Test (A-II)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MMSE</td>
<td></td>
</tr>
<tr>
<td>Kabau et al., 2002</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Not indicated</td>
<td>CDR</td>
<td>=0.5</td>
</tr>
<tr>
<td>Larrieu et al., 2002</td>
<td>Y</td>
<td>Y</td>
<td>MMSE&gt;1 SD</td>
<td>Not indicated</td>
<td>Henton Visual Retention test</td>
<td>≥1 SD below age adjusted norms</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>from age &amp; ed.</td>
<td></td>
<td>Structured Mental Status (SDSAMS)</td>
<td>≥1 SD below age &amp; ed. Adjusted means</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&gt;1 SD below average &amp; ed.</td>
<td></td>
</tr>
<tr>
<td>Fuxe et al., 2003</td>
<td>Examined as a clinical variable</td>
<td>Y</td>
<td>N</td>
<td>Not indicated</td>
<td>CDR Tests of explicit memory</td>
<td>≥1 SD below age adjusted norms</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Henton Visual Retention test</td>
<td>≥1 SD below age adjusted norms</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Structured Mental Status (SDSAMS)</td>
<td>≥1 SD below age &amp; ed. Adjusted means</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&gt;1 SD below average &amp; ed.</td>
<td></td>
</tr>
<tr>
<td>Rask et al., 2005</td>
<td>Examined as a clinical variable</td>
<td>N</td>
<td>Y</td>
<td>Not indicated</td>
<td>CERAD 10 word list delayed recall</td>
<td>&lt;1 SD below mean of cohort</td>
</tr>
<tr>
<td>Lambo et al., 2003</td>
<td>Y</td>
<td>N</td>
<td>MMSE&gt;24</td>
<td>Not indicated</td>
<td>Logical Memory (immediate) - DRS (memory)</td>
<td>≥1.5 SD below means norms</td>
</tr>
<tr>
<td>Lopez et al., 2003</td>
<td>Necessary for ‘probable’ MCI, but not ‘possible’ MCI</td>
<td>Y</td>
<td>Y</td>
<td>Not indicated</td>
<td>n/a</td>
<td>≥1.5 SD below a sample of 280 unimpaired subjects</td>
</tr>
<tr>
<td>Fadale et al., 2004</td>
<td>(InDDeX study)</td>
<td>Y</td>
<td>Y</td>
<td>Not indicated</td>
<td>CERD</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NYU Paragraph recall</td>
<td>&lt;9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HAM-D</td>
<td>&lt;13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HAM-D item 1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Feldman et al., 2004</td>
<td>N</td>
<td>Not indicated</td>
<td>Not indicated</td>
<td>Not indicated</td>
<td>CERD</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NYU paragraph recall</td>
<td>&lt;9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HAM-D</td>
<td>&lt;13</td>
</tr>
<tr>
<td>Gianguli et al., 2004</td>
<td>Y</td>
<td>Not indicated</td>
<td>N</td>
<td>MMSE&gt;24</td>
<td>CERAD 10 word list delayed recall</td>
<td>≥1 SD below mean of cohort</td>
</tr>
<tr>
<td>Grundman et al., 2004</td>
<td>(ADCS-MIS study)</td>
<td>Y</td>
<td>Not indicated</td>
<td>Y</td>
<td>MMSE&gt;24</td>
<td>≥1.5 SD below average score of the sample on each measure</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Logical Memory II</td>
<td>≥25(=16 yrs. ed.), ≥24 (8 to 15 yrs. Ed.)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hamilton Dep. Rating Scale</td>
<td></td>
</tr>
<tr>
<td>Royall et al., 2004</td>
<td>Not indicated</td>
<td>Examined as a clinical variable</td>
<td>Y</td>
<td>Not indicated</td>
<td>MMSEI</td>
<td>≥1 SD below average score of the sample on each measure</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CLOX 1, 2 (clock drawing test)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>COWA (verbal fluency)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EXIT 25 (executive interview)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Trail A = +1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CVLT (verbal memory)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DRS: MEM</td>
<td></td>
</tr>
<tr>
<td>Solfrizzi et al., 2004</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>MMSE&gt;1 SD from age &amp; ed.</td>
<td>Hatakawa Story Recall Test</td>
<td>Score in the lowest 10th percentile of the cohort</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>From age &amp; ed.</td>
<td></td>
</tr>
<tr>
<td>deJager et al., 2005</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Not indicated</td>
<td>CERAD 10 word list (free and delayed recall)</td>
<td>≥1 SD below means norms</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Verbal Paired Associates (cued recall test)</td>
<td>≥10 points below WAIS-R verbal IQ score. If no deficits on objective tests, memory complaint + informant’s confirmation of decline and functional decline</td>
</tr>
<tr>
<td>Devanand et al., 2005</td>
<td>Y</td>
<td>Y</td>
<td>MMSE&gt;22</td>
<td>MMSE&gt;1 SD from mean</td>
<td>CERAD 10 word list (free and delayed recall)</td>
<td>≥2 of 3 items recalled</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Or</td>
<td>≥1 SD below means norms</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Selective Reminding Task</td>
<td>≥10 points below WAIS-R verbal IQ score. If no deficits on objective tests, memory complaint + informant’s confirmation of decline and functional decline</td>
</tr>
<tr>
<td>Geslani et al., 2005</td>
<td>Y</td>
<td>Not indicated</td>
<td>N</td>
<td>WAIS-R similarities or digit symbol</td>
<td>≥1 SD below age &amp; ed adjusted norms</td>
<td></td>
</tr>
</tbody>
</table>

Table: MCI criteria used in earlier studies
risk of later being diagnosed with AD.31 On the other hand, their baseline tau levels considerably overlapped into the AD range, again raising the question of how appropriate is the concept of “conversion” compared with “reassignment of diagnostic category”. Similar results were seen in an earlier smaller study.32

**Clinical trials in MCI as a form of predictive validity**

The evaluation of predictive validity is another way in which cholinesterase inhibitor trials might be understood. In essence, the clinical trials can operate, at the level of the concept, as a diagnostic and therapeutic trial often operates at the level of the patient. Initial results of studies with cholinesterase inhibitors have yielded equivocal results.33-36 In consequence, it is not yet possible to conclude whether MCI is a form of dementia with a variable treatment response, or whether it is a separable entity for which treatment might or might not be preventive. The MCI studies themselves varied; while they demonstrated a strong relationship between apoE4 and the risk of developing clinically evident dementia,34 they varied importantly in the proportions who carried an ApoE4 allele.37 as well as the proportion who might, in other contexts, be described as having “very early dementia”.38

In dementia studies, there is a strong tradition of considering criterion validity chiefly in the form of referent validation, with the referent being a so-called “gold standard”. The “gold standard” was held to be neuropathology, although the existence of several sets of neuropathological criteria that can yield conflicting results diminishes the luster of this approach somewhat.39-40 Importantly, not all studies show a dose response between putative pathogenic factors and NCI, MCI and AD.41-43

On the other hand, neurocompensatory responses might be more specific, so that static levels might not reflect dynamic changes,44-46 making simple correlative studies- even with neuropathology - an inadequate test of validity.

**Conclusions**

In general, it seems logical to propose that cholinesterase inhibitors having not been demonstrably effective reflects either that: the MCI concept is valid, but the drugs are ineffective, or; MCI is valid, the drugs are effective, but the outcome measures used to detect treatment effects have been insensitive to clinically important change, or; the MCI concept is not valid enough to select operational criteria which pick out people with demonstrable memory impairment who might respond to treatment with a cholinesterase inhibitor. The present data do not readily allow for these possibilities to be distinguished, but they do show that even small differences in clinical trials enrollment criteria appear to make big differences in the rates of progression, even if attempting to have more rigorous criteria in population studies does not affect progression (or recovery).

An important policy consideration in knowing whether MCI is a valid target of treatment is that, as we have seen, MCI is not the only form of cognitive impairment that falls between normal cognitive function and dementia. In consequence, it is clear that questions about definition - which is central to the current controversies about MCI as a therapeutic target - are essential features and not regulatory niceties.37 Definitional nuances also have policy implications in that minor changes in how MCI is defined can result in three-fold differences in prevalence.13,47,48 If MCI is to become a treatable “disease” then careful attention needs to be paid to criteria, lest it become a rapidly spreading epidemic!
We note that MCI is potentially a target of therapy not just for cholinesterase inhibitors. There is now a set of publications which demonstrate a lack of depletion of choline acetyltransferase in MCI, or in mild AD brains, for that matter.\textsuperscript{44,49,50} Unimpressive symptomatic trials of ChEIs would support this view. Additionally, MCI is being evaluated as an entity driven by cerebrovascular disease, suggesting that it might also be a target for vascular risk factor modification.\textsuperscript{51} Other compounds that have been evaluated for use using current MCI criteria include rofecoxib\textsuperscript{52} and Vitamin E.\textsuperscript{33}

Many of the controversies about how to conceptualize MCI are amenable to pragmatic research programs, and should motivate better funding for systematic inquiry. Future studies might well make the controversy obsolete. For example, if an anti-amyloid compound that is potentially disease-modifying – that might, in fact, prevent dementia – were to be studied, then it might well be studied in people who were genetically at risk, even if they did not have MCI/very early Alzheimer’s disease. On the other hand, unless prevention of dementia is completely effective, it might well be that future ‘prevented dementia’ would look more like MCI (or CIND) than like normal cognitive function.

If many people with MCI actually have early dementia, which increasingly seems likely,\textsuperscript{53} then how to identify this group needs better attention. It is likely that clinically sensible tests of executive function will be of particular value.\textsuperscript{26} It might also be that the comparatively little attention paid to behavioural manifestations has lessened the sensitivity and specificity of the MCI construct.\textsuperscript{54-56} In this regard, an epidemiological and clinical research program that evaluated MCI in relation to depression would be of particular value.\textsuperscript{27} Further specification of the number and types of domains that are affected might also yield greater specificity.\textsuperscript{56} Hypotheses such as these – and the hypothesis that intervention can delay progression in very early dementia - merit specific testing, using both standard tests, and more sensitive measures than those now generally employed. Such a program would require a much more sophisticated evaluation of patients than is now routinely available outside of the academic research setting in Canada. In consequence, it appears that MCI as a high-risk state putatively distinct from very early dementia is best understood as an entity that requires additional research, including at the level of concept validation.\textsuperscript{59} Intervention studies should therefore be undertaken with a view to clarifying the concept, targets of treatment and identification of potential responder groups.

\textbf{DECLARATION}

Kenneth Rockwood is supported by an Investigator Award from the Canadian Institutes of Health Research (CIHR) and the Dalhousie Medical Research Foundation as the Kathryn Allen Weldon Professor of Alzheimer’s Research. Howard Chertkow is a “chercheur national” of the FRQS (Fonds de la Recherche en Santé du Québec).

Kenneth Rockwood, Howard Chertkow and Howard Feldman each receive research support from the CIHR. Dr. Rockwood receives research funding from the Alzheimer Society of Canada. Dr. Chertkow also receives research funding from the FRQS.

Kenneth Rockwood holds research grants from Janssen Ortho and Merek, and has received speakers fees from Janssen Ortho, Merck, Novartis and Pfizer. Howard Feldman has received grant funding from Pfizer, Eisai, and Janssen Pharma. These companies are manufacturers of drugs being investigated for the treatment of mild cognitive impairment.

Within the past 2 years Howard H. Feldman has served as a consultant, received grant funding or participated in CME programs sponsored by Astra Zeneca, Pfizer, Eisai, Novartis, Janssen, Servier, Sanofi Synthelabo, Lilly, GlaxoSmithKline, Myriad, Targacept, Forest, Lundbeck and Aponexy.

\textbf{REFERENCES}


