Enhancing detection of those at-risk for dementia:
A revised classification procedure for Mild Cognitive Impairment

by

Susan Diane Vanderhill
B.A., Yale University, 2001
M.A., University of Victoria, 2004

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Psychology

Enhancing detection of those at-risk for dementia:
A revised classification procedure for Mild Cognitive Impairment

by

Susan Diane Vanderhill
B.A., Yale University, 2001
M.A., University of Victoria, 2004

Supervisory Committee

Dr. Esther Strauss, Co-Supervisor
(Department of Psychology)

Dr. David. F. Hultsch, Co-Supervisor
(Department of Psychology)

Dr. Michael A. Hunter, Departmental Member
(Department of Psychology)

Dr. Denise Cloutier-Fisher, Outside Member
(Department of Geography)

Supervisory Committee

Dr. Esther Strauss, Co-Supervisor
(Department of Psychology)

Dr. David. F. Hultsch, Co-Supervisor
(Department of Psychology)

Dr. Michael A. Hunter, Departmental Member
(Department of Psychology)

Dr. Denise Cloutier-Fisher, Outside Member
(Department of Geography)

ABSTRACT

Evidence for the utility of the Mild Cognitive Impairment (MCI) classification as a predictor of impending dementia in older adults is somewhat limited. Although individuals with MCI show elevated rates of conversion to dementia at the group level, heterogeneity of outcomes is common at the individual level. Using data from a prospective five-year longitudinal investigation of cognitive change in a sample of 262 healthy older adults aged 64 to 92 years, this study was designed to address key limitations in current MCI classification procedures which tend to rely on single occasion assessment (Traditional MCI) by evaluating an alternate operational definition of MCI requiring evidence of persistent cognitive impairment over multiple testing sessions (Persistent MCI), and four subsequent variations of this operational definition. It was hypothesized that: (1) prevalence of Traditional MCI would exceed prevalence of Persistent MCI across all variations in the operational definition, (2a) both the Traditional MCI and Persistent MCI groups would show lower levels of performance and greater decline in both cognitive and functional status over five years relative to Controls,

(2b) the magnitude of these differences between those classified as Persistent MCI and Controls would exceed the magnitude of differences between those classified as Traditional MCI and Controls, and (3) the pattern of findings outlined in hypothesis 2 would persist under the four variations of the Traditional MCI and Persistent MCI inclusion criteria. Results were consistent with Hypothesis 1, and partially consistent with Hypotheses 2 and 3. In general, the Persistent MCI groups showed a lower mean baseline level of performance and a steeper trajectory of cognitive decline compared to the Control group and the Traditional MCI groups, although the sample-wide change in cognitive and functional status was small. There was some evidence that the variation of Persistent MCI classification which specified persistent memory impairment as an inclusion criteria achieved optimal prediction of cognitive and functional decline. Results are discussed with reference to retest effects, cognitive reserve, and clinical utility of the Persistent MCI concept for enhancing prediction of dementia in older adults.

Table of Contents

List of Tables

List of Figures

Acknowledgements

First and foremost, I would like to express my deepest gratitude to Dr. Esther Strauss, for taking me on as a graduate student so long ago and showing me every kindness, affording me every opportunity, and helping me grow immeasurably along the way. Equal thanks to Dr. David Hultsch for his patient guidance, teaching, and support. It has been an honour and a pleasure to train under two exceptional mentors who each exhibit the highest ideals of academic scholarship, enriched by near-implausible humility and genuine caring.

I would also like to offer very special thanks to Dr. Michael Hunter who has been exceptionally generous with his time and teaching over the years, and to Dr. Stuart Macdonald for his unexpected and substantial offering of time and teaching over the last year. Thanks to Dr. Denise Cloutier-Fisher, my outside member, and Dr. Grant Iverson, my external examiner, for sharing their time, insights and expertise.

I am grateful for the financial support of my doctoral training from the University of Victoria, the Michael Smith Foundation for Health Research, and the British Columbia Medical Services Foundation.

My special thanks go to the Project MIND participants and staff for their time and dedication. In particular, I'd like to thank project coordinator, Arlene Senft, and my graduate student colleagues, Allison Bielak and Jing Tan, for maintaining a delightful spirit of collaboration in the lab, both at work and at play.

Finally, I would especially like to thank my family for their constant support and encouragement and to my friends for the laughter and perspective. Most critically, I want to say thank you to my wife, Ashley, for her patience, love, and inspiration.

Introduction

Advances in treatments for Alzheimer's disease and related dementias have shifted research efforts toward the early identification of individuals likely to develop these disorders. The term Mild Cognitive Impairment (MCI) has been used as a label for individuals who show cognitive impairment relative to their healthy peers, but do not meet full criteria for any dementia syndrome (Flicker, Ferris, & Reisberg, 1991; Smith, Petersen, Parisi, & Ivnik, 1996; Zaudig, 1992). Although MCI has been conceptualized as a precursor to dementia (Petersen et al., 1999), evidence for the utility of MCI as a predictor of impending dementia has been found to be somewhat limited. While individuals with MCI do show elevated rates of conversion to dementia at the group level, heterogeneity of outcomes is common at the individual level (for reviews see Bruscoli & Lovestone, 2004; Petersen, 2004; Tuokko & McDowell, 2006). Some individuals with MCI develop dementia, some remain stable for long periods, and some revert to unimpaired status. The aim of the current study was to improve existing methods for identifying those at greatest risk of dementia.

MCI: Overview

Although the incidence of cognitive disorders such as Alzheimer's disease and related dementias increases dramatically with age (Petersen, 2003), dementia is no longer thought to be an inevitable outcome in aging. However, as most dementia syndromes are characterized by an insidious onset and progressive course, intense effort has been devoted toward clarifying the boundary between the subtle cognitive changes associated with normal aging and cognitive changes that may represent earliest stages of a dementia

syndrome. As early as 1962, Kral differentiated "Benign" and "Malignant" Senescent Forgetfulness, the latter characterized by its progressive course and poor prognosis (Kral, 1962). Since that time, a host of labels and classification schemes have been proposed to characterize those individuals showing cognitive functioning below expected levels, but insufficient to impair their functional abilities and thereby warrant a dementia diagnosis (see Table 1). These labels include Age Associated Memory Impairment (AAMI; Crook, Bartus, Ferris, & Whitehouse, 1986), Late-Life Forgetfulness (Blackford & la Rue, 1989), Mild Cognitive Decline (WHO, 1993), Aging-Associated Cognitive Decline (AACD; Levy, 1994), Aging-Related Cognitive Decline (ARCD; APA, 1994), Mild Neurocognitive Decline (MND; APA, 1994), Cognitive Impairment No Dementia (CIND; Ebly, Hogan, & Parhad, 1995; Graham et al., 1997), and Mild Cognitive Impairment (MCI; Petersen et al., 1999; Smith et al., 1996). Although these classification schemes differ somewhat with respect to their inclusion/exclusion criteria (e.g., severity of impairment required for diagnosis, extent to which other potential causes of impairment must be ruled out) and underlying pathological/prognostic assumptions (e.g., AAMI describes "normal" decline associated with aging while MND implies a pre-dementia state), they share a common goal of characterizing older adults with suboptimal cognitive functioning in the context of some consideration of, but failure to meet full criteria for, a dementia syndrome (for reviews see Davis & Rockwood, 2004; Tuokko & McDowell, 2006).

Table 1. Classification Schemes for Cognitive Disorders of Aging of Insufficient Severity to Warrant Dementia Diagnosis

Diagnostic Label	Diagnostic Criteria
Benign Senescent Forgetfulness (Kral, 1962)	Difficulty with recall of relatively unimportant elements of past experience. No assumption of a progressive course or underlying disease.
Malignant Senescent Forgetfulness (Kral, 1962)	Memory loss associated with forgetfulness, disorientation, confabulation, and poor memory testing performance. Progressive course and presumed underlying disease.
Age Associated Memory Impairment (Crook et al., 1996)	Gradual onset memory complaints confirmed by memory test performance 1.0 SD below mean test value for young adults. Ages 50+.
Late-Life Forgetfulness (Blackford & la Rue, 1989)	Preserved general intelligence and perceived decrease in everyday memory confirmed by standardized self-report questionnaire. Performance between 1 and 2 SDs below the mean for own age on 50% or more of (4+) memory tests administered. Ages 50-79.
Mild Cognitive Disorder (WHO, 1993)	Decline in cognitive performance (confirmed objectively) not attributable to other mental or behavioral disorders identified in ICD-10. May be reversible.
Aging-Associated Cognitive Decline (Levy, 1994)	Gradual decline in any one of memory and learning, attention and concentration, thinking, language, or visuospatial functioning present for at least 6 months. Performance at least 1 SD below normative mean for age on relevant cognitive tests.
Aging-Related Cognitive Decline (APA, 1994)	Objectively identified decline in cognitive functioning that is within normal limits relative to same-aged peers.
Mild Neurocognitive Decline (APA, 1994)	Presence of two or more areas (e.g., memory, executive functioning, attention, processing speed, perceptual-motor abilities, or language) of objective cognitive impairment or decline lasting most of the time for at least 2 weeks (reported by person or informant). Objective evidence of a neurological or general medical condition that is judged to be etiologically related to the cognitive disturbance.
Cognitive Impairment No Dementia (Ebly et al., 1995, Graham et al., 1997)	Presence of objective cognitive impairment but no dementia diagnosis. Etiology not specified (i.e., may be caused by depression, alcoholism, vascular, mental illness, mental retardation, etc.).
Mild Cognitive Impairment (Petersen et al., 1999, Smith et al., 1996)	Complaint of memory impairment, normal activities of daily living, normal general cognitive functioning, abnormal memory function for age, and absence of dementia.

The term MCI is most commonly employed in contemporary research literature and will be used in this paper. It should be noted that some authors have advocated strongly for the exclusive use of the term MCI to describe a specific clinical syndrome, presumed to be a precursor to Alzheimer's disease, characterized by subjective and objective memory impairment in the absence of other cognitive impairment or functional decline (Petersen et al., 1999). A recent consensus report from the International Working Group on Mild Cognitive Impairment recommended a similar, yet broadened conceptualization of MCI characterized by subjective and objective impairment (or documented decline) in *any* cognitive domain in the absence of significant functional decline (Winblad et al., 2004) with guidelines for sub-classifying individuals with MCI according to the nature (e.g., amnestic versus nonamnestic) and extent (e.g., single cognitive domain or multiple cognitive domains) of impairment. Although these recommendations and guidelines are informative, elements of these specific criteria (particularly the requirements for subjective complaint and lack of significant functional decline) have been disputed or inconsistently applied in many studies, leading many authors to use the term MCI as a generic label to refer to non-demented individuals showing non-normative impairment in cognitive functioning (e.g., Bruscoli & Lovestone, 2004; Luis, Loewenstein, Acevedo, Barker, & Duara, 2003; Tuokko & McDowell, 2006) as will be done in the current paper.

MCI: Prevalence, Incidence, and Conversion to Dementia

Evidence to support the potential utility of MCI as a predictor of dementia comes from longitudinal studies of individuals with MCI using both clinical and

epidemiological samples. Such studies have produced varying prevalence and incidence rates for MCI and varying rates of conversion from MCI to dementia. A recent review of nearly 40 major studies of MCI identified prevalence estimates for MCI in older adult samples ranging from 1 to 36% (Tuokko & McDowell, 2006). Incidence rates for MCI, less frequently reported, ranged from 8-77/1000 per year. Rates of conversion from MCI to dementia ranged from 5-30% per year. The authors noted that variation in prevalence, incidence, and conversion rates may be accounted for by differences in syndromes studied (e.g., AAMI versus AACD), operational definitions of "impairment", statistical methodologies, presentation of results, study design, and sampling and measurement techniques.

In a prior review of 26 MCI conversion studies, Bruscoli and Lovestone (2004) compared studies on the following: specific criteria for inclusion in the MCI group, mean age of participants, sample source (e.g., clinic versus community), mean follow-up interval, and specific criteria for conversion to dementia diagnosis. These authors reported that, across studies, annual conversion rates ranged from 2-31%, with a mean annual conversion rate of 10% per year. Conversion rates showed no significant relation to age of sample or length of follow-up period, but did differ significantly by sample source. That is, the mean annual conversion rate across clinical samples (15%) was approximately twice as high as the rate among community volunteer samples (8%). Of the seven variables that were studied in at least two of the reviewed studies as potential predictors of conversion to dementia among those with MCI (e.g., age, gender, APOE4 status, education, neuropsychological test scores, neuroimaging, and EEG), only baseline neuropsychological testing was an unequivocal predictor of negative course in MCI.

That is, of the 15 studies that investigated baseline neuropsychological testing as a predictor of conversion, all studies showed that, among individuals with MCI, those who go on to convert to dementia have lower baseline neuropsychological performance than those who do not convert.

MCI: Instability of Classification

While variations in rates of conversion to dementia raise questions about the utility of the MCI classification as an indicator of impending dementia, perhaps more striking are the pervasive reports that a significant minority of individuals classified as MCI at one time point fail to demonstrate cognitive impairment at a subsequent time point. Such individuals are said to "revert" from MCI to unimpaired or cognitively normal status. Rates of reversion have been reported in a number of major epidemiological samples. Ritchie and colleagues reported a 15% rate of reversion across a one year follow up interval in a French sample (Ritchie, Artero, & Touchon, 2001). Rates of reversion in a second, independent French sample ranged from 32-41% over two years, depending on the operational definition of MCI (Larrieu et al., 2002). In a German sample, rates of reversion across a one and a half year follow up interval ranged from 18-22%, again depending on the operational definition of MCI (Busse, Hensel, Guhne, Angermeyer, & Riedel-Heller, 2006). In an American sample, 28% of individuals with MCI reverted to normal after a two year follow up interval (Ganguli, Dodge, Shen, & DeKosky, 2004). Two recent studies from the Canadian Study of Health and Aging documented five-year rates of reversion from CIND (operationalized as exclusion of dementia plus clinical impression of some cognitive decline) and four different versions

of amnestic MCI (with and without subjective memory impairment, with and without intact instrumental activities of daily living or IADLs). Results of these studies indicated that 13% of individuals characterized as CIND were cognitively normal at follow up (Tuokko et al., 2003). Of the individuals characterized as amnestic MCI, 26-32% (depending on amnestic MCI version under consideration) were cognitively normal at five-year follow-up (Fisk, Merry, & Rockwood, 2003).

Epidemiological studies, because of the large sample sizes involved, often employ less detailed cognitive testing protocols than those used in some smaller cohort studies and most clinical settings. As such, some of the instability of MCI classification in these large studies might be attributed to unreliability of cognitive measures employed within the classification scheme, rather than real change in an individual's status from cognitively impaired to not impaired. However, data from two smaller-scale cohort studies employing more detailed, and presumably more reliable, cognitive assessment in the classification of MCI document nontrivial rates of reversion. In a community cohort of 157 older adults tested on a battery of seven neuropsychological measures of memory, de Jager and Budge (2005) documented a reversion rate of 35% over a two year interval. In a sample of 70 research participants with MCI identified by clinical consensus review of neurological and neuropsychological evaluation data, 7% of those identified as amnestic MCI and 17% of those identified as nonamnestic MCI reverted to normal at one year follow up (Loewenstein, Acevedo, Agron, & Duara, 2007). Even within clinical samples, where the overall rates of conversion to dementia have been shown to be higher compared to community samples (Bruscoli & Lovestone, 2004), the phenomenon of reversion has been documented. One study followed a clinic sample of 127 outpatients

with MCI (operationalized as Clinical Dementia Rating score CDR= 0.5). They found that of the 75 patients with complete follow-up data, almost half (44.4%, or 26.2% of the original sample) reverted to nonimpaired status (CDR=0) after one year (Devanand, Folz, Gorlyn, & Moeller, 1997).

MCI: Limitations of Single-Session Assessment and Single-Test Impairment

The same factors that may account for variation in conversion rates may also contribute to variation in reversion rates in MCI (e.g., varying inclusion/exclusion criteria for type of cognitive impairment and age of participants, use of clinic versus community samples, differing methods of identifying cognitive impairment, varying reliability of cognitive assessment measures, sample sizes, statistical methods, length of longitudinal follow-up, etc.). However, an important and largely overlooked issue in MCI research that may be a significant contributor to classification instability is the reliance on single-session assessment and single-test impairment. That is, although wide variation in operational definitions of MCI exists across studies, in practice, the vast majority of the recent studies reviewed above specify an objective cognitive impairment inclusion criterion for MCI which is operationalized as *impaired performance on a single psychometric test at a single time point* (Busse et al., 2006; de Jager & Budge, 2005; Ganguli et al., 2004; Larrieu et al., 2002; Loewenstein et al., 2007; Ritchie et al., 2001). This practice, although common, is highly problematic. Large-scale normative data collection efforts have demonstrated that isolated "impaired" scores on neuropsychological measures are relatively common among "normal" samples. For example, within the carefully screened, neurologically normal sample employed by

Heaton and colleagues (1991) to generate comprehensive normative data for the Halsted-Reitan neuropsychological battery, ninety percent of participants obtained at least one score in the "abnormal" range (i.e., T-score less than or equal to 39).

More recent data specific to older adults come from two studies by Brooks and colleagues (Brooks, Iverson, Holdnack, & Feldman, 2008; Brooks, Iverson, & White, 2007) who examined base rates of low memory scores among neurologically normal older adults in the Neuropsychological Assessment Battery (NAB; Stern & White, 2003) normative sample (age 55-79, N=742) as well as the Wechsler Memory Scale – Third Edition (WMS-III, Wechsler, 1997) normative sample (age 55-87, N=550). On the NAB Memory Module, a battery of four memory tests (i.e., List Learning, Shape Learning, Story Learning, Daily Living Memory) that yields ten subtest T scores based on age-, gender-, and education-corrected norms, over half (55.5%) of the "normal" individuals had at least one of ten subtest scores greater than 1 SD below the mean for their demographic group, 30.8% of individuals had one score greater than 1.5 SD below their group mean, and 16.4% had at least one score greater than 2 SD below their group mean (Brooks et al., 2007). The proportion of individuals scoring below specified cut-off scores was greatly influenced by estimated intelligence level. For example, 80.1% of individuals with low-average intellectual functioning obtained one score below a 1 SD cutoff compared with 46.4% of individuals with high-average intellectual functioning. Similarly, on the WMS-III, a battery of memory measures comprised of 11 subtests (4 of which were considered in the present study, yielding a total of 8 scores), over half (64.1%) of the "normal" individuals had at least one of eight subtest scores greater than 1 SD below the mean for their age group and 70% of the sample had at least one of eight

scores greater than 1 SD below the mean for their demographic group (Brooks et al., 2008). Similarly, 12.9% had at least one of eight subtest scores greater than 2 SD below the mean for their age group and 21.6% had at least one of eight scores greater than 2 SD below the mean for their demographic group. Again, the proportion of individuals scoring below cut-off scores was strongly influenced by estimated intelligence level. The authors discussed the implications of these findings in terms of the risk for identifying normal individuals as "accidental MCI," a term that was first coined by de Rotrou and colleagues (2005) to refer to individuals who are diagnosed with MCI at one time point, but later show cognitive test performance in the normal range.

Although base-rate data from large-scale, single-battery normative samples are informative, many neuropsychologists take a "flexible-battery" approach (c.f., Lezak, Howieson, & Loring, 2004; Strauss, Sherman, & Spreen, 2006), employing a series of specific tests from a range of sources to selectively tailor their clinical data gathering to the cognitive domains of interest. Palmer and colleagues had this approach in mind when they examined psychometric test scores in a carefully screened healthy, neurologically normal, older adult sample (age 50-80; N=132) across a selection of neuropsychological tests commonly employed in clinical practice (Palmer, Boone, Lesser, & Wohl, 1998). The authors tested individuals on a relatively brief (two and a half hour) session which included the following tests: the Mini-Mental State Exam (Folstein, Folstein, & McHugh, 1975), Weschler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981) Digit Span and Digit Symbol, Weschler Memory Scale-Revised (WMS-R; Wechsler, 1987) Logical Memory and Visual Reproduction, Stroop Words and Colors (Goodglass & Kaplan, 1979), Boston Naming Test (Kaplan, Goodglass, & Weintraub, 1983), Controlled Oral

Word Generation Test (Lezak, 1995), Rey-Osterreith Complex Figure Test (Lezak, 1995), Recognition Memory Test (Warrington, 1984), Auditory Consonant Trigrams (Stuss et al., 1982), and Wisconsin Card Sorting test (Heaton, Chelune, Talley, Kay, & Curtiss, 1993). Results indicated that 73% of individuals obtained at least one score in the range traditionally considered "below normal limits" (e.g., 1.3 SD below the normative mean; below the 9th percentile) and 48% had two scores in this range. Further, 37% of individuals had at least one score in the "impaired" range (e.g., 2 SD below the mean; below the 3rd percentile) and 24% had two scores in this range. Proposed explanations for the high proportion of purportedly normal individuals evidencing impaired performance include transient fatigue, low mood, or loss of motivation on the part of the participant.

Taken together, these findings suggest that current single-session assessment practices with single-test impairment inclusion criteria for MCI will likely lead to elevated rates of false positives that may, in turn, account for some of the heterogeneity of outcomes for individuals diagnosed as MCI. Two tactics may be employed to address these issues. First, multiple-session assessments may be employed to improve differentiation of those individuals with stable neuropsychological impairment or decline from those showing "accidental" poor performance on testing due to transient factors such as low mood or motivation, or fatigue. Second, multiple-test impairment inclusion criteria could be employed to improve differentiation of those individuals with robust neuropsychological impairment from those showing isolated low scores due to accidental poor performance. Of these two tactics, the former may be best suited to epidemiological and community-based longitudinal research studies where individual cognitive test

batteries may be minimal, but data are collected at multiple time points. The latter may be better suited to clinical studies where emphasis may be placed on gathering more complete cognitive assessment data to make more reliable diagnoses of MCI.

To date, one existing study has employed a multiple-session assessment approach in a community sample. Collie and colleagues performed repeat assessment of older adults semi-annually over a period of one year (Collie, Maruff, & Currie, 2002). They found that while roughly 20% of study participants met criteria for MCI (defined as performance 1.5 SD below normative mean on a measure of delayed verbal recall) at any one testing session, only 13% met criteria at all three sessions. The latter, "Persistent MCI" group, captured those individuals with more consistent cognitive difficulty including those whose already low cognitive ability was on a declining trajectory. These authors suggested that individuals with Persistent MCI are likely at greater risk of dementia than those showing transient poor performance during a single assessment. The current study was designed to replicate this work using slightly modified procedures and extend these findings by examining the longitudinal course of cognitive and functional outcomes for those identified as Persistent MCI.

Objectives and Hypotheses

Using data from a prospective five-year longitudinal study of cognitive change in community-dwelling older adults, the current study was designed to accomplish the three objectives listed below.

Objective 1: Prevalence of Traditional MCI versus Persistent MCI

The first Study Objective was to replicate the findings of Collie et al. (2002) regarding the discrepancy in prevalence rates for Traditional MCI versus Persistent MCI. Using procedures adapted from Collie et al., algorithms were applied to the current sample to identify the rates of cognitive impairment observed at a single measurement occasion (i.e., Traditional MCI) versus the rates of cognitive impairment observed at consecutive measurement occasions (i.e., Persistent MCI). Follow up exploratory analyses were performed using systematic variations in classification algorithms to provide prevalence estimates for MCI groups operationalized with varying severity, duration, pervasiveness, and specificity of cognitive impairment.

Hypothesis 1: Base-rates of impaired performance on any one cognitive test at any one time point (Traditional MCI) were expected to exceed base-rates of persistent impaired performance observed on consecutive assessment sessions (Persistent MCI). This discrepancy was expected to hold across all operational definitions of Traditional and Persistent MCI.

Objective 2: External Validity of Persistent MCI Classification

The second Study Objective was to provide evidence for the external validity of the Persistent MCI classification procedures. Because previous studies have demonstrated that those with Traditional MCI go on to develop dementia, a syndrome characterized by impairment in cognitive and functional status, at higher rates relative to controls, it was expected that:

Hypothesis 2a: Both the Traditional MCI and Persistent MCI groups would show lower levels of performance and greater decline in both cognitive and functional status over five years relative to Controls.

Further, and to the extent that the Persistent MCI classification improves upon limitations in Traditional MCI classification, thereby more accurately capturing those at risk of dementia, it was expected that:

Hypothesis 2b: The magnitude of difference in baseline level of performance and longitudinal trajectory of decline in cognitive and functional status between those classified as Persistent MCI and Controls would exceed the magnitude of difference in baseline level of performance and longitudinal trajectory of decline in cognitive and functional status between those classified as Traditional MCI and Controls.

Objective 3: Robustness of Persistent MCI Classification

The final Study Objective was to explore the robustness of the Persistent MCI classification scheme by varying elements of the ("Basic") classification algorithm for Traditional and Persistent MCI in four ways. Because there is much debate in the literature regarding the optimal operational definition for MCI, cognitive and functional outcomes under four variations of the Traditional and Persistent MCI inclusion criteria were investigated. Under the first variation ("Severity"), a more stringent normative cutoff for inclusion in the MCI groups (i.e., from 1 SD to 1.5 SD) was applied. Under the second variation ("Duration"), the duration of cognitive impairment required for inclusion in a Persistent MCI group was increased (i.e., from two to three consecutive assessments). Under the third variation ("Pervasiveness"), the MCI groups were sub-

classified according to the pervasiveness of cognitive impairment (i.e., impaired performance on single test versus multiple tests). Under the fourth variation ("Specificity"), the MCI groups were sub-classified according to the specific cognitive domain of observed impairment (i.e., amnestic versus nonamnestic impairment).

Hypothesis 3: It was expected that the overall pattern of findings outlined in previous hypotheses would persist under alternate variations of the Traditional MCI and Persistent MCI inclusion criteria. That is, regardless of severity, duration, pervasiveness, or specificity of impairment required for inclusion in the MCI groups, (a) the Traditional MCI and Persistent MCI groups would show lower levels of performance and greater decline in both cognitive and functional status over five years relative to Controls and (b) the magnitude of difference in baseline level of performance and longitudinal trajectory of decline in cognitive and functional status between those classified as Persistent MCI and Controls would exceed the magnitude of difference in baseline level of performance and longitudinal trajectory of decline in cognitive and functional status between those classified as Traditional MCI and Controls.

Methods

Participants

Data for this study were drawn from Years 1-6 of Project MIND, an ongoing prospective investigation of short-term fluctuations in performance and long-term cognitive change in older adults. At baseline, participants were 304 Caucasian community-dwelling adults (208 women and 96 men) aged 64 to 92 years (M = 74.0, SD = 6.0) recruited through local media advertisements seeking individuals who were

concerned about their cognitive functioning, but not diagnosed with any neurological disorder. Exclusionary criteria included physician-diagnosed dementia or a Mini Mental State Examination (MMSE; Folstein et al., 1975) score less than 24, a history of significant head injury (e.g., loss of consciousness greater than 5 minutes), other neurological or major medical illnesses (e.g., Parkinson's disease, heart disease, cancer), severe sensory impairment (e.g., difficulty reading newspaper-size print, difficulty hearing a normal conversation), drug or alcohol abuse, current psychiatric diagnosis, psychotropic drug use, and lack of fluency in English.

Because the Persistent MCI classification scheme under investigation in the current study required cognitive testing data from two serial annual assessments, 42 Project MIND participants who completed Year 1, but not Year 2 cognitive assessment were excluded from the final sample. The excluded participants did not differ from the final sample in terms of age, education, gender, health status (number of chronic conditions), or functional status (activities of daily living). However, the excluded participants had a slightly lower average Mini-Mental State Exam (Folstein et al., 1975) score (M = 28.4, SD = 1.7), relative to the final sample (M = 28.8, SD = 1.1), $F(1, 303) = 4.392$, $p < .05$, $\eta^2 = .01$). The final sample comprised 262 adults (180 women and 82 men) aged 64 to 92 years (M = 73.8, SD = 5.8).

Procedure

Prior to enrolling in the Project MIND, potential participants were screened for inclusion and exclusion criteria by telephone interview. This was followed by two testing sessions (one group and one individual) scheduled over approximately three months. The

group testing session was held in the laboratory and the individual testing sessions were conducted in the participants' homes. At the group session, participants provided demographic and health information, and completed a series of cognitive benchmark measures designed to assess multiple abilities. During the individual session, participants completed a series of individually administered cognitive and functional measures. Over the course of the project, these group and individual testing sessions were repeated annually at Years 1 through 4, and again at Year 6. During Year 5, the face-to-face testing sessions were replaced with a health update and cognitive screening completed via telephone.

Background Information

Demographic information (age, gender, years of education, marital status, native language, ethnic background) was obtained from participants during the initial group testing session and verified annually during the individual sessions. Additional background information collected at baseline and included in the current study included self-reports of memory, health, and depressive affect.

Self-Reported Memory

Self-reported memory was assessed using two items from the Memory Functioning Questionnaire (Gilewski, Zelinski, & Schaie, 1990) in which participants rated their memory compared to (a) one year ago and (b) their same aged peers. Higher scores indicate less decline in memory and better memory relative to peers.

Self-Reported Health

Self-reported health was assessed in multiple ways. A general health

questionnaire in which participants indicated whether they had been diagnosed by a medical practitioner with any of fifteen specified major chronic health conditions (Hultsch, Hertzog, Dixon, & Small, 1998) was used. In addition, participants rated their current state of health on a five point Likert-type scale relative to (a) a perfect state of health and (b) their same-age peers (Hultsch et al., 1998). Higher scores indicate better perceived health.

Self-Reported Depression

Self-reported depression was assessed using the Depressive Affect Subscale of Centre for Epidemiological studies Depression Scale (Hertzog, Van Alstine, Usala, & Hultsch, 1990; Radloff, 1977), on which participants rated the frequency of depressive symptoms experienced over the past week. Higher scores indicate greater severity of depression.

Cognitive Benchmark Measures used for MCI Classification

A series of cognitive benchmark measures were administered at annual group testing sessions. These measures were used to classify participants into MCI groups at Years 1, 2, and 3 according to the algorithms outlined in the upcoming section titled, *MCI Classification Procedures*.

Perceptual Speed

The WAIS-R Digit Symbol Substitution task (Wechsler, 1981) was used to assess perceptual processing speed. A coding key pairing nine numbers (1 through 9) with nine symbols was presented. Underneath the key were rows of randomly ordered numbers with empty boxes below. The participant was asked to copy as many symbols as possible

into the empty boxes based on the digit-symbol pairings in the coding key for 90 seconds. The number of correctly completed items was recorded.

Reasoning

The Letter Series test (Thurstone, 1962) was used to assess inductive reasoning. A series of letters following a distinct pattern was presented. The participant was asked to decipher the pattern in the target string and provide the next letter in the sequence. The number of correct responses out of 20 was recorded.

Episodic Memory

A word recall task (Hultsch, Hertzog, & Dixon, 1990) was used to assess episodic memory. One list of 30 English words selected from a total set of six lists was presented. The list contained six words from each of five taxonomic categories typed on a single page in unblocked order. The participant was given two minutes to study the list and five minutes to write the words they could recall in any order. The number of correctly recalled words was used as the measure.

Verbal Fluency

The Controlled Associations test from the Educational Testing Service (ETS) kit of factor-references cognitive tests (Ekstrom, French, Harman, & Dermen, 1976) was used to assess verbal fluency. The participant was given six minutes to generate as many synonyms as possible in response to a set of target words. The total number of correct synonyms was recorded.

Vocabulary

A recognition vocabulary test was used to assess vocabulary. The test was composed by combining three 18-item tests from the ETS kit of factor-references

cognitive tests (Ekstrom et al., 1976). The participant was given 15 minutes to complete a 54-item multiple-choice task. The number of correct items was recorded.

Cognitive and Functional Outcome Measures

As Alzheimer's disease and related dementias are defined by impairments in cognitive and functional status, a series of measures designed to assess these domains of functioning were administered repeatedly at the annual testing sessions. These measures were selected from available research to map on to key diagnostic features of dementia (APA, 1994) and Alzheimer's disease (Dubois et al., 2007; McKhann et al., 1984), and to assess domains of functioning known to show early decline in Alzheimer's disease (Albert, 2008). Cognitive assessment included measures of global cognitive status, visual recall, and executive functioning. Functional assessment included measures of global functional status and applied problem solving. Specific measures are described below. The schedule of administration is summarized in the Table 2.

Table 2. Administration Schedule of Outcome Measures

Outcome Measure	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6
Cognitive						
Mini-Mental State Exam	x	x	x	x	x *	x
Trail Making Test	x	x	x	x		
Digit Symbol Coding Recall	x	x	x	x		x
Functional						
Activities of Daily Living	x	x	x	x		x
Everyday Problems Test	x	x	x	x		x

Note: Year 5 Mini-Mental State Exam score derived from Telephone Interview for Cognitive Status

Global Cognitive Status

Mini-Mental State Exam (MMSE): The MMSE is a widely used brief screening measure for cognitive impairment (Folstein et al., 1975). This measure surveys cognitive functioning in the domains of orientation, memory, attention, language, and visuoconstructive ability. Higher scores indicate better global cognitive functioning. The MMSE was administered at study Years 1 through 4, and again at Year 6. During Year 5, the Telephone Interview for Cognitive Status (TICS), a telephone adaptation of the MMSE that correlates highly with the original measure and can be used to generate MMSE-equivalent scores (Brandt, Spencer, & Folstein, 1988), was used in place of an individually-administered MMSE. Higher scores on the MMSE indicate higher cognitive functioning.

Executive Functioning

Trailing Making Test Part B (Trails B): The Trail Making Test (Reitan & Wolfson, 1985) is a paper and pencil measure of perceptual speed, sequencing, and

mental flexibility. In the first portion of this task, Part A, the participant is asked to connect 25 randomly arranged encircled numbers as quickly as possible by drawing pencil lines from one to another in numeric order. In the second portion of this task, Part B, the page contains both numbers and letters and the participant must draw pencil lines connecting the number and letters in alternating sequence. Trails A can be considered a measure of simple attention and perceptual speed, whereas Trails B measures these faculties as well as executive functions such as sequencing and mental flexibility. Time to completion for each section is measured in seconds; higher scores indicate poorer performance. This measure was administered at Years 1 through 4 of the project.

Visual Memory

Digit Symbol Coding Recall: The WAIS-R Digit Symbol Substitution task (Wechsler, 1981) was used annually to assess perceptual processing speed as outlined previously. Following the 90 second coding portion of the task (during which participants marked the appropriate symbol below the associated number), participants were presented with a sheet containing the nine symbols and asked to recall the number that had been paired with the symbol. This number of items drawn correctly was used as a measure of incidental visual recall.

Global Functional Status

Basic and Instrumental Activities of Daily Living (ADLs)*:* Participants were asked to rate their level of difficulty with nine basic and instrumental activities of daily living (ADLs: i.e., walking across a room, bathing self, dressing self, getting up from a bed or chair, climbing stairs, walking several blocks, managing finances, performing household activities, and driving a car), on a scale from 0 to 2 (0 = no difficulty, 1 =

some difficulty, 2 = a lot of difficulty) (Rodgers & Miller, 1997). This measure was administered at Years 1 through 4, and again at Year 6. A total score was obtained by summing participants' responses across the nine activities. Higher scores indicate greater difficulty with ADLs.

Everyday Problem Solving

Everyday Problems Test (EPT)*:* The EPT (Willis & Marsiske, 1993) is a paper and pencil test of everyday cognitive ability administered at Years 1 through 4, and Year 6. This measure uses 21 printed stimuli designed to closely mimic items encountered in daily life (e.g., medication label, pay telephone instructions) and requires participants to solve problems pertaining to the stimuli. Test items cover major components of instrumental activities of daily living including medication use, meal preparation, telephone use, shopping, financial management, household management and transportation. Two of the 42 items in the original EPT were omitted from the final version as a substantial number of participants disputed interpretation of the stimuli. Possible scores ranged from 0 to 40. Higher scores indicate better performance.

Relationship of Benchmark Measures to Cognitive and Functional Outcome Measures

Table 3 presents the zero-order correlations between the baseline raw scores on the cognitive benchmark measures used for MCI Classification and the baseline performance on the cognitive and functional outcome measures. Correlations between the benchmark and outcome measures were generally in the small to medium range and all were in the expected direction (i.e., worse benchmark cognitive status was associated with lower performance on cognitive and functional outcome measures.)

Table 3. Correlations between Baseline Cognitive Benchmark Measures and Baseline Cognitive and Functional Outcome Measures

Outcome Measure	Perceptual Speed	Reasoning	Episodic Memory	Verbal Fluency	Vocabulary
Cognitive					
Mini-Mental State Exam	.20**	.33**	.27**	.19**	.21**
Trail Making Test – Part B	-.49**	-.51**	-.28**	-.25**	-.11
Digit Symbol Coding Recall	.43**	.33**	.34**	.22**	.12*
Functional					
Activities of Daily Living	-.35**	-.25**	-.24**	-.25**	-.12*
Everyday Problems Test	.31**	.58**	.42**	.33**	.50**

Note: * $p < .05$, ** $p < .01$

MCI Classification Procedures

Classification of cognitive group status was determined based on performance on the five cognitive benchmark tasks (i.e., perceptual speed, reasoning, episodic memory, verbal fluency, and vocabulary.) Normative data were drawn from an independent sample of adults aged 65 to 94 years recruited from the same population (The Victoria Longitudinal Study, Dixon & de Frias, 2004)[1]. The use of local normative data derived for all tasks on the same population is preferred to the use of other published normative data given the close demographic match of the local sample to the current sample and the ability to make more accurate comparisons across tasks. Normative data for perceptual speed, reasoning, verbal fluency and vocabulary tasks are based on data from 445

[1] We thank Dr. Roger Dixon for the use of data from the Victoria Longitudinal Study to establish norms for the classification of cognitive status of the present sample.

individuals (282 women, 163 men) and normative data for the episodic memory task are based on data from 194 individuals (125 women, 69 men).

Figure 1 is a schematic representation of the algorithm applied for classifying individuals according to Traditional MCI and Persistent MCI classification. As a starting point ("Basic" classification scheme), a relatively liberal operational definition of MCI based on criteria adapted from Levy (1994) was employed. That is, participants were classified as having a cognitive impairment if they obtained scores more than 1 SD below the mean of their age- and education-matched peers on any one of the cognitive benchmark tasks. For the Traditional MCI categorization, data from the Year 1 baseline cognitive assessment was considered. Individuals who demonstrated cognitive impairment at Year 1 were categorized as Traditional MCI and individuals who did not demonstrate impairment were categorized as $Controls_{1y}$. For the Persistent MCI categorization, data from the first two annual cognitive assessments were considered. Individuals who demonstrated cognitive impairment at each of the first two assessments (Year 1 and Year 2) were categorized as $Persistent_{2y}$ MCI, those who demonstrated impairment at one, but not both time points were categorized as Unstable, and those who did not demonstrate impairment at any point were categorized as $Controls_{2y}$.

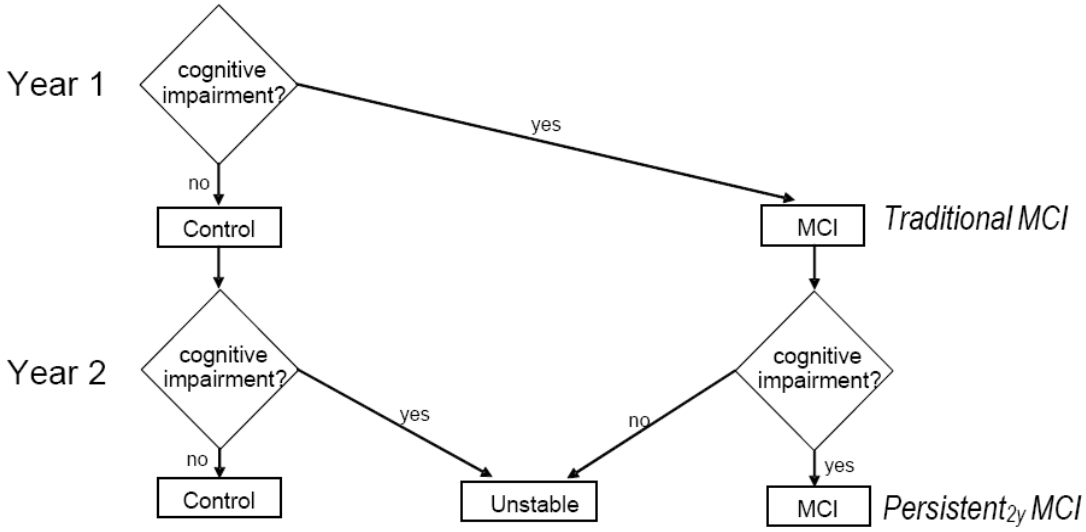


Figure 1. Classification of Traditional MCI and Persistent MCI (Basic variation)

The operational definition of MCI was varied systematically as outlined in hypothesis three to explore the robustness of findings associated with the traditional versus Persistent MCI categorization. Under the first variation ("Severity"), the same procedures as outlined in Figure 1 were employed, but a more stringent normative cutoff for impairment was applied at each annual assessment. That is, participants were classified as having a cognitive impairment if they obtained scores more than 1.5 SD below the mean of their age- and education-matched peers on any one of the cognitive benchmark tasks. Those who demonstrated cognitive impairment at baseline assessment were categorized as Traditional $MCI_{1.5\ SD}$ and individuals who demonstrated cognitive impairment at each of the first two assessments (Year 1 and Year 2) were categorized as $Persistent_{2y}\ MCI_{1.5\ SD}$.

Under the second variation ("Duration"), shown in Figure 2, an increased duration of impairment was required for inclusion in the Persistent MCI group. A third

assessment was added in the categorization of Persistent MCI. Individuals who demonstrated cognitive impairment at each of the first three assessments (Years 1, 2, and 3) were categorized as Persistent$_{3y}$ MCI, those who demonstrated impairment (using a 1 SD cutoff) at one, but not all time points were categorized as Unstable, and those who did not demonstrate impairment at any point were categorized as Controls.

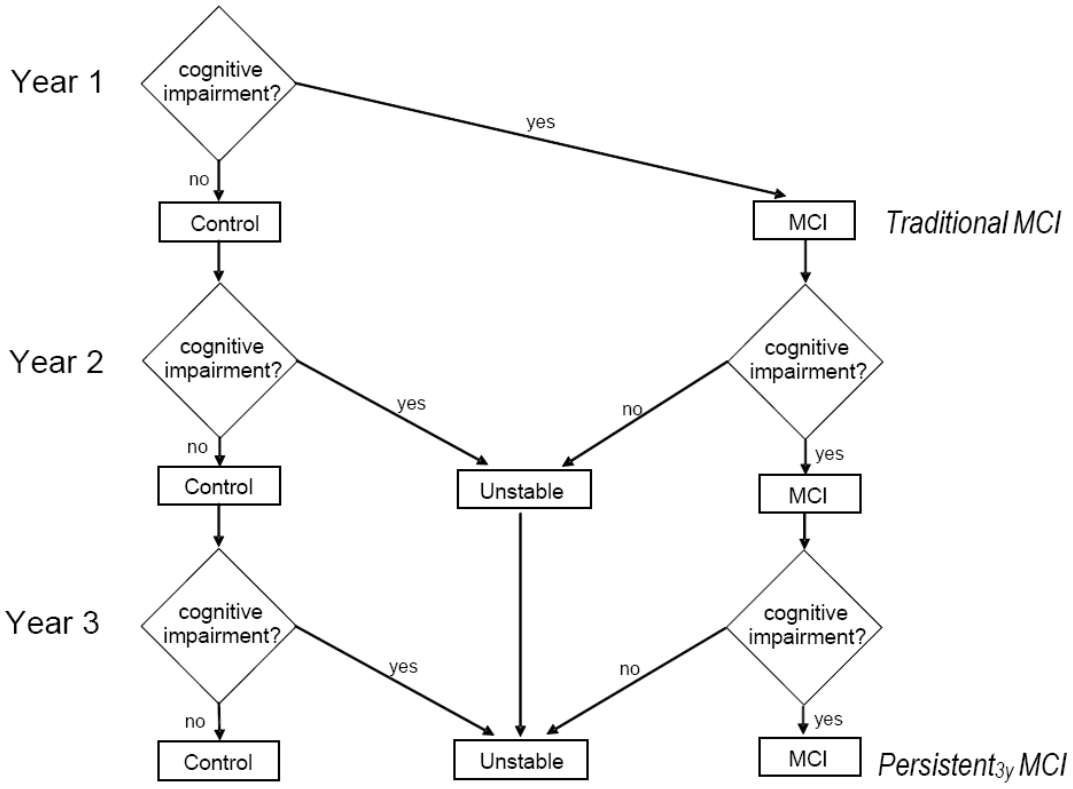


Figure 2. Classification of Traditional MCI and Persistent MCI, with the addition of a third cognitive assessment (Duration variation)

Under the third variation ("Pervasiveness"), outlined in Figure 3, the MCI groups were subdivided according to the pervasiveness of cognitive impairment. Individuals who demonstrated cognitive impairment (using a 1 SD cutoff) on multiple tests (i.e., two or more tests) at both of the first two assessments (Years 1 and 2) were categorized as

Persistent$_{2y}$ MCI-Multiple (P-MCI-M), and those who demonstrated impairment on at least one test at both of the first two assessments were categorized as Persistent$_{2y}$ MCI-Single (P-MCI-S). Those who demonstrated impairment on at least one test, but not at both time points were categorized as Unstable, and those who did not demonstrate impairment at any point were categorized as Controls.

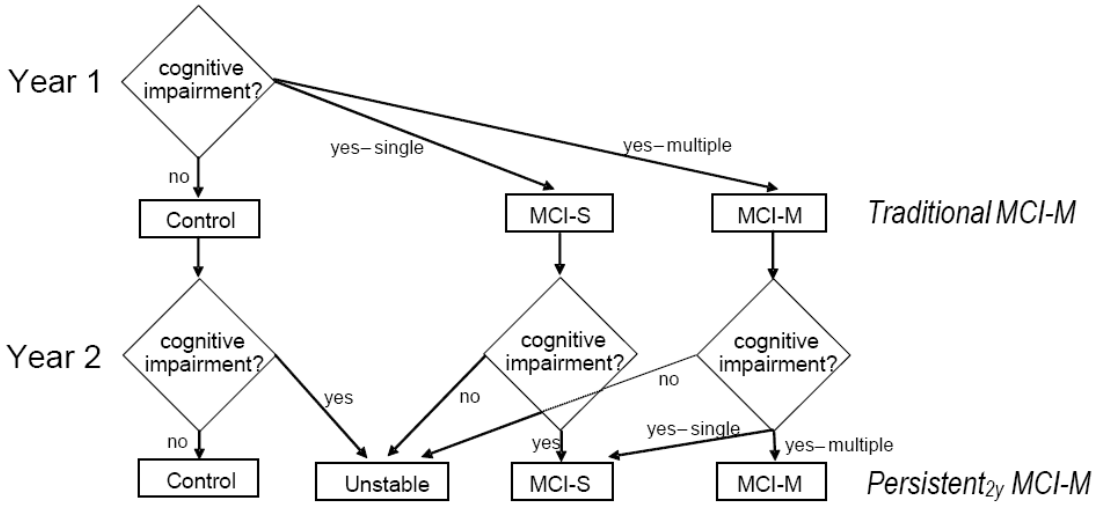


Figure 3. Classification of Traditional MCI and Persistent MCI, subdividing MCI by number of tests showing impairment (Pervasiveness variation)

Under the fourth variation ("Specificity"), outlined in Figure 4, the MCI groups were subdivided according to the specific domain of cognitive impairment. Individuals who demonstrated cognitive impairment (using a 1 SD cutoff) on the memory measure at both of the first two assessments (Years 1 and 2) were categorized as Persistent$_{2y}$ amnestic-MCI (P-aMCI), and those who demonstrated impairment on at least one non-memory measure at both of the first two assessments were categorized as Persistent$_{2y}$ nonamnestic-MCI (P-nMCI). Those who demonstrated impairment on at least one test,

but not at both time points were categorized as Unstable, and those who did not demonstrate impairment at any point were categorized as Controls.

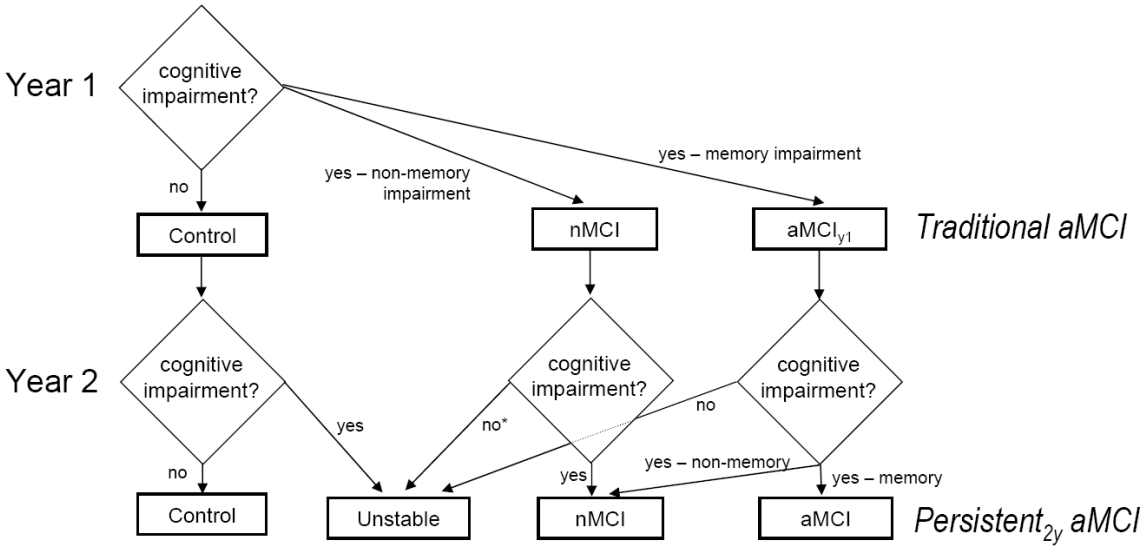


Figure 4. Classification of Traditional MCI and Persistent MCI, subdividing MCI by specific cognitive domain showing impairment (Specificity variation)

Results

Overview

Results are presented in three sections, corresponding with the three Study Objectives. The first section addresses the question of prevalence of Persistent MCI in comparison to Traditional MCI. The second section addresses the issue of the external validity of the Persistent$_{2y}$ MCI classification using five-year cognitive and functional outcome data. The third section addresses the robustness of the Persistent MCI classification by replicating analyses performed in the second section, varying elements of the operational definition of MCI in four ways.

Objective 1: Prevalence of Traditional MCI versus Persistent MCI

Analysis Plan

To address the hypotheses outlined in Objective 1, descriptive statistics are used to report base rates of Traditional MCI versus $Persistent_{2y}$ MCI (as outlined in Figure 1) and rates of conversion between the various groups. Next, base rates of Traditional and Persistent MCI, according to the four permutations of the classification scheme outlined in Objective 3 (as outlined in Figures 1, 2, 3, and 4) are reported. Between-group comparisons of key demographic and background measures are reported to provide a more complete understanding of the cognitive status groups.

Base Rates of Traditional versus $Persistent_{2y}$ MCI

Table 4 documents base rates of Traditional MCI versus $Persistent_{2y}$ MCI (according to the classification procedures outlined in Figure 1 and using a 1 SD cutoff for impairment). As expected, the base rate of Traditional MCI (N=144; 55% of the sample) was higher than that of $Persistent_{2y}$ MCI (N=97; 37% of the sample). At Year 2, the Unstable group (N=74) comprised a substantial 28% of the sample. An examination of the rates of conversion to the Unstable group revealed that a greater proportion of individuals from the Traditional MCI group converted to the Unstable group (N=47; 33% of the MCI group; 18% of the sample), compared to the proportion of individuals who converted from the $Control_{1y}$ group to the Unstable group (N=27; 23% of the Control group; 10% of the sample). This difference was statistically significant in terms of the overall sample (McNemar $X^2 = 4.878$, $p < 0.05$). That is, the proportion of the overall sample that showed impairment at Year 1, but not Year 2 was significantly greater than the proportion that did not show impairment at Year 1, but were impaired at Year 2.

Table 4. Prevalence of Traditional MCI versus Persistent$_{2y}$ MCI (Basic variation)

MCI Classification Algorithm	Control	Unstable	MCI
Year 1 - Traditional MCI	118 (45%)	not defined	144 (55%)
Year 2 - Persistent$_{2y}$ MCI	91 (35%)	74 (28%)	97 (37%)

Notes: Values are N's with percentage of original sample (N = 262) in parentheses. Rates of conversion to the Unstable group for the Year 1 Control and MCI groups were calculated by subtracting the Year 2 group sample size from the Year 1 group sample size and expressing this difference as a percentage of the Year 1 group sample size. For example, the rate of conversion from the Year 1 Traditional MCI group to the Year 2 Unstable group is equal to (144 – 97)/144 = 47/144 = 33%.

Table 5 provides a breakdown of key demographic and background variables by cognitive status group. The overall pattern of group differences was similar under both the Traditional and Persistent MCI classification. That is, the Control groups tended to be slightly younger, more educated, less depressed, healthier, and have better perceived memory ability. Only a small portion of these differences, however, were statistically significant. Under both classifications, the Control groups rated their own health as significantly closer to a perfect state of health than did the MCI groups. Under the Traditional MCI classification scheme, the Control group rated their own memory as higher relative to their peers than did the Traditional MCI group; under the Persistent MCI classification the Control group rated their own memory as higher relative to their peers than did the Unstable group. The P-MCI group had fewer years of education than both the Control and the Unstable group although it should be noted that their mean education was nonetheless quite high.

Table 5. Background Measures by Cognitive Status Group for Persistent$_{2y}$ MCI Classification (Basic variation)

Background Measure	Traditional MCI			Persistent$_{2y}$ MCI			
	Control	T-MCI	Contrasts	Control	Unstable	P-MCI	Contrasts
Age (years)	73.3 (5.3)	74.3 (6.3)		72.7 (5.0)	74.2 (5.8)	74.6 (6.6)	
Education (years)	15.5 (3.0)	14.9 (3.1)		15.5 (3.1)	15.7 (3.1)	14.5 (3.0)	C, U > P
Depressive Affect	7.9 (2.0)	8.2 (2.3)		7.8 (1.9)	7.8 (1.8)	8.4 (2.6)	
Number of Chronic Conditions	3.0 (2.3)	3.5 (2.5)		3.0 (2.4)	3.0 (2.2)	3.7 (2.7)	
SR Health cf Perfect	4.3 (0.6)	4.2 (0.7)	C > T	4.4 (0.6)	4.2 (0.6)	4.1 (0.7)	C > P
SR Health cf Peers	4.6 (0.6)	4.4 (0.6)		4.6 (0.6)	4.5 (0.6)	4.4 (0.6)	
SR Memory cf Previous Year	3.9 (0.8)	3.7 (1.1)		3.9 (0.7)	3.9 (0.9)	3.6 (1.1)	
SR Memory cf Peers	4.4 (0.9)	4.0 (0.9)	C > T	4.4 (0.9)	4.0 (0.9)	4.2 (1.1)	C > U
Gender (% Female)	70	68		69	70	67	

Note: Values are means with standard deviations in parentheses. Contrasts represent one way ANOVA (two group) and post hoc (three group) comparisons (LSD, $p < .05$). SR = Self-Rated, C = Control, T = Traditional MCI, U = Unstable, P = Persistent MCI.

Base Rates of MCI with Increased Severity Inclusion Criteria

Table 6 documents base rates of Traditional MCI and Persistent$_{2y}$ MCI according to the first variation in classification procedures outlined in Objective 3. In this case, a more stringent normative cutoff was applied (according to the classification procedures shown in Figure 1). Cognitive impairment was operationalized as a score greater than 1.5 SD below age and education referenced normative data on any one of five cognitive benchmark measures. As expected, the base rate of Traditional MCI$_{1.5SD}$ (N=77; 29% of the sample) was higher than that of Persistent$_{2y}$ MCI$_{1.5SD}$ (N=45; 17% of the sample). An examination of the rates of conversion to the Unstable group showed that the proportion

of individuals from the Traditional MCI group who converted to the Unstable group (N=32; 42% of the MCI group; 12% of the sample), did not significantly differ from the proportion of individuals who converted from the Control group to the Unstable group (N=23; 12% of the Control group; 9% of the sample) in terms of the overall sample (McNemar $X^2 = 1.164$, $p = 0.28$).

Table 6. Prevalence of Traditional MCI versus Persistent$_{2y}$ MCI$_{1.5\ SD}$ (Severity variation)

MCI Classification Algorithm	Control	Unstable	MCI
Year 1 - Traditional MCI $_{1.5\ SD}$	185 (71%)	not defined	77 (29%)
Year 2 - Persistent$_{2y}$ MCI $_{1.5\ SD}$	162 (62%)	55 (21%)	45 (17%)

Notes: Values are N's with percentage of original sample (N = 262) in parentheses. Rates of conversion to the Unstable group for the Year 1 Control and MCI groups were calculated by subtracting the Year 2 group sample size from the Year 1 group sample size and expressing this difference as a percentage of the Year 1 group sample size.

Table 7 provides a breakdown of key demographic and background variables by cognitive status group. Under the Traditional MCI classification, Controls differed from the T-MCI group in terms of education, self-rated health relative to a perfect state, and self-rated memory relative to one year ago (all lower in T-MCI), and number of chronic conditions (higher in T-MCI). Under the Persistent MCI classification, Controls differed from the Unstable and the P-MCI group in terms of education (lower in Unstable and P-MCI groups). Controls also rated their health as closer to a perfect state than did the P-MCI group.

Table 7. Background Measures by Cognitive Status Group for Persistent$_{2y}$ MCI$_{1.5SD}$ Classification (Severity variation)

Background Measure	Traditional MCI			Persistent$_{2y}$ MCI $_{1.5\ SD}$			
	Control	T-MCI	Contrasts	Control	Unstable	P-MCI	Contrasts
Age (years)	73.7 (5.8)	74.2 (6.0)		73.8 (5.8)	73.8 (5.9)	74.0 (6.4)	
Education (years)	15.5 (3.1)	14.3 (2.9)	C > T	15.6 (3.1)	14.5 (3.3)	14.4 (2.6)	C > U, P
Depressive Affect	7.9 (2.0)	8.3 (2.5)		8.0 (2.1)	8.0 (1.9)	8.4 (2.8)	
Number of Chronic Conditions	3.1 (2.4)	3.8 (2.6)	C < T	3.1 (2.4)	3.4 (2.6)	3.8 (2.5)	
SR Health cf Perfect	4.3 (0.6)	4.1 (0.7)	C > T	4.3 (0.6)	4.1 (0.6)	4.1 (0.8)	C > P
SR Health cf Peers	4.5 (0.6)	4.4 (0.6)		4.5 (0.6)	4.5 (0.5)	4.4 (0.7)	
SR Memory cf Previous Year	3.9 (0.8)	3.5 (1.2)	C > T	3.9 (0.8)	3.7 (1.0)	3.6 (1.3)	
SR Memory cf Peers	4.3 (0.9)	4.0 (1.2)		4.2 (0.9)	4.1 (1.0)	4.2 (1.2)	
Gender (% Female)	69	68		70	62	73	

Note: Values are means with standard deviations in parentheses. Contrasts represent one way ANOVA (two group) and post hoc (three group) comparisons (LSD, $p < .05$). SR = Self-Rated, C = Control, T = Traditional MCI, U = Unstable, P = Persistent MCI.

Base Rates of MCI with Increased Duration Inclusion Criteria

Table 8 documents base rates of Traditional MCI versus Persistent MCI when the criteria for Persistent MCI were altered to require increased duration (3 years) of impairment for inclusion in the Persistent MCI group (according to the classification procedures outlined in Figure 2). Data from Year 2 classification is also included in the table for reference. It is noted that 15 participants (6% of original sample; 3 from Year 1 Control group, 12 from Year 1 Traditional MCI group) left the study between Year 2 and 3, so the overall sample size under consideration for analyses in this section was 247.

Again, as expected, the base rate of Traditional MCI (N=132; 53% of the sample) was markedly higher than that of Persistent$_{3y}$ MCI (N=72; 29% of the sample). An examination of the rates of conversion to the Unstable group shows that a significantly greater proportion of individuals from the Traditional MCI group converted to the Unstable$_{3y}$ group (N=60; 45% of the MCI group; 24% of the sample), compared to the proportion of individuals who converted from the Control$_{1y}$ group to the Unstable$_{3y}$ group (N=36; 31% of the Control group; 15% of the sample) in terms of the overall sample (McNemar $X^2 = 5.470$, $p < 0.05$). Rates of change from Year 2 classification to Year 3 classification were also explored. Interestingly, the proportion of individuals who showed cognitive impairment at Year 2, but not Year 3 (N=29; 12% of the sample) did not significantly differ from the proportion of individuals who did not show impairment at Year 2, but did show impairment at Year 3 (N=28; 11% of the sample) in terms of the overall sample (McNemar $X^2 = 0.000$, $p = 1.00$).

Table 8. Prevalence of Traditional MCI versus Persistent$_{3y}$ MCI (Duration variation)

MCI Classification Algorithm	Control	Unstable	MCI
Year 1 - Traditional MCI	115 (47%)	-	132 (53%)
Year 2 - Persistent$_{2y}$ MCI	90 (36%)	69 (28%)	88 (36%)
Year 3 - Persistent$_{3y}$ MCI	79 (32%)	96 (39%)	72 (29%)

Notes: Fifteen participants (6% of original sample; 3 from Year 1 Control group, 12 from Year 1 Traditional MCI group) left the study between Year 2 and 3. Values in table are N's with percentage of the Year 3 sample (N = 247) in parentheses. Rates of conversion to the Unstable group for the Year 1 Control and MCI groups were calculated by subtracting the Year 3 group sample size from the Year 1 group sample size and expressing this difference as a percentage of the Year 1 group sample size, after accounting for attrition noted above.

Table 9 provides a breakdown of key demographic and background variables by cognitive status group. Under the Traditional MCI classification, Controls differed from the T-MCI group in terms of self-rated health relative to a perfect state, and self-rated memory relative to prior year (both lower in T-MCI). Under the Persistent MCI classification, Controls differed from the Unstable and the P-MCI group in terms of self-rated health relative to a perfect state (lower in Unstable and P-MCI groups). Controls differed from the P-MCI group in terms of number of chronic conditions (higher in the P-MCI group), self-rated health relative to same aged peers, self-rated memory relative to prior year (both lower in the P-MCI group). Controls differed from the Unstable group in terms of their self-rated memory relative to same aged peers (lower in Unstable group).

Table 9. Background Measures by Cognitive Status Group for Persistent$_{3y}$ MCI Classification (Duration variation)

Background Measure	Traditional MCI			Persistent$_{3y}$ MCI			
	Control	T-MCI	Contrasts	Control	Unstable	P-MCI	Contrasts
Age (years)	73.3 (5.3)	74.3 (6.3)		72.7 (5.0)	73.4 (5.4)	74.9 (6.6)	
Education (years)	15.5 (3.0)	14.9 (3.1)		15.3 (2.9)	15.6 (3.2)	14.6 (3.0)	
Depressive Affect	7.9 (2.0)	8.2 (2.3)		7.7 (1.9)	8.1 (1.9)	8.5 (2.8)	
Number of Chronic Conditions	3.0 (2.3)	3.5 (2.5)		2.8 (2.3)	3.2 (2.4)	3.9 (2.7)	C < P
SR Health cf Perfect	4.3 (0.6)	4.2 (0.7)	C > T	4.5 (0.6)	4.2 (0.6)	4.1 (0.7)	C > U, P
SR Health cf Peers	4.6 (0.6)	4.4 (0.6)		4.6 (0.5)	4.5 (0.6)	4.4 (0.6)	C > P
SR Memory cf Previous Year	3.9 (0.8)	3.7 (1.1)		3.9 (0.7)	3.8 (0.9)	3.5 (1.0)	C > P
SR Memory cf Peers	4.4 (0.9)	4.0 (0.9)	C > T	4.4 (1.0)	4.0 (0.9)	4.2 (1.0)	C > U
Gender (% Female)	70	68		72	67	69	

Note: Values are means with standard deviations in parentheses. Contrasts represent one way ANOVA (two group) and post hoc (three group) comparisons (LSD, $p < .05$). SR = Self-Rated, C = Control, T = Traditional MCI, U = Unstable, P = Persistent MCI.

Base Rates of MCI with Increased Pervasiveness Inclusion Criteria

Table 10 documents base rates of Traditional MCI versus Persistent MCI when the criteria for Persistent MCI were altered to subdivide the MCI groups according to number of cognitive domains showing impairment (according to the classification procedures outlined in Figure 3). As expected, the base rate of Traditional MCI-Single (N= 81; 31% of the sample) was higher than that of Persistent MCI-Single (N=59; 23% of the sample). Similarly, the base rate of Traditional MCI-Multiple (N=63; 24% of the sample) was higher than that of Persistent MCI-Multiple$_{2y}$ (N=38; 15% of the sample).

Examination of the rates of conversion within the Controls and the Traditional MCI-S groups showed a familiar pattern. That is, a greater proportion of individuals from the Traditional MCI-S group converted to the Unstable group (N=41; 51% of the MCI-S group; 16% of the sample), compared to the proportion of individuals who converted from the Control group to the Unstable group (N=27; 23% of the Control group; 10% of the sample) (McNemar $X^2 = 4.516$, $p < 0.05$). However, the proportion of individuals from the Traditional MCI-M group who converted to the Unstable group (N=6; 10% of the Traditional MCI-M group; 2% of the sample) did not differ from the proportion of the Control group who showed impairment in multiple cognitive domains at Year 2 (classified as Unstable; N=4, 3% of the Control group; 2% of the sample) (McNemar $X^2 = 0.400$, $p = 0.53$). Similarly, the proportion of individuals from the Traditional MCI-M group who converted to the Persistent MCI-S group (N=19; 30% of the Traditional MCI-M group; 7% of the sample) did not differ from the proportion of the Traditional MCI-S group who showed impairment in multiple cognitive domains at Year 2 (classified as Persistent MCI-S; N=14, 17% of the Traditional MCI-S group; 5% of the sample) (McNemar $X^2 = 0.485$, $p = 0.49$).

Globally, a comparison of the overall proportion of individuals across the sample who showed instability of cognitive impairment in a positive direction (i.e., showed less pervasive, or no impairment on Year 2 testing; N=66; 25% of the sample) relative to the overall proportion of individuals who showed instability of cognitive impairment in a negative direction (i.e., showed more pervasive impairment on Year 2 testing; N=41; 16% of the sample) was not statistically significant (McNemar-Bowker $X^2 = 6.220$, $p = 0.10$).

Table 10. Prevalence of Traditional MCI and Persistent MCI (Pervasiveness variation)

MCI Classification Algorithm	Control	Unstable	MCI-Single	MCI-Multiple
Year 1 - Traditional MCI	118 (45%)	-	81 (31%)	63 (24%)
Year 2 - Persistent$_{2y}$ MCI	91 (35%)	74 (28%)	59 (23%)	38 (15%)

Notes: Values are N's with percentage of original sample (N = 262) in parentheses. Rate of conversion to the Unstable group for the Year 1 Control group was calculated by subtracting the Year 2 group sample size from the Year 1 group sample size and expressing this difference as a percentage of the Year 1 group sample size. Rates of conversion to the Unstable group for the Year 1 MCI-Single and MCI-Multiple were calculated similarly, after accounting for the proportion of individuals who converted from the MCI-Multiple to the MCI-Single group (N=19) and individuals who converted from the MCI-Multiple group to the Unstable group (N=6).

Table 11 provides a breakdown of key demographic and background variables by cognitive status group. Under the Traditional MCI classification, Controls differed from the T-MCI-M group in terms of self-rated health relative to a perfect state, and self-rated memory relative to same aged peers (both lower in T-MCI-M). Under the Persistent MCI classification, both the Control and the Unstable group had more years of education than the P-MCI-M group.

Table 11. Background Measures by Cognitive Status Group for Persistent MCI Classification (Pervasiveness variation)

Background Measure	Traditional MCI (single vs multiple)				Persistent MCI_{2y} (single vs multiple)				
	Control	T-MCI-S	T-MCI-M	Contrasts	Control	Unstable	P-MCI-S	P-MCI-M	Contrasts
Age	73.3 (5.3)	73.6 (6.0)	75.1 (6.6)		72.7 (5.0)	74.2 (5.8)	74.1 (6.2)	75.5 (7.1)	
Education	15.5 (3.0)	15.3 (3.1)	14.5 (3.1)		15.5 (3.1)	15.7 (3.1)	14.7 (2.8)	14.2 (3.1)	C, U > PM
Depressive Affect	7.9 (2.0)	8.1 (2.0)	8.3 (2.7)		7.8 (1.9)	7.8 (1.8)	8.4 (2.5)	8.4 (2.8)	
# Chronic Conditions	3.0 (2.3)	3.2 (2.3)	3.9 (2.7)		3.0 (2.4)	3.0 (2.2)	3.8 (2.4)	3.7 (3.0)	
SR Health cf Perfect	4.3 (0.6)	4.2 (0.6)	4.1 (0.7)	C > TM	4.4 (0.6)	4.2 (0.6)	4.1 (0.7)	4.1 (0.7)	
SR Health cf Peers	4.6 (0.6)	4.4 (0.6)	4.4 (0.6)		4.6 (0.6)	4.5 (0.6)	4.4 (0.6)	4.4 (0.7)	
SR Memory cf Previous Year	3.9 (0.8)	3.7 (1.0)	3.7 (1.1)		3.9 (0.7)	3.9 (0.9)	3.7 (1.1)	3.6 (1.2)	
SR Memory cf Peers	4.4 (0.9)	4.0 (1.0)	4.1 (1.1)	C > TS	4.4 (0.9)	4.0 (0.9)	4.2 (1.0)	4.2 (1.2)	
Gender (% Female)	70	74	60		69	70	68	66	

Note: Values are means with standard deviations in parentheses. Contrasts represent ANOVA post hoc comparisons (LSD, $p < .05$). SR = Self-Rated, C = Control, T = Traditional MCI, U = Unstable, TS = Traditional MCI-Single, TM = Traditional MCI-Multiple, PS = Persistent MCI-Single, PM = Persistent MCI-Multiple,

Base Rates of MCI with Modified Specificity Inclusion Criteria

Table 12 documents base rates of Traditional MCI versus Persistent MCI when the criteria for Persistent MCI were altered to subdivide the MCI groups according to specific cognitive domain(s) showing impairment (according to the classification procedures outlined in Figure 4). As expected, the base rate of Traditional Nonamnestic-

MCI (nMCI; N=112; 43% of the sample) was higher than that of Persistent nMCI (N=87; 33% of the sample). Similarly, the base rate of Traditional Amnestic-MCI (aMCI, N=32; 12% of the sample) was higher than that of Persistent aMCI (N=10; 4% of the sample).

Examination of the rates of conversion to the Unstable group shows that the relative proportion of individuals from the Traditional nMCI group who converted to the Unstable group (N=41; 37% of the Traditional nMCI group; 16% of the sample), compared to the proportion of individuals from the Year 1 Control group who showed cognitive impairment in a non-memory domain at Year 2 (classified as Unstable; N=27; 23% of the Control group; 10% of the sample) was not significantly different (McNemar $X^2 = 3.409$, $p = 0.07$). Similarly, the relative proportion of individuals from the Traditional aMCI group who converted to the Unstable group (N=6; 19% of the Traditional aMCI group; 2% of the sample) compared to the proportion of individuals from the Year 1 Control group who showed a memory impairment at Year 2 evaluation (classified as Unstable; N=2, 2% of the Control group; <1% of the sample) was not significantly different (McNemar $X^2 = 2.000$, $p = 0.16$). Finally, the relative proportion of individuals from the Traditional aMCI group who converted to the Persistent nMCI group (N=16; 50% of the Traditional aMCI group; 6% of the sample) compared to the proportion of individuals from the Traditional nMCI group who showed a memory impairment at Year 2 evaluation (classified as Persistent nMCI; N=11, 10% of the Traditional nMCI group; 4% of the sample) was not significant different (McNemar $X^2 = 0.593$, $p = 0.44$).

Table 12. Prevalence of Traditional MCI and Persistent MCI (Specificity variation)

MCI Classification Algorithm	Control	Unstable	nMCI	aMCI
Traditional MCI	118 (45)		112 (43)	32 (12)
$Persistent_{2y}$ MCI	91 (35)	74 (28)	87 (33)	10 (4)

Notes: Values are N's with percentage of original sample (N = 262) in parentheses. nMCI = nonamnestic MCI, aMCI = amnestic MCI. Rate of conversion to the Unstable group for the Year 1 Control group was calculated by subtracting the Year 2 group sample size from the Year 1 group sample size and expressing this difference as a percentage of the Year 1 group sample size. Rates of conversion to the Unstable group for the Year 1 nMCI and aMCI were calculated similarly, after accounting for the proportion of individuals who converted from aMCI to the nMCI group (N=16) and individuals who converted from the aMCI to the Unstable group (N=6).

Table 13 provides a breakdown of key demographic and background variables by cognitive status group. Under the Traditional MCI classification, Controls rated their health (compared to perfect state and to peers) and memory (compared to peers) significantly higher than those in the Traditional Amnestic MCI group. Controls also rated their memory (compared to peers) higher than did the Traditional Nonamnestic MCI group. The Traditional Nonamnestic MCI group rated their health (compared to peers) as better than did the Traditional Amnestic MCI group. Under the Persistent MCI classification, the Controls rated their health (compared to a perfect state) as better than did both the Persistent Nonamnestic and Amnestic MCI groups. The Controls rated their memory (compared to peers) as better than did the Unstable group. Note that the small size of the Amnestic MCI groups was likely associated with reduced power to detect statistically significant group differences associated with this group, particularly under the Persistent MCI classification.

Table 13. Background Measures by Cognitive Status Group for Persistent MCI Classification (Specificity variation)

Background Measure	Traditional MCI (nonamnestic vs amnestic)				Persistent MCI_{2y} (nonamnestic vs amnestic)				
	Control	T-nMCI	T-aMCI	Contrasts	Control	Unstable	P-nMCI	P-aMCI	Contrasts
Age	73.3 (5.3)	74.4 (6.0)	73.8 (7.4)		72.7 (5.0)	74.2 (5.8)	74.6 (6.2)	74.8 (9.4)	
Education	15.5 (3.0)	15.0 (3.0)	14.6 (3.4)		15.5 (3.1)	15.7 (3.1)	14.5 (3.0)	15.0 (2.2)	
Depressive Affect	7.9 (2.0)	8.2 (2.3)	8.0 (2.2)		7.8 (1.9)	7.8 (1.8)	8.5 (2.7)	7.6 (0.7)	
# Chronic Conditions	3.0 (2.3)	3.5 (2.4)	3.4 (3.0)		3.0 (2.4)	3.0 (2.2)	3.9 (2.7)	2.7 (2.2)	
SR Health cf Perfect	4.3 (0.6)	4.2 (0.7)	4.0 (0.6)	C > Ta	4.4 (0.6)	4.2 (0.6)	4.1 (0.7)	3.8 (0.8)	C > Pn,Pa
SR Health cf Peers	4.6 (0.6)	4.5 (0.6)	4.2 (0.6)	C,Tn > Ta	4.6 (0.6)	4.5 (0.6)	4.4 (0.6)	4.1 (0.7)	
SR Memory cf Previous Year	3.9 (0.8)	3.8 (1.0)	3.5 (1.3)		3.9 (0.7)	3.9 (0.9)	3.7 (1.0)	3.5 (1.8)	
SR Memory cf Peers	4.4 (0.9)	4.1 (1.0)	4.0 (1.1)	C > Tn,Ta	4.4 (0.9)	4.0 (0.9)	4.2 (1.0)	3.9 (1.5)	C > U
Gender (% Female)	70	71	60		69	70	68	60	

Note: Values are means with standard deviations in parentheses. Contrasts represent ANOVA post hoc comparisons (LSD, $p < .05$). SR = Self-Rated, C = Control, T = Traditional MCI, nMCI = Nonamnestic MCI, aMCI = Amnestic MCI, U = Unstable, Tn = Traditional Nonamnestic-MCI-Single, Ta = Traditional Amnestic-MCI, Pn = Persistent Nonamnestic-MCI, Pa = Persistent amnestic MCI.

Section Summary

The first study objective was to replicate findings from a single prior study which employed a Persistent MCI classification scheme (Collie et al., 2002). It was hypothesized that base rates of impaired performance on any one cognitive test at any one time point in the current study (Traditional MCI) would exceed base rates of persistent

impaired performance observed on consecutive assessment sessions (Persistent MCI). This discrepancy was expected to hold across the five proposed operational definitions of Traditional and Persistent MCI. Results were consistent with this hypothesis and are summarized in Table 14. That is, regardless of which specific variation of Persistent versus Traditional MCI classification scheme was under consideration, the prevalence of Traditional MCI exceeded that of Persistent MCI. The proportion of individuals who met criteria for MCI at baseline (i.e., met criteria for Traditional MCI) but did not meet criteria for Persistent MCI ranged from 33-67%, depending on specific classification scheme under consideration.

Table 14. Summary of Prevalence Rates of Traditional versus Persistent MCI across Collie et al., 2002 and Current Study

MCI variation	Traditional MCI (% of sample)	Persistent MCI (% of sample)	Percent of T-MCI who did not meet P-MCI criteria
Collie et al., 2002	20	13	35
Basic	55	37	33
Severity	29	17	41
Duration	53	29	45
Pervasiveness [a]	24	15	40
Specificity [b]	12	4	67

Notes: N=262 for Basic, Severity, Pervasiveness, and Specificity variations, N=247 for Duration variation. [a] For Pervasiveness variation, MCI group in table is the MCI-Multiple group. [b] For Specificity variation, MCI group in table is the Amnestic-MCI group.

The nature of instability in classification rates revealed an overall trend for a reduced proportion of individuals showing impairment on re-evaluation. For example, in the first (Basic) MCI classification scheme, the overall proportion of individuals who showed impairment at Year 1, but not Year 2 was significantly greater than the

proportion of individuals who showed impairment at Year 2, but not Year 1. This finding was universally supported in terms of trend across the four remaining MCI classification scheme variations (i.e., Severity, Duration, Pervasiveness, and Specificity). Results were statistically significant under the Basic and Duration variations. Under the Pervasiveness variation, this finding was significant for the Control versus MCI-Single comparison, but not for either the Control versus MCI-Multiple comparison or the MCI-Single versus MCI-Multiple comparison. Results were not significant under either the Severity or the Specificity conditions.

Comparison of cognitive status group differences in demographic and background variables revealed an overall trend for the Control groups to be slightly younger, more educated, less depressed, healthier, and have better perceived memory ability. In general, the Unstable groups tended to perform at a level intermediate to the Control and the MCI groups in terms of background measures. Relatively few of these group differences were statistically significant. Across the five variations in MCI classification schemes, Control groups did tend to show significantly better self-rated health and self-rated memory relative to both the Traditional and the Persistent MCI groups, though this finding was not consistent. Under the Basic, Severity, and Pervasiveness variations, the Control groups had significantly higher average levels of education than did the Persistent MCI groups, as well as the the Traditional MCI group in the Severity variation.

Objective 2: External Validity of Persistent MCI Classification

The current section addresses the external validity of the Persistent$_{2y}$ MCI classification using five-year cognitive and functional outcome data.

Analysis Plan

Hierarchical Linear Modeling (HLM) was used to examine cognitive status group differences in baseline level and longitudinal trajectory of each cognitive (MMSE, Trails B, and Coding Recall) and functional (ADL and EPT) outcome measure. HLM provides a particularly appropriate means of analyzing longitudinal data because it can account for both within-person level and trajectories of performance over time and between-person influences (e.g., cognitive status group) on individual performance (Raudenbush & Bryk, 2002). In the case of longitudinal data, this task is accomplished by modeling observed variance at two levels: within-person and between-person. An added advantage of HLM is that such models, estimated according to full maximum likelihood procedures, offer a built-in means to preserve data from individuals who may have incomplete participation over the course of a longitudinal study by assigning increased weight to those with no missing data when calculating parameter estimates.

For the current study, a two-level HLM model was used. At the first level (within-person), each person's data was fit to an individual linear growth curve in the form:

$$\textit{Level 1} \qquad Y_{ij} = \beta_{0i} + \beta_{1i}(X_{ij}) + r_{ij}$$

where Y_{ij} is the value of the outcome variable of interest for person *i* at time *j*. The parameter β_{0i} represents the value of Y at an individual's intercept, or starting point. The parameter β_{1i} represents the rate of change for an individual per one unit increase in the independent (X_{ij}) variable. As a primary item of interest in the current study is change in outcome measures over time, the independent (X_{ij}) variable is used to model the effect of time, which could be operationalized in a number of ways (e.g., testing wave, time in

study, participant age, etc.). The parameter r_{ij} is an error term representing unexplained within-person variance.

At the second level (between-person), parameters in the Level 1 individual growth curve equations were specified as a function of sample level predictors in the form:

Level 2 $\beta_{0i} = \gamma_{00} + u_{0i}$

$\beta_{1i} = \gamma_{10} + u_{1i}$

where γ_{00} is the average intercept for the sample and γ_{10} is the average slope for the sample. The parameters u_0 and u_1 are error terms representing unexplained between-person variance in their associated parameters.

The above HLM equations were used as a reference point in the analyses that follow. That is, for each outcome measure, Level 1 and Level 2 models in the form above were fit to provide estimates of the sample-wide average baseline performance and rate of change over time. The next step in the analyses was to fit more complex Level 2 models in order to determine whether individuals in the sample varied in their baseline performance point and rate of change as a function of their cognitive status. Prior to moving onto this next step, preliminary investigations were conducted in order to determine the optimal metric for time in the HLM models.

Determination of Optimal Time Metric: Prior to the investigation of cognitive status differences in outcome measures of interest, preliminary analyses were conducted to determine the optimal metric for the time variable in individual growth curves. Unlike Repeated Measures ANOVA models, which assume equal time intervals between measurement occasions across subjects, HLM can accommodate varying time intervals between measurement occasions by subject. That is, HLM allows specification of a time

variable at an individual level, which allows individual differences in test-retest interval to be accounted for.

Exploratory analyses were undertaken to determine the optimal time metric for the planned HLM analyses. Three options were available with the current data. First, a simple "Study Year" variable could be used to model outcome as a function of study year (e.g., 0=Year 1, 1=Year 2, 2=Year 3, etc.). Second, a "Time in Study" variable could be used to model outcome as a function of each individual's test re-test interval (e.g., 0=Year 1 testing, or baseline; 1.2=Year 2 testing, 1.2 years following baseline; 1.9=Year 2, 1.9 years following baseline, etc.). Third, an "Age" variable could be used to model outcome as a function of each individual's exact age at each testing occasion. To ensure meaningful interpretation of intercept coefficients, the time variables were centered at the lowest available time point. In the case of age, this meant that a value of 64.0 (the age of the youngest participant at baseline) was subtracted from each individual's age at each time point.

To illustrate the implications of these alternate time metrics visually, the following three figures show individual Level 1 regression equations, using MMSE as an outcome variable, under each of the three different available time metrics. Each line on a graph represents the Level 1 regression equation for one study participant. In Figure 5, Study Year is used as an index of time. In Figure 6, Time in Study is used as an index of time. In Figure 7, Age is used as an index of time.

A simple visual comparison of the Study Year (Figure 5) and the Time in Study (Figure 6) figures is informative. The differences in lengths of the individual regression lines in the Time in Study figure demonstrate that individuals varied in the test-retest

interval from baseline to study end. (Individuals also varied considerably in their annual test-retest intervals.) It is plausible, then, that the Time in Study metric may more accurately capture individual's trajectory of change in MMSE score over time than does the simple Study Year metric.

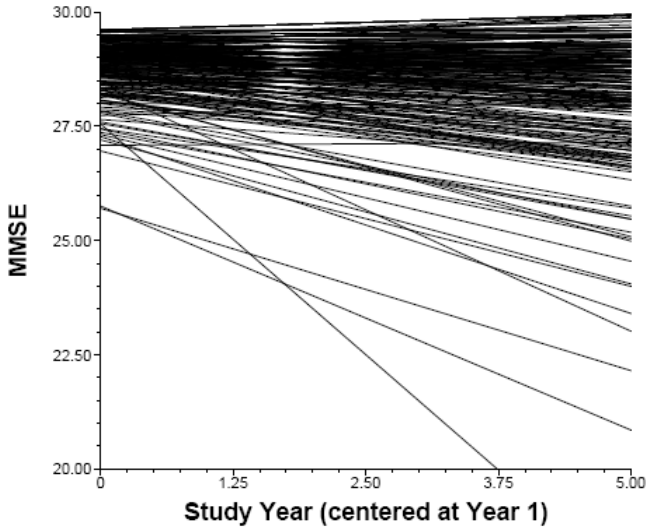


Figure 5. HLM Level 1 MMSE equations as a function of Study Year

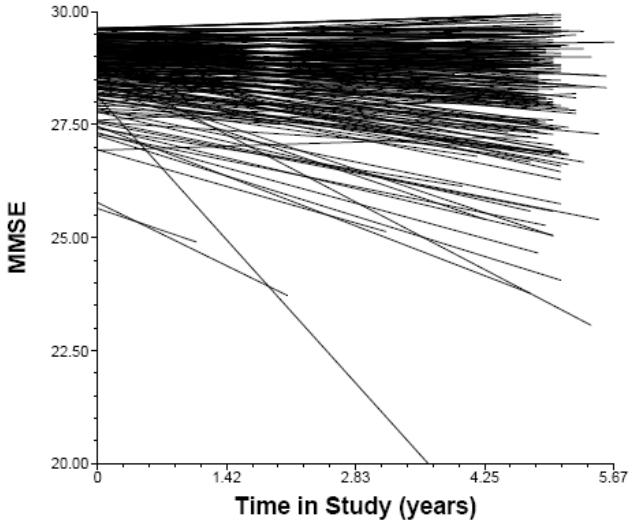


Figure 6. HLM Level 1 MMSE equations as a function of Time in Study

The Age figure (Figure 7) is a compelling demonstration of variability in age across the sample relative to the length of longitudinal follow-up across the study. That is, the difference in years between the age of the youngest participant at the first time point and the age of the oldest participant at the latest time point is approximately 33 years. An examination of Figure 7 shows that, relative to the 33 year overall span of ages in the sample, the 5 year study duration is relatively short. Thus, the Age time metric offers the potential to expand investigation of observed effects over a 33-year age span, rather than the observed 5-year span of the study. However, this approach decreases the number of data points available at each time point in the observed scale, and consequently decreases power to detect reliable effects in the data.

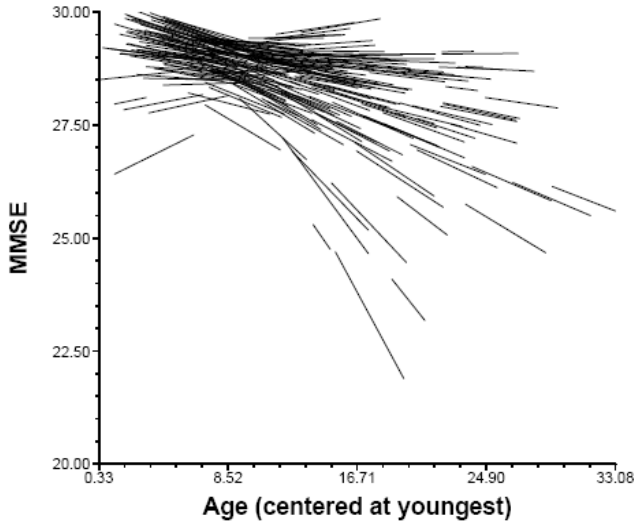


Figure 7. HLM Level 1 MMSE equations as a function of Age

Final selection of optimal time metric was completed by comparing model fit parameters for the three models shown above, as well as for identical sets of models run for each outcome measure (not shown). For the majority of outcome measures, lowest Akaike's Information Criteria (AIC) values (indicating best fitting models) were associated with the Age time metric. However, for a number of outcome variables the reliability estimates for the Level 1 coefficient were unacceptably low in the Age models. The Time in Study models, on the other hand, did not present the problem of low Level 1 reliability estimates and had consistently lower AIC values than the Study Year models. Thus, the decision was made to use Time in Study as the time metric for all planned Level 1 equations.

Investigation of Group Differences in Baseline Status and Rates of Change: As a next step, to determine whether individuals varied in their starting point and rate of change as a function of their cognitive status, new HLM models were fit in which a cognitive status group variable was added to the Level 2 equations as follows:

Level 2 $\beta_{0i} = \gamma_{00} + \gamma_{01}$ (Cognitive Status Group) + u_{0i}

$\beta_{1i} = \gamma_{10} + \gamma_{11}$ (Cognitive Status Group) + u_{1i}

The Cognitive Status Group variable was dummy coded to allow meaningful interpretation of estimated parameters. For example, in the evaluation of starting point and trajectory of change in an outcome variable for the Control versus Traditional MCI classification, the Control group was coded "0" and the Traditional MCI group was coded "1." Therefore, the parameter estimate for γ_{00} can be interpreted as the average performance at baseline (intercept) for individuals in the Control group while the parameter estimate for γ_{01} can be interpreted as the difference between average performances at baseline (intercept) for the Traditional MCI group compared to the Control group. Similarly, the parameter γ_{10} represents the average rate of change per unit increase in time for the Control group while the parameter γ_{11} represents the difference between average rates of change per unit increase in time for the Traditional MCI group compared to the Control group.

Final HLM Models

To test hypotheses associated with study Objective 2, a series of HLM models were run for each outcome measure. As outlined previously, in each case an initial, "simple" model with a single within-person Time in Study predictor was fit in order to estimate the sample-wide average intercept and slope for the outcome variable of interest:

Simple Model: Time in Study

Level 1 $Y = \beta_0 + \beta_1 (\text{Time in Study}) + r$

Level 2 $\beta_0 = \gamma_{00} + u_0$

$\beta_1 = \gamma_{10} + u_1$

Two separate "complex" models were then fit for each outcome variable. In the first complex model, a between-subjects cognitive status group variable was fit capturing individual membership in the Control versus the Traditional MCI grouping.

Complex Model 1: Traditional MCI

Level 1 $Y = \beta_0 + \beta_1 (\text{Time in Study}) + r$

Level 2 $\beta_{0i} = \gamma_{00} + \gamma_{01} (\text{Ctrl. vs. T-MCI}) + u_{0i}$

$\beta_{1i} = \gamma_{10} + \gamma_{11} (\text{Ctrl. vs. T-MCI}) + u_{1i}$

The "Ctrl vs. T-MCI" variable was dummy coded, as outlined previously, to allow meaningful interpretation of estimated parameters.

In the second complex model, a between-subjects cognitive status group variable was fit capturing individual membership in the Control versus the Persistent_{2y} MCI grouping. In order to dummy code the three level Persistent_{2y} cognitive status group variable (e.g., Control, Unstable, Persistent_{2y} MCI), this variable was modeled as two

separate dichotomous variables (shown below) in order to capture all possible contrasts of interest:

Complex Model 2: Persistent$_{2y}$ MCI

Level 1 $Y = \beta_0 + \beta_1 (\text{Time in Study}) + r$

Level 2 $\beta_0 = \gamma_{00} + \gamma_{01} (\text{Ctrl. vs. Unst.}) + \gamma_{02} (\text{Ctrl. vs. } P_{2y}\text{-MCI}) + u_0$

$\beta_1 = \gamma_{10} + \gamma_{11} (\text{Ctrl. vs. Unst.}) + \gamma_{12} (\text{Ctrl. vs. } P_{2y}\text{-MCI}) + u_1$

Model Comparisons: The simple Time in Study model was used as a benchmark against which the two complex models were compared. The two complex models were then compared to each other. Although these comparisons do not provide a direct test of the hypotheses of interest, it was expected that the two complex models would both provide an overall improved accounting of observed variance (i.e., yield better overall model fit parameters as described below) compared to the simple model. Further, it was expected that the Persistent$_{2y}$ MCI model would yield better overall model fit parameters compared to the Traditional MCI model. Because these comparisons involved a mix of nested and non-nested model comparisons, Akaike's Information Criteria (AIC) was used to compare model fit. The AIC is calculated by adding the deviance value associated with a given model to the number of parameters estimated in the model multiplied by two. A lower AIC can be interpreted as a better fitting model. The AIC thus penalizes complex models and can be used to evaluate whether any improvement in overall model fit associated with a more complex model is of a magnitude beyond that which would be expected from simply adding more parameters to the model.

Parameter Estimates and Group Comparisons: Parameter estimates for the complex model Level 2 equations were used to determine average baseline performance

and trajectory of change in outcome measures by cognitive status group. These values were also used to test hypotheses of interest for the current Study Objective. That is, to evaluate Hypothesis 2a, statistical comparisons of discrepancies observed between baseline level of performance and trajectory of decline in the (a) Control versus Traditional MCI and (b) Control versus $Persistent_{2y}$ MCI groupings were undertaken and are reported in the following section. Evidence for Hypothesis 2b is restricted to descriptive comparison of the magnitude of discrepancy of baseline performance level and trajectory of decline relative to Controls in the Traditional MCI versus the $Persistent_{2y}$ MCI groupings, as there is no available statistical test to compare model parameters across models with non-mutually exclusive groups.

Results for Cognitive Outcome Measures

Global Cognitive Functioning (MMSE): In terms of overall model fit, the simple Time in Study model yielded an AIC value of 4853.70. Both the Traditional MCI model (AIC = 4827.13) and the Persistent MCI model (AIC = 4816.01) had lower AIC values, indicating better model fit in the MCI models relative to the Time in Study model. The Persistent MCI model had a lower AIC value than the Traditional MCI model, indicating best fit of individual performance curves in the Persistent MCI model.

Parameter estimates for HLM Level 2 equations were used to derive average baseline scores and rates of change over time for each cognitive status group. These data are shown graphically in Figure 8. In the Traditional MCI model (Figure 8, left) the Control group had an average baseline MMSE score of 29.04 and an average rate of decline of -0.05 points per year. The Traditional MCI group had an average baseline score of 28.66 and an average rate of decline of -0.19 points per year. In the $Persistent_{2y}$

MCI model (Figure 8, right), the Control group had an average baseline MMSE score of 29.09 and an average rate of decline of -0.03 points per year. The Persistent MCI group had an average baseline score of 28.53 and an average rate of decline of -0.22 points per year. The Unstable group had an average baseline score of 28.90 and an average rate of decline of -0.12 points per year. For reference (data not shown), the sample-wide average baseline score was 28.83 (SE = 0.07), with an average rate of decline of -0.13 points per year (SE = 0.02, t = -5.06, $p < .001$).

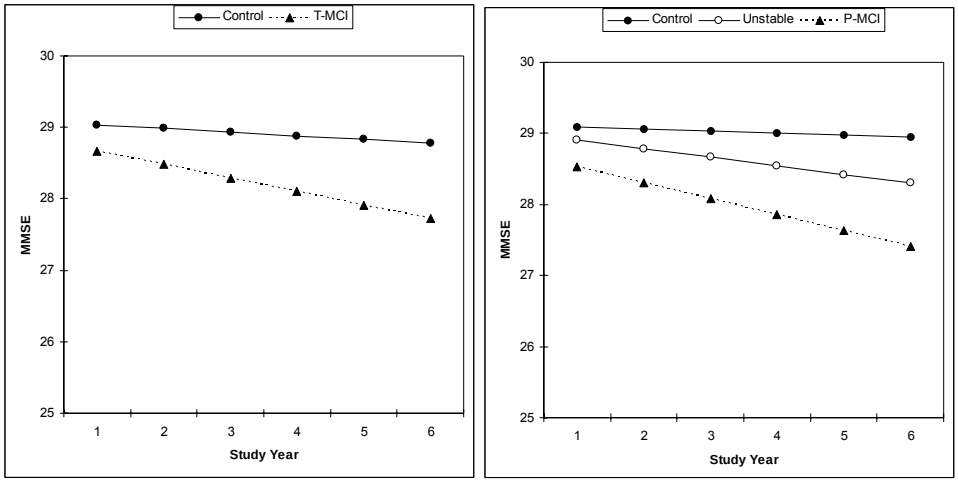


Figure 8. Group average MMSE performance across study for Traditional MCI (left) and Persistent$_{2y}$ MCI (right) classification schemes

Table 15 presents statistical comparisons of the magnitude of baseline and slope discrepancies between the cognitive status groups in the Traditional MCI model and the Persistent MCI model. Consistent with the hypothesis 2a, these comparisons show that both the Traditional MCI and the Persistent MCI groups had significantly lower baseline MMSE scores and greater longitudinal decline in MMSE scores relative to the Control groups. Although no specific hypotheses were made regarding performance of the

Unstable group in the Persistent MCI model, performance of this group, relative to the Control and Persistent MCI groups, is also presented in the table.

Table 15. Group Differences in MMSE Baseline Score and Rate of Change for Traditional MCI and Persistent$_{2y}$ MCI Models

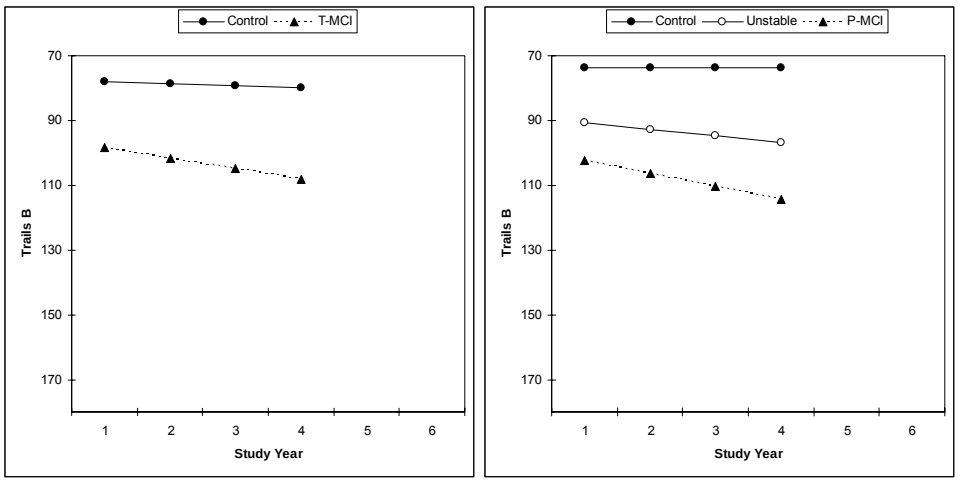
HLM Model	Baseline Difference	SE	*t*	Slope Difference	SE	*t*
Traditional MCI						
Ctrl vs. T-MCI	-.37	.14	-2.76 **	-.14	.05	-2.81 **
Persistent$_{2y}$ MCI						
Ctrl vs. P-MCI	-.56	.16	-3.55 **	-.19	.06	-3.42 **
Ctrl vs. Unst	-.19	.17	-1.13	-.09	.06	-1.51
Unst vs. P-MCI	-.37	.17	-2.22 *	-.10	.06	-1.71

Note: ** $p < .01$. * $p < .05$. Higher scores on MMSE indicate better performance. Negative baseline difference values indicate poorer baseline performance in the second group of comparison. Negative slope difference values indicate poorer trajectory of performance in second group of comparison.

Inspection of the values in the above table shows that, consistent with hypothesis 2b, the raw absolute magnitude of the discrepancy in both baseline performance and longitudinal rate of decline was larger for the Persistent MCI versus Control comparison than for the Traditional MCI versus Control comparison.

Executive Functioning (Trails B): Both the Traditional MCI model (AIC = 9635.15) and the Persistent MCI model (AIC = 9628.21) had lower AIC values compared to the simple Time in Study model (AIC = 9660.34). The Persistent MCI model had a lower AIC value than the Traditional MCI model, indicating best fit of individual performance curves for the Persistent MCI model.

Figure 9 shows average baseline scores and rates of change in Trails B performance for each cognitive status group. In the Traditional MCI model (left) the Control group had an average baseline score of 77.94 seconds and an average rate of decline of 0.60 seconds per year. The Traditional MCI group had an average baseline score of 98.30 seconds and an average rate of decline of 3.28 seconds per year. In the Persistent MCI model (right), the Control group had an average baseline score of 73.83 seconds and an average rate of change of 0.00 seconds per year. The Persistent MCI group had an average baseline score of 102.35 seconds and an average rate of decline of 6.25 seconds per year. The Unstable group had an average baseline score of 90.66 seconds and an average rate of decline of 2.06 seconds per year. For reference (data not shown in figure), the sample-wide average baseline score was 89.13 seconds (SE = 2.19), with an average rate of decline of 2.07 seconds per year (SE = 0.79, $t = 2.61$, $p < .01$).



Note: Y axes are reversed such that good performance appears at the top of the figure and poor performance appears at bottom of figure; negative slopes indicate decline over time.

Figure 9. Group average Trails B performance across study for Traditional MCI (left) and Persistent$_{2y}$ MCI (right) classification scheme

Table 16 presents comparisons of the magnitude of baseline and slope discrepancies between the cognitive status groups in the Traditional MCI model and the Persistent MCI model. These results were partially consistent with hypothesis 2a, in that both the Traditional MCI and the Persistent MCI groups had significantly lower baseline Trails B scores relative to the Control groups. The Persistent$_{2y}$ MCI group showed a significantly greater rate of decline compared to the Control group. The difference between the rate of change in the Traditional MCI compared to the Control group, on the other hand, was not statistically significant.

Table 16. Group Differences in Trails B Baseline Score and Rate of Change for Traditional MCI and Persistent$_{2y}$ MCI Models

HLM Model	Baseline Difference	SE	*t*	Slope Difference	SE	*t*
Traditional MCI						
Ctrl vs. T-MCI	20.36	4.21	4.84 **	2.69	1.58	1.70
Persistent$_{2y}$ MCI						
Ctrl vs. P-MCI	28.52	4.86	5.87 **	4.03	1.86	2.16 *
Ctrl vs. Unst	16.83	5.20	3.24 **	2.06	1.97	1.05
Unst vs. P-MCI	11.69	5.14	2.28 *	1.96	1.99	.99

Note: ** $p < .01$. * $p < .05$. Higher scores on Trails B indicate poorer performance. Positive baseline difference values indicate poorer performance in the second group listed in comparison. Positive slope difference values indicate poorer trajectory of performance over time in the second group.

As with the MMSE models, and consistent with hypothesis 2b, the absolute magnitude of the discrepancy in both baseline performance and longitudinal rate of decline was larger for the Persistent MCI versus Control comparison than for the Traditional MCI versus Control comparison.

Visual Memory (Coding Recall): Both the Traditional MCI model (AIC = 5117.05) and the Persistent MCI model (AIC = 5102.24) had lower AIC values compared to the simple Time in Study model (AIC = 5140.71, 6 estimated parameters). The Persistent MCI model had a lower AIC value than the Traditional MCI model.

Figure 10 shows average baseline scores and rates of change on the Coding Recall measure for each cognitive status group. In the Traditional MCI model (left) the Control group had an average baseline score of 6.97 and an average rate of change of 0.05 points per year (positive values indicate improved scores over time). The Traditional MCI group had an average baseline score of 5.67 and an average rate of change of 0.00 points per year. In the Persistent MCI model (right), the Control group had an average baseline score of 7.18 and an average rate of change of 0.10 points per year. The Persistent MCI group had an average baseline score of 5.46 and an average rate of decline of -0.03 points per year. The Unstable group had an average baseline score of 6.16 and an average rate of decline of -0.02 points per year. For reference (data not shown in figure), the sample-wide average baseline score was 6.25 (SE = 0.13), with an average rate of change of 0.02 points per year (SE = 0.03, $t = 0.60$, $p = 0.55$).

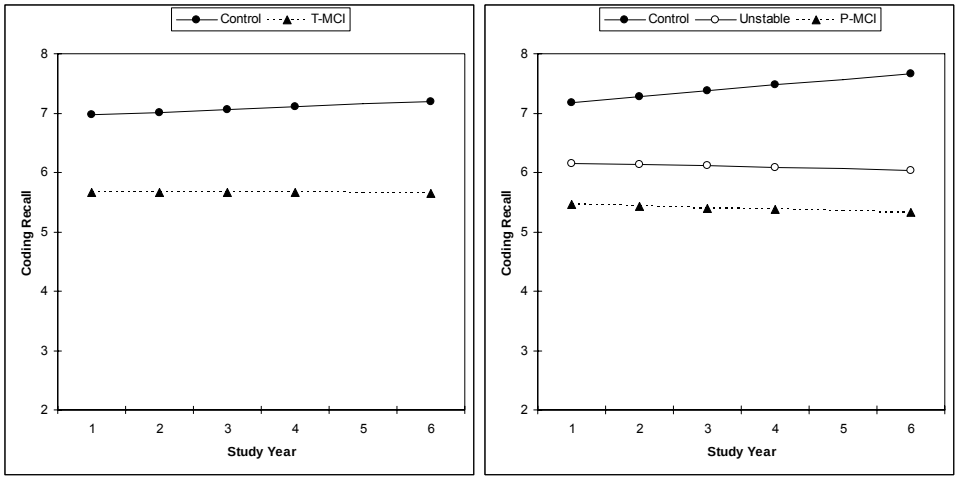


Figure 10. Group average Coding Recall performance across study for Traditional MCI (left) and Persistent$_{2y}$ MCI (right) classification scheme

Table 17 presents comparisons of the magnitude of baseline and slope discrepancies between the cognitive status groups in the Traditional MCI model and the Persistent MCI model. These results were partially consistent with the hypothesis 2a, in that both the Traditional MCI and the Persistent MCI groups had significantly lower baseline Coding Recall scores relative to the Control groups. Neither the Traditional MCI nor the Persistent MCI group showed a significantly greater rate of decline compared to the Control group.

Table 17. Group Differences in Coding Recall Baseline Score and Rate of Change for Traditional MCI and Persistent$_{2y}$ MCI Models

HLM Model	Baseline Difference	SE	*t*	Slope Difference	SE	*t*
Traditional MCI						
Ctrl vs. T-MCI	-1.30	.26	-5.07 **	-.05	.07	-.70
Persistent$_{2y}$ MCI						
Ctrl vs. P-MCI	-1.73	.30	-5.77 **	-.12	.08	-1.52
Ctrl vs. Unst	-1.02	.32	-3.21 **	-.12	.09	-1.42
Unst vs. P-MCI	-.70	.32	-2.22 *	.00	.09	-.04

Note: ** $p < .01$. * $p < .05$. Higher scores on Coding Recall indicate better performance. Negative baseline difference values indicate poorer baseline performance in the second group of comparison. Negative slope difference values indicate poorer trajectory of performance in second group of comparison.

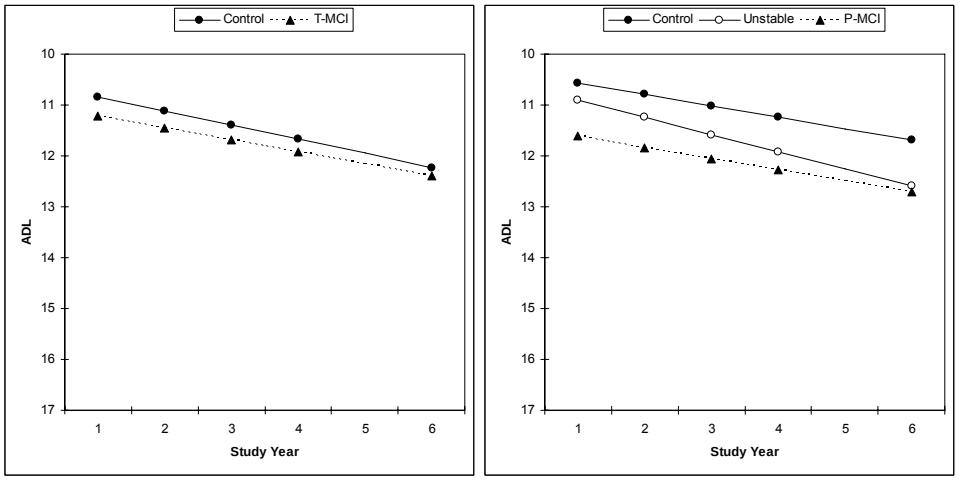
Again, consistent with hypothesis 2b, the absolute magnitude of the discrepancy in baseline performance was larger for the Persistent MCI versus Control comparison than for the Traditional MCI versus Control comparison. As neither the Traditional MCI nor the Persistent MCI groups significantly differed from Controls in terms of their rate of change over time, the relative magnitude of these differences was not evaluated.

Results for Functional Outcome Measures

Global Functional Status (ADL): Both the Traditional MCI model (AIC = 4929.07) and the Persistent MCI model (AIC = 4923.84) had lower AIC values compared to the simple Time in Study model (AIC = 4932.97). The Persistent MCI model had a lower AIC value than the Traditional MCI model.

Figure 11 shows average baseline scores and rates of change on the ADL measure for each cognitive status group. In the Traditional MCI model (left), the Control group

had an average baseline score of 10.34 (higher scores indicate poorer performance) and an average rate of decline of 0.28 points per year (positive values indicate decline over time). The Traditional MCI group had an average baseline score of 10.72 and an average rate of change of 0.23 points per year. In the Persistent MCI model (right), the Control group had an average baseline score of 10.07 and an average rate of change of 0.22 points per year. The Persistent MCI group had an average baseline score of 11.12 and an average rate of decline of 0.22 points per year. The Unstable group had an average baseline score of 10.41 and an average rate of decline of 0.34 points per year. For reference (data not shown in figure), the sample-wide average baseline score was 10.55 (SE = 0.15), with an average rate of decline of 0.25 points per year (SE = 0.03, $t = 7.52$, $p < .001$).



Note: Y axes are reversed such that good performance appears at the top of the figure and poor performance appears at bottom of figure; negative slopes indicate decline over time.

Figure 11. Group average ADL performance across study for Traditional MCI (left) and Persistent$_{2y}$ MCI (right) classification scheme

Table 18 presents comparisons of the magnitude of baseline and slope discrepancies between the cognitive status groups in the Traditional MCI model and the

Persistent MCI model. In this case, only the Persistent MCI group had a significantly lower baseline ADL performance relative to the Control group. The Traditional MCI group did not differ from Controls in terms of baseline performance or trajectory of change over time. The Persistent MCI group did not differ from Controls in terms of trajectory of change.

Table 18. Group Differences in ADL Baseline Score and Rate of Change for Traditional MCI and Persistent$_{2y}$ MCI Models

HLM Model	Baseline Difference	SE	*t*	Slope Difference	SE	*t*
Traditional MCI						
Ctrl vs. T-MCI	.38	.31	1.25	-.04	.07	-.63
Persistent$_{2y}$ MCI						
Ctrl vs. P-MCI	1.05	.36	2.94 **	-.01	.08	-.08
Ctrl vs. Unst	.34	.38	.90	.11	.08	1.32
Unst vs. P-MCI	.71	.38	1.87	-.12	.09	-1.38

Note: ** $p < .01$. * $p < .05$. Higher scores on ADL indicate poorer performance. Positive baseline difference values indicate poorer performance in the second group of comparison. Positive slope difference values indicate poorer trajectory of performance in second group of comparison.

Again, consistent with hypothesis 2b, the absolute magnitude of the discrepancy in baseline performance was larger for the Persistent MCI versus Control comparison than for the Traditional MCI versus Control comparison. As neither the Traditional MCI nor the Persistent MCI groups differed from Controls in terms of their rate of change over time, the relative magnitude of these differences was not evaluated.

Applied Problem Solving (EPT): In the case of the EPT measure, both the Traditional MCI model (AIC = 6318.26) and the Persistent MCI model (AIC = 6301.03) had lower AIC values compared to the simple Time in Study model (AIC = 6357.67). Similarly, the Persistent MCI model had a lower AIC value than the Traditional MCI model.

Figure 12 shows average baseline scores and rates of change over time for each cognitive status group. In the Traditional MCI model (left), the Control group had an average baseline score of 33.39 and an average rate of decline of -0.14 points per year. The Traditional MCI group had an average baseline score of 29.50 and an average rate of decline of -0.27 points per year. In the Persistent MCI model (right), the Control group had an average baseline score of 33.81 and an average rate of decline of -0.09 points per year. The Persistent MCI group had an average baseline score of 28.35 and an average rate of decline of -0.30 points per year. The Unstable group had an average baseline score of 31.91 and an average rate of decline of -0.23 points per year. For reference (data not shown in figure), the sample-wide average baseline score was 31.25 (SE = 0.35), with an average rate of decline of -0.21 points per year (SE = 0.05, $t = -3.87$, $p < .001$).

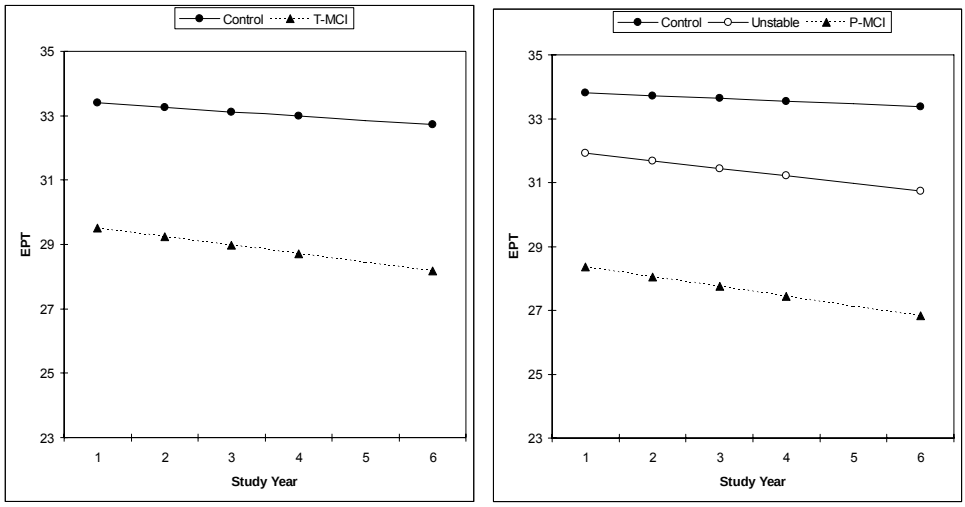


Figure 12. Group average EPT performance across study for Traditional MCI (left) and Persistent$_{2y}$ MCI (right) classification scheme

Table 19 presents comparisons of the magnitude of baseline and slope discrepancies between the cognitive status groups in the Traditional MCI model and the Persistent MCI model. Consistent with hypothesis 2a, both the Traditional MCI and the Persistent MCI groups had significantly lower baseline EPT performance relative to their respective Control groups. However, neither MCI group differed from Controls in terms of their trajectory of change over time.

Table 19. Group Differences in EPT Baseline Score and Rate of Change for Traditional MCI and Persistent$_{2y}$ MCI Models

HLM Model	Baseline Difference	SE	*t*	Slope Difference	SE	*t*
Traditional MCI						
Ctrl vs. T-MCI	-3.89	.67	-5.83 **	-.13	.11	-1.23
Persistent$_{2y}$ MCI						
Ctrl vs. P-MCI	-5.46	.76	-7.17 **	-.22	.13	-1.74
Ctrl vs. Unst	-1.90	.81	-2.33 *	-.15	.13	-1.14
Unst vs. P-MCI	-3.56	.80	-4.43 **	-.07	.13	-.53

Note: ** $p < .01$. * $p < .05$. Higher scores on EPT indicate better performance. Negative baseline difference values indicate poorer baseline performance in the second group of comparison. Negative slope difference values indicate poorer trajectory of performance in second group of comparison.

Again, consistent with hypothesis 2b, the absolute magnitude of the discrepancy in baseline performance was larger for the Persistent MCI versus Control comparison than for the Traditional MCI versus Control comparison. As neither the Traditional MCI nor the Persistent MCI groups differed from Controls in terms of their rate of change over time, the relative magnitude of these differences was not evaluated.

Section Summary

In terms of overall model fit, a similar pattern of findings was observed across outcome measures. That is, for all outcome measures, overall model fit parameters (i.e., AIC values) were improved by the addition of a either a Traditional MCI or a Persistent MCI cognitive status grouping variable to the Level 2 models. This indicated that a significant portion of within-subjects variance in individual performance curves on each outcome measure could be accounted for by cognitive status group membership.

Similarly, for all outcome measures, overall model fit parameters were lower for the Persistent MCI models relative to the Traditional MCI models. This indicated that the Persistent MCI cognitive status group variable provided a better accounting of observed individual performance curves than did the Traditional MCI cognitive status group variable. Although these model fit findings do not represent a direct test of the hypotheses outlined in Objective 2, they are in keeping with what might be expected if the Persistent MCI classification more accurately captured those at risk of decline.

In terms of Level 2 parameter estimates, or values for cognitive status group baseline performance and rate of change over time, a similar pattern was again observed across most measures. Consistent with the initial part of hypothesis 2a, baseline performance differences were consistently observed for Traditional MCI versus Control comparisons (with the exception of the ADL measure) and for Persistent MCI versus Control comparisons. However, findings were only somewhat consistent with the second part of hypothesis 2a. It was hypothesized that the Traditional and Persistent MCI groups would differ from Controls in terms of their average rates of change over time, but this finding was only seen in the MMSE measure (for the Traditional and Persistent MCI groups) and in the Trails B measure (for the Persistent MCI group). It is noted that, at the sample level, a significant negative overall trajectory of change was observed for all measures except the Coding Recall measure, which had an overall trajectory of change that did not differ significantly from zero.

Table 20 presents a summary of cognitive status group differences in baseline performance and rate of change over time as derived from HLM models outlined above. As parameter estimates in the previous section were presented in raw score units for each

outcome measure, the values in the following table are presented on a Z score metric (i.e., $M = 0$, $SD = 1$; baseline mean and standard deviation values for the Year 1 Basic Control group used as a normative reference) to facilitate informal comparison of baseline and rate of change differences across groups and across measures. For example, the difference between the T-MCI group and Controls on the MMSE at baseline was equal to 0.36 SD (relative to the Control group baseline performance). The difference between the Traditional MCI group and Controls in terms of five year rate of change, on the other hand, was equal to 0.66 SD. These values are presented for summary purposes only. No statistical comparisons were conducted to (a) compare effects between measures or (b) compare magnitude of baseline versus slope discrepancies.

Table 20. Summary of Cognitive Status Group Differences in Baseline Performance and Rate of Change across Cognitive and Functional Outcome Measures

Outcome Measure	Standardized Absolute Difference [a] Relative to Controls			
	Baseline Performance		Five-Year Rate of Change	
Cognitive	T-MCI	P_{2y}-MCI	T-MCI	P_{2y}-MCI
Mini-Mental State Exam	**.36**	**.54**	**.66**	**.93**
Trail Making Test Part B	**.77**	**1.08**	.51	**.76**
Digit Symbol Coding Recall	**.59**	**.78**	.11	.28
Functional				
Activities of Daily Living	.17	**.48**	.10	.02
Everyday Problems Test	**.92**	**1.29**	.15	.26

Note: [a] Values represent absolute raw score differences converted to Z score metric by standardizing relative to Control group (mean and SD) baseline performance. In all cases where results were statistically reliable, the MCI groups performed more poorly at baseline and had a steeper trajectory of decline over time relative to the Control groups. Significant values ($p < .05$) are shown in bold.

Hypothesis 2b posited that the magnitude of discrepancies in baseline performance and rate of change over time between Persistent MCI versus Control groups would exceed discrepancies between Traditional MCI versus Controls groups. Because there is no appropriate statistical technique to test this hypothesis, these results were evaluated using descriptive comparison of parameter estimates. This hypothesis was almost universally supported, with the exception of the ADL slope values.

Objective 3: Robustness of Persistent MCI Classification

The third section addresses the robustness of the Persistent MCI classification by replicating analyses performed in the previous section, systematically varying elements of the operational definition of MCI.

Analysis Plan

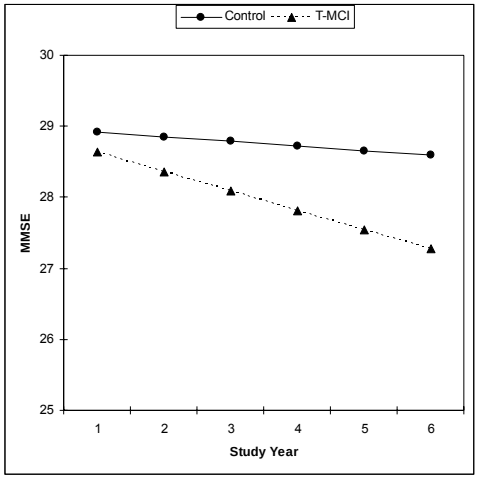
HLM analyses in the form outlined in the previous section were replicated, varying elements of the operational definition of MCI in four ways. Under the first variation (Severity), a more stringent normative cutoff for inclusion in the MCI groups (e.g., from 1 SD to 1.5 SD) was applied. Under the second variation (Duration), an increased duration of cognitive impairment was required for inclusion in a Persistent MCI group (e.g., from two to three consecutive assessments). Under the third variation (Pervasiveness), the MCI groups were sub-classified according to the pervasiveness of cognitive impairment (i.e., impaired performance on single test versus multiple tests). Under the fourth variation (Specificity), the MCI groups were sub-classified according to the specific cognitive domain of observed impairment (i.e., amnestic versus nonamnestic impairment). Results for all four variations are presented simultaneously by outcome

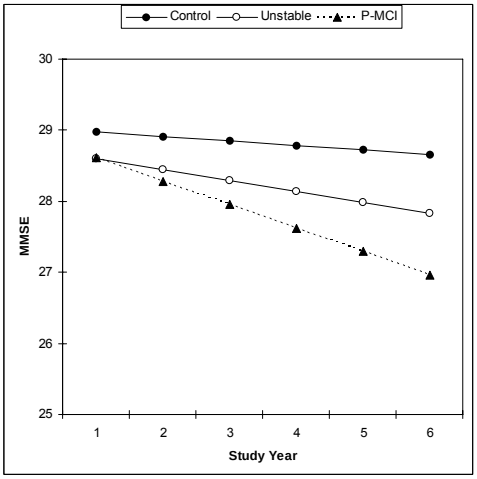
measure. Because of the quantity of analyses performed, presentation of results was condensed relative to that of the previous section. Average cognitive status group baseline performance and rates of change over time are omitted from the text, instead presented exclusively in figures. Model fit statistics are also omitted. In general, these model fit statistics followed patterns outlined in the previous section (i.e., best model fit found for Persistent MCI models, compared to Traditional MCI models and simple Time in Study models).

Results for Cognitive Outcome Measures

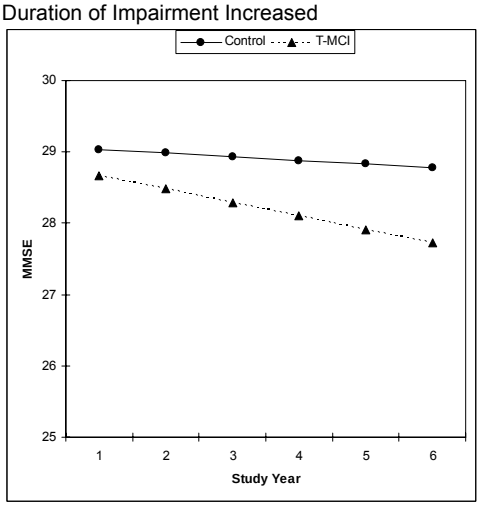
Global Cognitive Functioning (MMSE): Parameter estimates for HLM Level 2 equations were used to derive average baseline scores and rates of change over time for each cognitive status group under each variation in MCI inclusion criteria. These data are presented visually in an array of graphs (Figure 13). On the left, average performance for cognitive status groups under the Traditional MCI classification are presented. On the right, average performance for cognitive status groups under the Persistent MCI classification are presented. Headings indicate which permutation of MCI inclusion criteria is under consideration.

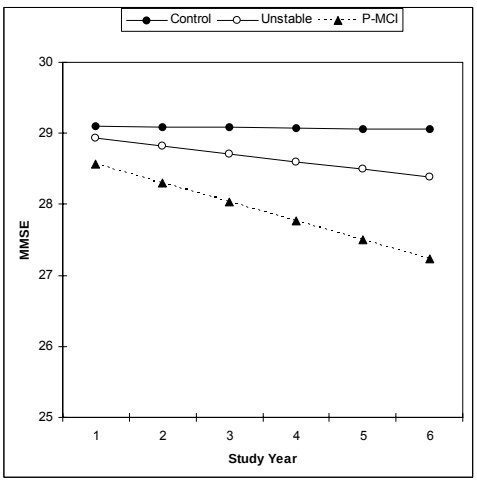
Severity of Impairment Increased



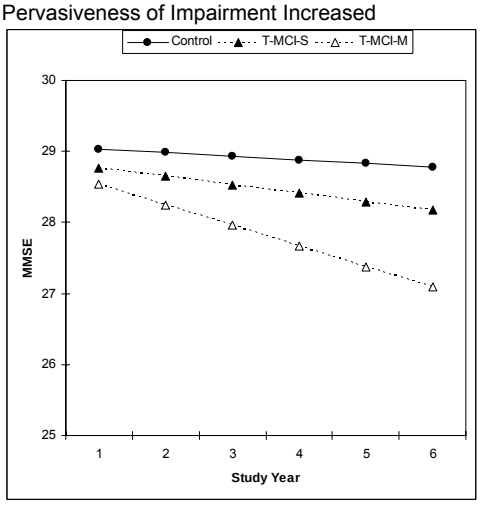


Duration of Impairment Increased





Pervasiveness of Impairment Increased



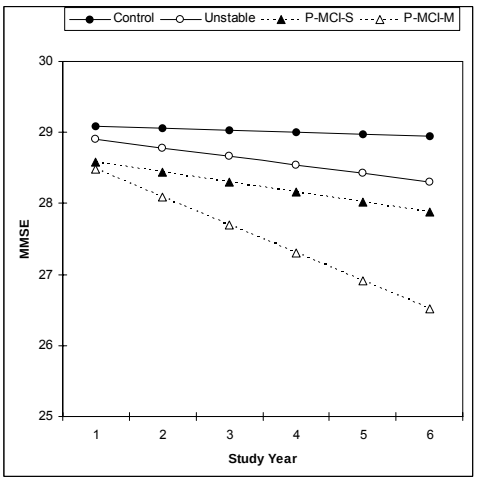


Figure continues

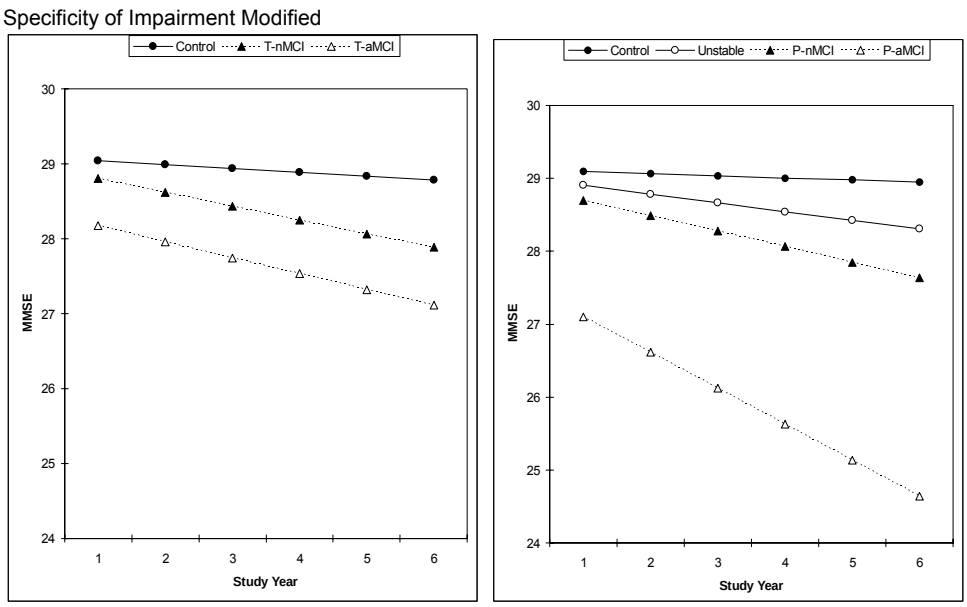


Figure 13 Cognitive Status Group performance on MMSE across variations in MCI Inclusion Criteria. Traditional MCI shown at left, Persistent MCI shown at right

Table 21 presents results of statistical comparisons of cognitive status group differences in baseline performance and rate of change over time for the Traditional and Persistent MCI models under the four variations for MCI inclusion criteria. These comparisons provide a test of Hypothesis 3a. That is, it was expected that, for each variation in MCI inclusion criteria, both the Traditional MCI and the Persistent MCI groups would show worse baseline performance and greater decline over time on the MMSE relative to their respective Control group. In general, this pattern of results was observed, although the difference in baseline performance in the Control versus T-$MCI_{1.5\ SD}$ groups was not statistically significant and the difference in rate of change over time between the Control versus the T-aMCI group was not statistically significant.

Table 21. Cognitive Status Group Differences in MMSE Baseline Score and Rate of Change for Alternative Operational Definitions of Traditional and Persistent MCI Classifications

MCI Variation	Baseline Difference	SE	*t*	Slope Difference	SE	*t*
Severity						
Ctrl vs. T-MCI$_{1.5\ SD}$	-.28	.15	-1.90	-.21	.05	-3.90 **
Ctrl vs. P-MCI$_{1.5\ SD}$	-.37	.18	-1.98 *	-.27	.07	-3.90 **
Duration						
Ctrl vs. T-MCI	-.33	.13	-2.50 *	-.13	.05	-2.57 *
Ctrl vs. P$_{3y}$-MCI	-.53	.17	-3.19 **	-.26	.06	-4.16 **
Pervasiveness						
Ctrl vs. T-MCI-M	-.50	.17	-2.92 **	-.24	.06	-3.80 **
Ctrl vs. P-MCI-M	-.61	.21	-2.88 **	-.37	.08	-4.59 **
Specificity						
Ctrl vs. T-aMCI	-.87	.22	-4.00 **	-.16	.09	-1.89
Ctrl vs. P-aMCI	-1.99	.36	-5.55 **	-.46	.16	-2.85 **

Note: ** $p < .01$. * $p < .05$. Higher scores on MMSE indicate better performance. Negative baseline difference values indicate poorer baseline performance in the second group of comparison. Negative slope difference values indicate poorer trajectory of performance in second group of comparison.

Descriptive comparisons of baseline and slope differences for MMSE scores between the Traditional MCI group versus Controls and the Persistent MCI versus Controls (i.e., tests of Hypothesis 3b) yielded a familiar pattern of results. That is, the Persistent MCI groups were consistently poorer in terms of their baseline MMSE status and rate of change relative to Controls than were the Traditional MCI groups.

Executive Functioning (Trails B): Parameter estimates for HLM Level 2 equations were used to derive average baseline scores and rates of change over time for each cognitive status group under each variation in MCI inclusion criteria. These data are presented visually in an array of graphs (Figure 14). On the left, Level 2 average performances for cognitive status groups under the Traditional MCI classification are presented. On the right, Level 2 average performances for cognitive status groups under the Persistent MCI classification are presented.

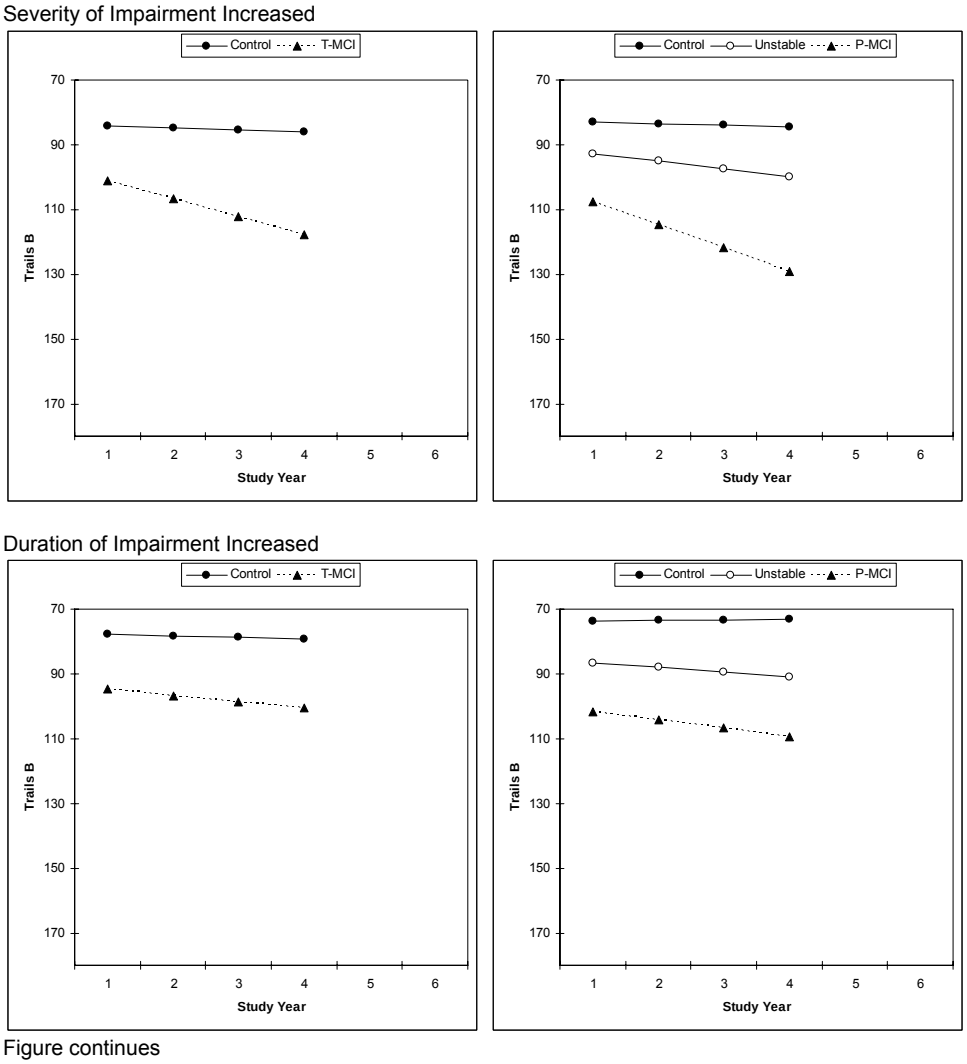


Figure continues

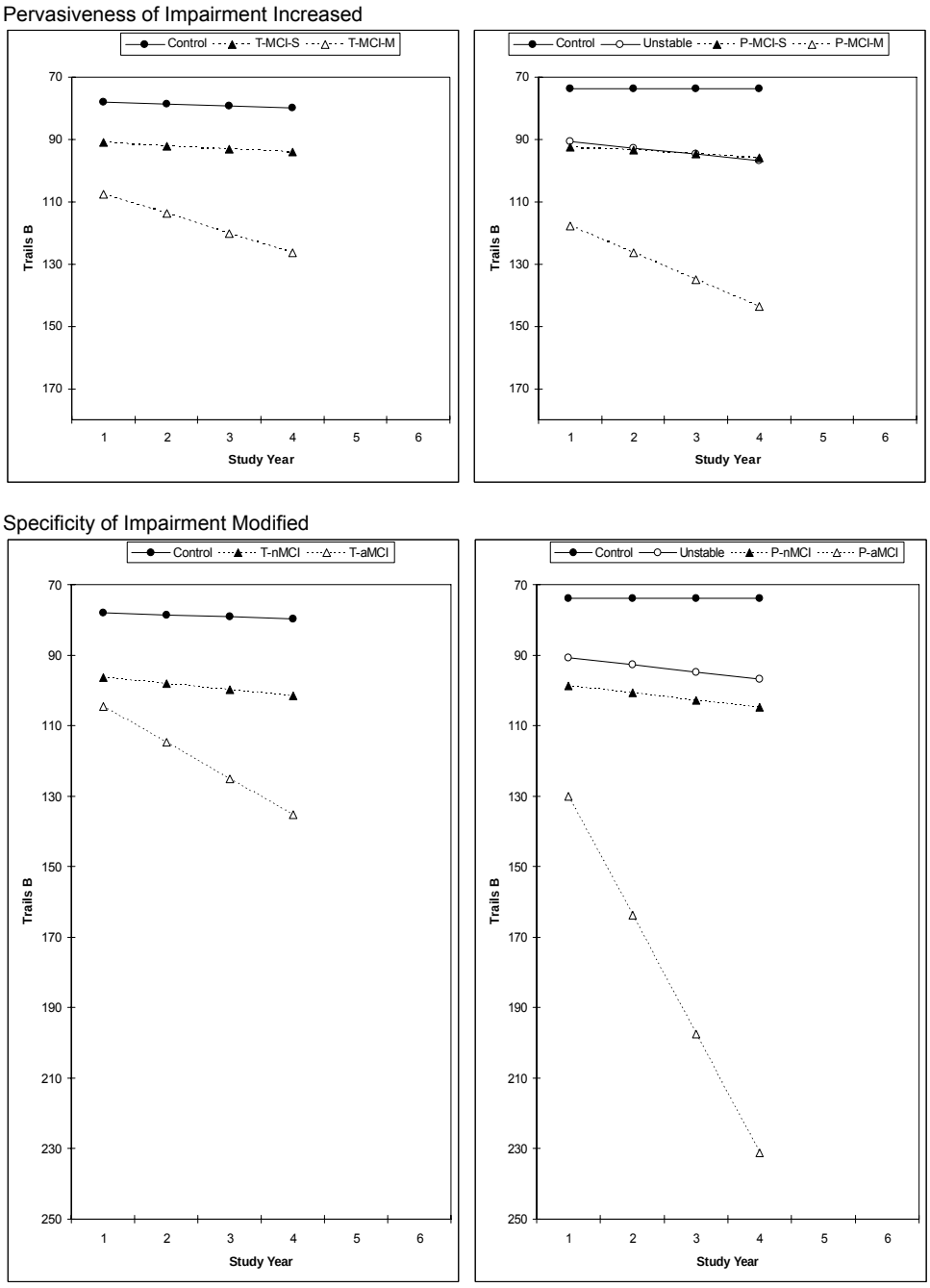


Figure 14. Cognitive Status Group performance on Trails B across variations in MCI Inclusion Criteria. Traditional MCI shown at left, Persistent MCI shown at right

Table 22 presents results of statistical comparisons of cognitive status group differences in baseline performance and rate of change over time on the Trails B measure for the Traditional and Persistent MCI models under the four variations for MCI

inclusion criteria. Both the Traditional MCI and the Persistent MCI groups consistently demonstrated poorer baseline performance relative to Controls. Differences in rates of change over time between the MCI and Control groups were observed in the increased Severity, the increased Pervasiveness, and the cognitive domain Specificity MCI inclusion variations, but not the increased Duration MCI inclusion variation.

Table 22. Cognitive Status Group Differences in Trails B Baseline Score and Rate of Change for Alternative Operational Definitions of Traditional and Persistent MCI Classifications

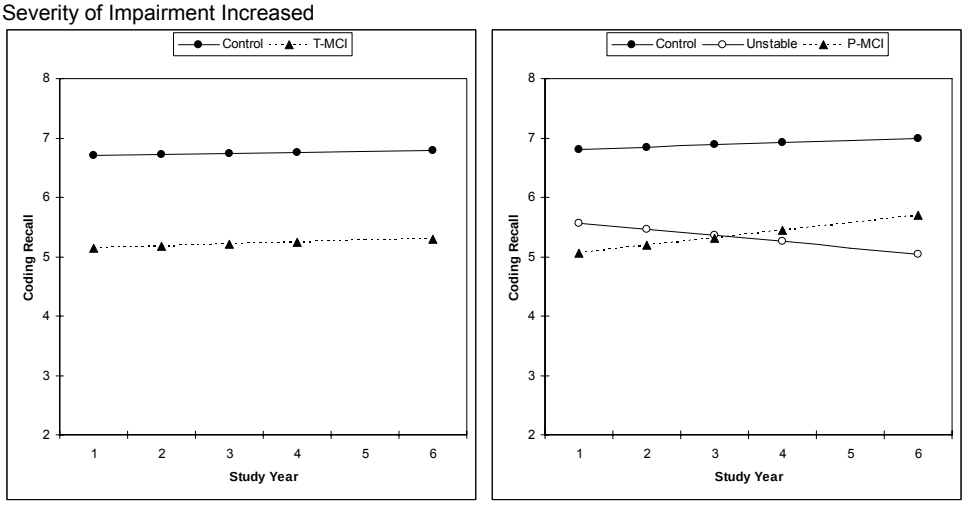
MCI Variation	Baseline Difference	SE	*t*	Slope Difference	SE	*t*
Severity						
Ctrl vs. T-MCI$_{1.5\ SD}$	16.94	4.69	3.61 **	4.91	1.71	2.87 **
Ctrl vs. P-MCI$_{1.5\ SD}$	24.66	5.77	4.27 **	6.63	2.20	3.01 **
Duration						
Ctrl vs. T-MCI	16.91	3.65	4.63 **	1.50	1.17	1.28
Ctrl vs. P$_{3y}$-MCI	28.14	4.53	6.21 **	2.69	1.50	1.79
Pervasiveness						
Ctrl vs. T-MCI-M	29.64	5.22	5.68 **	5.65	2.01	2.81 **
Ctrl vs. P-MCI-M	43.95	6.28	7.00 **	8.67	2.53	3.43 **
Specificity						
Ctrl vs. T-aMCI	26.63	6.70	3.98 **	9.63	2.81	3.42 **
Ctrl vs. P-aMCI	56.34	10.77	5.23 **	33.66	5.40	6.24 **

Note: ** $p < .01$. * $p < .05$. Higher scores on Trails B indicate poorer performance. Positive baseline difference values indicate poorer performance in the second group listed in comparison. Positive slope difference values indicate poorer trajectory of performance over time in the second group.

Descriptive comparisons of baseline and slope differences between the Traditional MCI group versus Controls and the Persistent MCI versus Controls revealed that, consistent with Hypothesis 3b, the Persistent MCI groups were consistently poorer relative to Controls than were the Traditional MCI groups.

Visual Memory (Coding Recall): Parameter estimates for HLM Level 2 equations were used to derive average baseline scores and rates of change over time for each cognitive status group under each variation in MCI inclusion criteria. These data are presented visually in an array of graphs (Figure 15).

Severity of Impairment Increased



Duration of Impairment Increased

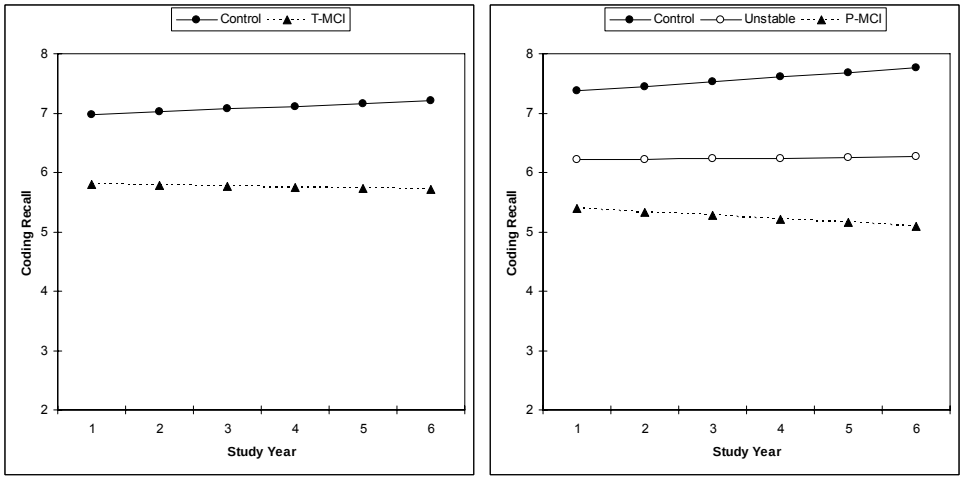


Figure continues

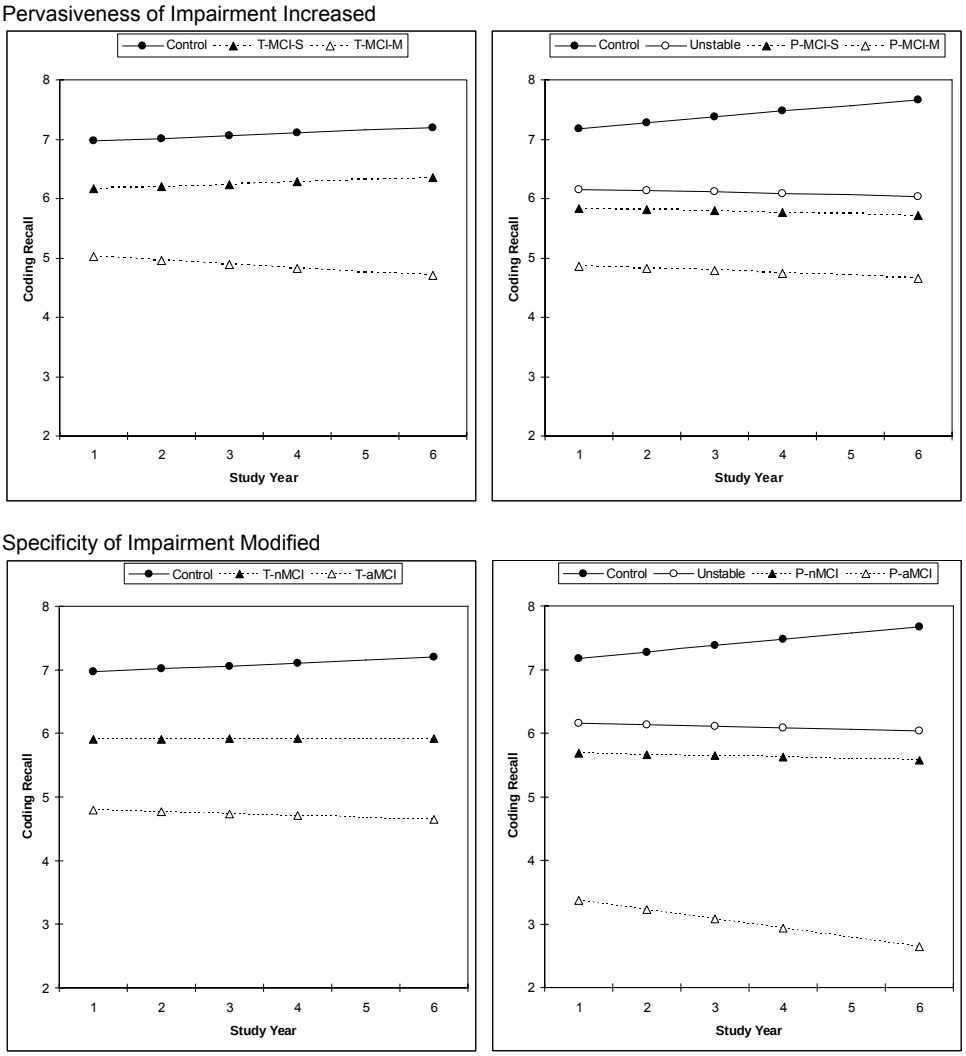


Figure 15. Cognitive Group performance on Coding Recall across variations in MCI Inclusion Criteria. Traditional MCI shown at left, Persistent MCI shown at right

Table 23 presents results of statistical comparisons of cognitive status group differences in baseline performance and rate of change over time on the Coding Recall measure for the Traditional and Persistent MCI models under the four variations for MCI inclusion criteria. Both the Traditional MCI and the Persistent MCI groups consistently demonstrated poorer baseline performance relative to Controls. No differences were observed in rates of change over time between the MCI and Control groups.

Table 23. Cognitive Status Group Differences in Coding Recall Baseline Score and Rate of Change for Alternative Operational Definitions of Traditional and Persistent MCI Classifications

MCI Variation	Baseline Difference	SE	*t*	Slope Difference	SE	*t*
Severity						
Ctrl vs. T-MCI$_{1.5\ SD}$	-1.57	.28	-5.62 **	.01	.08	.19
Ctrl vs. P-MCI$_{1.5\ SD}$	-1.75	.35	-5.05 **	.09	.10	.94
Duration						
Ctrl vs. T-MCI	-1.18	.26	-4.52 **	-.06	.07	-.92
Ctrl vs. P$_{3y}$-MCI	-1.98	.32	-6.09 **	-.14	.09	-1.54
Pervasiveness						
Ctrl vs. T-MCI-M	-1.95	.32	-6.15 **	-.11	.09	-1.21
Ctrl vs. P-MCI-M	-2.32	.39	-5.89 **	-.14	.12	-1.17
Specificity						
Ctrl vs. T-aMCI	-2.17	.41	-5.29 **	-.08	.12	-.63
Ctrl vs. P-aMCI	-3.81	.68	-5.60 **	-.24	.23	-1.08

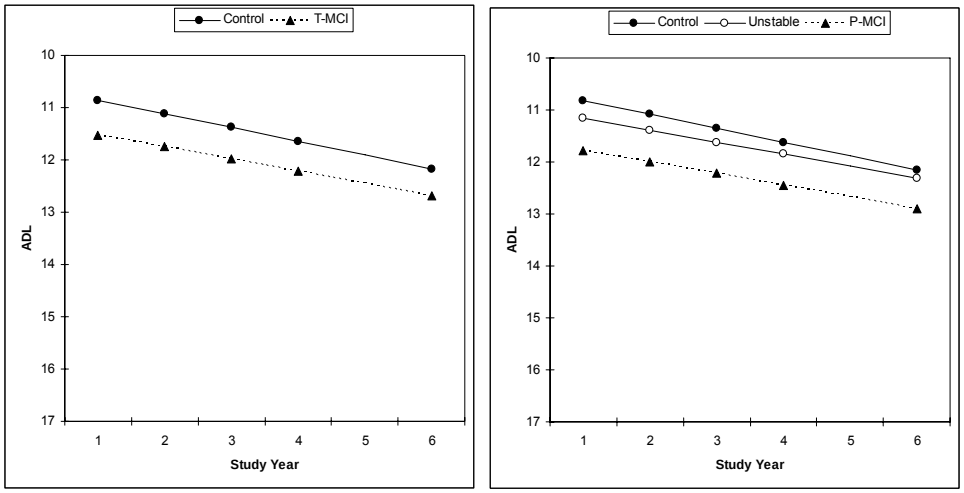
Note: ** $p < .01$. * $p < .05$. Higher scores on Coding Recall indicate better performance. Negative baseline difference values indicate poorer baseline performance in the second group of comparison. Negative slope difference values indicate poorer trajectory of performance in second group of comparison.

Descriptive comparisons of baseline and slope differences between the Traditional MCI group versus Controls and the Persistent MCI versus Controls revealed that, consistent with Hypothesis 3b, the Persistent MCI groups were consistently poorer relative to Controls than were the Traditional MCI groups.

Results for Functional Outcome Measures

Global Functional Status (ADL): Parameter estimates for HLM Level 2 equations were used to derive average baseline scores and rates of change over time for each cognitive status group under each variation in MCI inclusion criteria. These data are presented visually in an array of graphs (Figure 16).

Severity of Impairment Increased



Duration of Impairment Increased

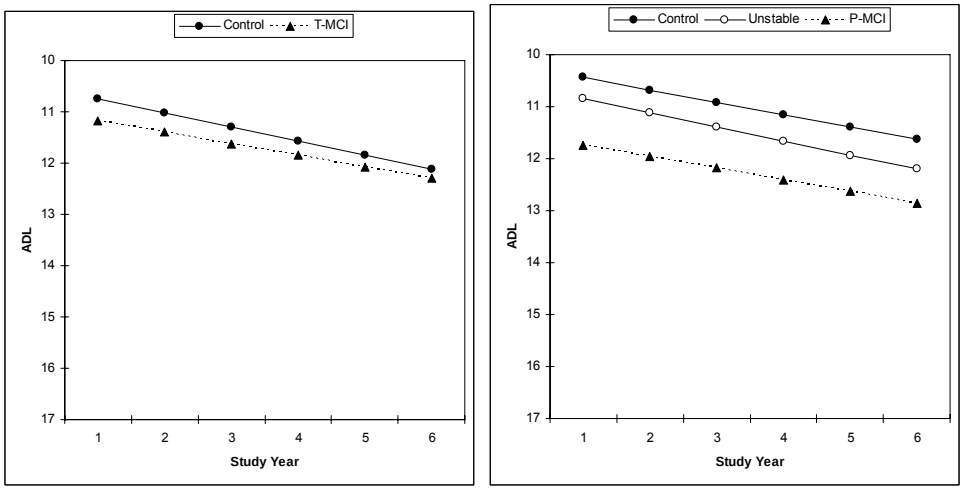


Figure continues

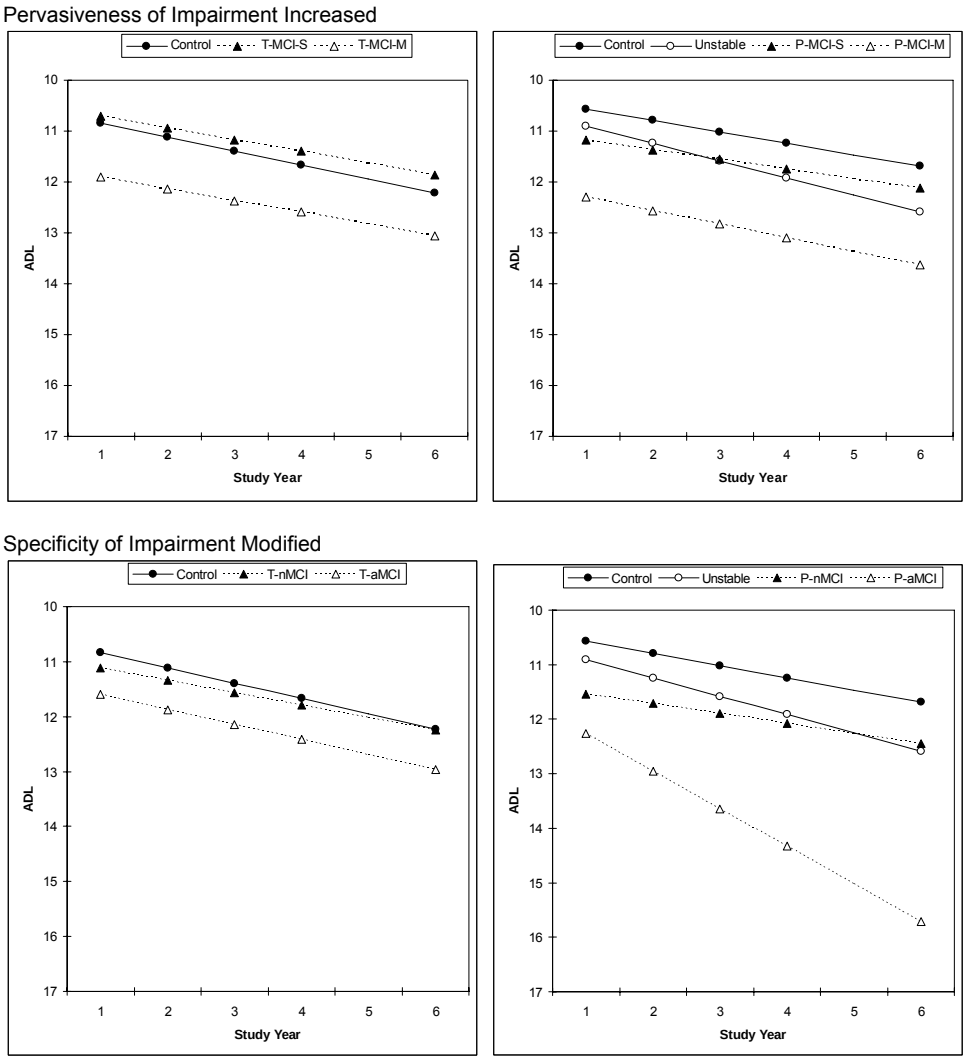


Figure 16. Cognitive Status Group performance on ADL across variations in MCI Inclusion Criteria. Traditional MCI shown at left, Persistent MCI shown at right

Table 24 presents results of statistical comparisons of cognitive status group differences in baseline performance and rate of change over time on the ADL measure for the Traditional and Persistent MCI models under the four variations for MCI inclusion criteria. In general, the Traditional MCI and the Persistent MCI groups demonstrated poorer baseline performance relative to Controls. There were no differences observed in rates of change over time between the MCI and Control groups, with the exception of a

statistically significant difference between the Persistent MCI group and Controls in the Specificity variation.

Table 24. Cognitive Status Group Differences in ADL Baseline Score and Rate of Change for Alternative Operational Definitions of Traditional and Persistent MCI Classifications

MCI Variation	Baseline Difference	SE	*t*	Slope Difference	SE	*t*
Severity						
Ctrl vs. T-MCI$_{1.5\ SD}$	.67	.34	1.99 *	-.03	.08	-.42
Ctrl vs. P-MCI$_{1.5\ SD}$	.96	.42	2.31 *	-.05	.10	-.47
Duration						
Ctrl vs. T-MCI	.42	.31	1.35	-.05	.07	-.71
Ctrl vs. P$_{3y}$-MCI	1.30	.39	3.34 **	-.02	.09	-.18
Pervasiveness						
Ctrl vs. T-MCI-M	1.06	.38	2.77 **	-.04	.09	-.49
Ctrl vs. P-MCI-M	1.73	.47	3.68 **	.04	.11	.38
Specificity						
Ctrl vs. T-aMCI	.76	.50	1.53	.00	.12	-.02
Ctrl vs. P-aMCI	1.70	.82	2.07 *	.46	.22	2.07 *

Note: ** $p < .01$. * $p < .05$. Higher scores on ADL indicate poorer performance. Positive baseline difference values indicate poorer performance in the second group of comparison. Positive slope difference values indicate poorer trajectory of performance in second group of comparison.

Comparisons of baseline and slope differences between the Traditional MCI group versus Controls and the Persistent MCI versus Controls revealed that, consistent

with Hypothesis 3b, the Persistent MCI groups were generally poorer relative to Controls than were the Traditional MCI groups.

Applied Problem Solving (EPT): Parameter estimates for HLM Level 2 equations were used to derive average baseline scores and rates of change over time for each cognitive status group under each variation in MCI inclusion criteria. These data are presented visually in an array of graphs (Figure 17).

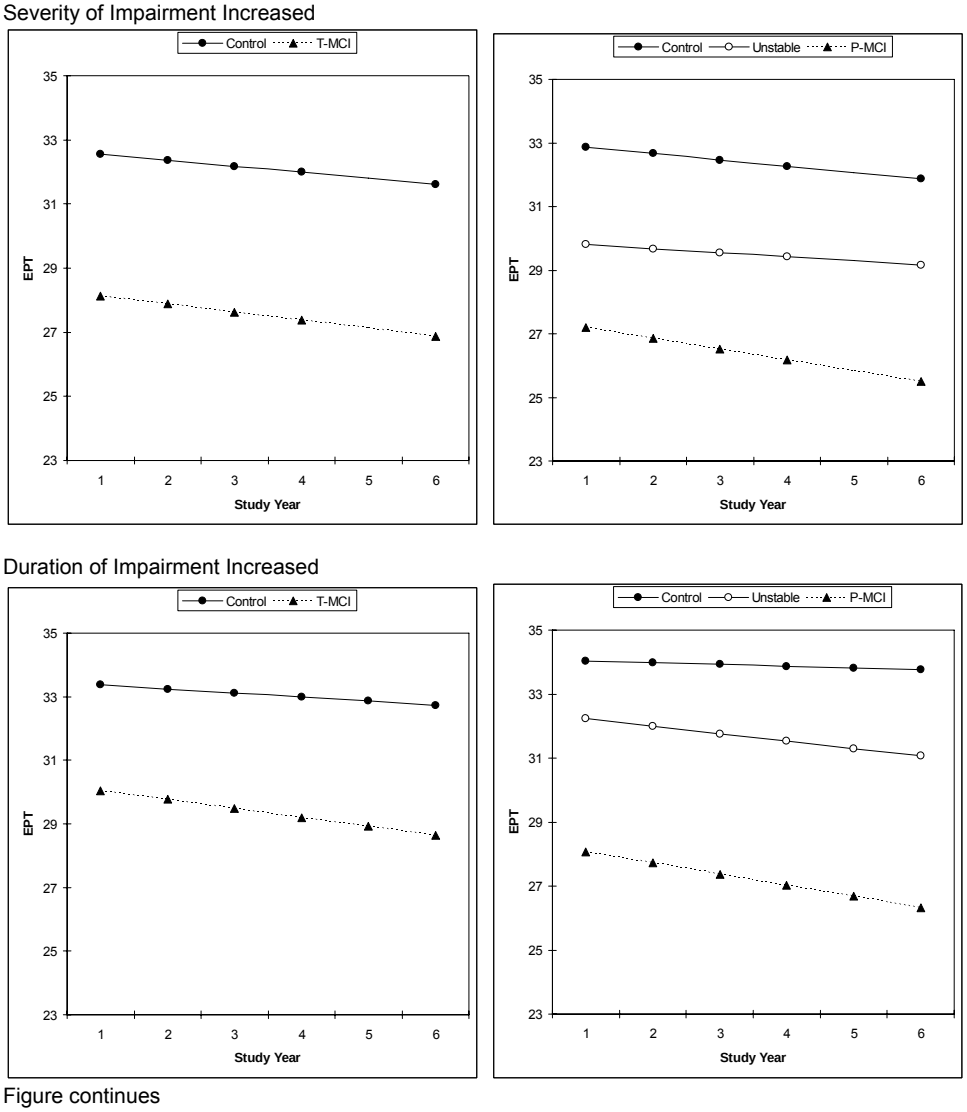


Figure continues

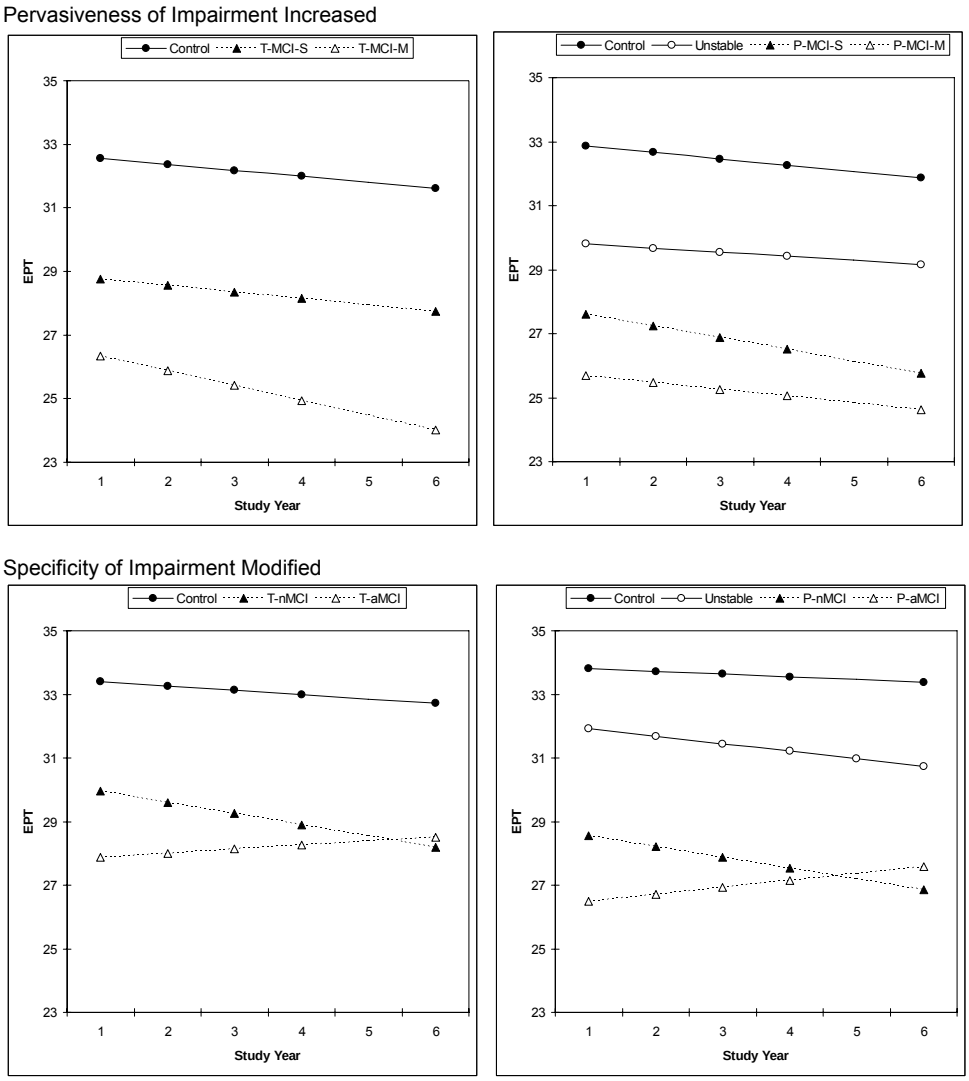


Figure 17. Cognitive Status Group performance on EPT across variations in MCI Inclusion Criteria. Traditional MCI shown at left, Persistent MCI shown at right

Table 25 presents results of statistical comparisons of cognitive status group differences in baseline performance and rate of change over time on the EPT measure for the Traditional and Persistent MCI models under the four variations for MCI inclusion criteria. The Traditional MCI and the Persistent MCI groups consistently demonstrated poorer baseline performance relative to Controls. A statistically significant difference in rate of change over time was observed between the Persistent MCI and Control groups in the increased Duration MCI inclusion variation, but this was not observed between the

parallel Traditional MCI and Control groups, nor was this observed in the Severity, Pervasiveness, or Specificity MCI inclusion variations.

Table 25. Cognitive Status Group Differences in EPT Baseline Score and Rate of Change for Alternative Operational Definitions of Traditional and Persistent MCI Classifications

MCI Variation	Baseline Difference	SE	*t*	Slope Difference	SE	*t*
Severity						
Ctrl vs. T-MCI$_{1.5\ SD}$	-4.42	.73	-6.08 **	-.06	.12	-.54
Ctrl vs. P-MCI$_{1.5\ SD}$	-5.68	.89	-6.37 **	-.14	.15	-.89
Duration						
Ctrl vs. T-MCI	-3.33	.66	-5.06 **	-.15	.11	-1.44
Ctrl vs. P$_{3y}$-MCI	-5.97	.79	-7.54 **	-.29	.14	-2.08 *
Pervasiveness						
Ctrl vs. T-MCI-M	-6.06	.81	-7.47 **	-.12	.14	-.81
Ctrl vs. P-MCI-M	-8.36	.97	-8.58 **	-.12	.19	-.56
Specificity						
Ctrl vs. T-aMCI	-5.52	1.07	-5.15 **	.26	.19	1.37
Ctrl vs. P-aMCI	-7.31	1.76	-4.14 **	.30	.39	.78

Note: ** $p < .01$. * $p < .05$. Higher scores on EPT indicate better performance. Negative baseline difference values indicate poorer baseline performance in the second group of comparison. Negative slope difference values indicate poorer trajectory of performance in second group of comparison.

Descriptive comparisons of baseline and slope differences between the Traditional MCI group versus Controls and the Persistent MCI versus Controls revealed

that, consistent with Hypothesis 3b, the Persistent MCI groups were generally poorer relative to Controls than were the Traditional MCI groups.

Section Summary

Although model fit statistics were not reported in detail in Objective 3 results, these statistics followed patterns outlined in Objective 2. Specifically, model fit was better in the Persistent MCI and Traditional MCI models relative to the simple Time in Study models. Moreover, model fit was best in the Persistent MCI models compared to the Traditional MCI models. Although these model fit findings do not represent a direct test of the hypotheses outlined in Objective 3, they are in keeping with what might be expected if the Persistent MCI classification more accurately captured those at risk of decline.

In terms of Level 2 parameter estimates, or values for cognitive status group baseline performance and rate of change over time, a similar pattern was again observed across most measures. Consistent with the first part of Hypothesis 3a, baseline performance differences were almost universally observed for Traditional MCI versus Control comparisons. Two exceptions included a failure to find significant differences on the MMSE between the Traditional MCI and Control groups under the Severity variation of MCI inclusion criteria, and on the ADL under the Duration and Specificity variations of MCI inclusion criteria. Findings were only partially consistent with the second part of hypothesis 3a, which posited that the Traditional and Persistent MCI groups would differ from Controls in terms of their average rates of change over time. The Traditional and Persistent MCI groups differed from Controls on the MMSE measure

(for all variations of the Traditional and Persistent MCI groups except the Traditional MCI group in the Specificity variation), on the Trails B measure (for the Severity, Pervasiveness, and Specificity variations of both the Traditional and Persistent MCI groups), on the ADL measure (for the Specificity variation of the Persistent MCI group), and on the EPT measure (for the Duration variation of the Persistent MCI group).

Tables 26 through 29 present a summary of cognitive status group differences in baseline performance and rate of change over time as derived from HLM models outlined in this section. Similar to the section summary for Objective 2 results, parameter estimates were converted from raw score units to a Z score metric relative to Year 1 Basic Control group (as shown in Figure 1) performance on each outcome measure to facilitate informal comparison of baseline and rate of change differences across groups, measures, and variations of MCI inclusion criteria. These values are presented for summary purposes only.

Table 26. Summary of Cognitive Status Group Differences in Baseline Performance and Five Year Rate of Change across Cognitive and Functional Outcome Measures under More Stringent Normative Cutoff for MCI (Severity variation)

Outcome Measure	Standardized Absolute Difference [a] Relative to Controls			
	Baseline Performance		Five-Year Rate of Change	
Cognitive	T-MCI	P_{2y}-MCI	T-MCI	P_{2y}-MCI
Mini-Mental State Exam	.27	**.35**	**1.00**	**1.28**
Trail Making Test Part B	**.64**	**.94**	**.93**	**1.26**
Digit Symbol Coding Recall	**.70**	**.79**	.03	.21
Functional				
Activities of Daily Living	**.30**	**.44**	.07	.10
Everyday Problems Test	**1.04**	**1.35**	.08	.16

Note: [a] Values represent absolute raw score differences converted to Z score metric by standardizing relative to Year 1 Control group (Figure 1) baseline performance. In all cases where results were statistically reliable, the MCI groups performed more poorly at baseline and had a steeper trajectory of decline over time relative to the Control groups. Significant values ($p < .05$) are shown in bold.

Table 27. Summary of Cognitive Status Group Differences in Baseline Performance and Five Year Rate of Change across Cognitive and Functional Outcome Measures under Increased Duration of Impairment for Persistent MCI (Duration variation)

Outcome Measure	Standardized Absolute Difference [a] Relative to Controls			
	Baseline Performance		Five-Year Rate of Change	
Cognitive	T-MCI	P_{3y}-MCI	T-MCI	P_{3y}-MCI
Mini-Mental State Exam	**.36**	**.51**	**.66**	**1.24**
Trail Making Test Part B	**.77**	**1.07**	.51	.51
Digit Symbol Coding Recall	**.59**	**.89**	.11	.31
Functional				
Activities of Daily Living	.17	**.59**	.10	.04
Everyday Problems Test	**.92**	**1.41**	.15	**.34**

Note: [a] Values represent absolute raw score differences converted to Z score metric by standardizing relative to Year 1 Control group (Figure 1) baseline performance. In all cases where results were statistically reliable, the MCI groups performed more poorly at baseline and had a steeper trajectory of decline over time relative to the Control groups. Significant values ($p < .05$) are shown in bold.

Table 28. Summary of Cognitive Status Group Differences in Baseline Performance and Five Year Rate of Change across Cognitive and Functional Outcome Measures when MCI Groups Sub-Classified by Pervasiveness of Impairment (Pervasiveness variation)

Outcome Measure	Standardized Absolute Difference [a] Relative to Controls			
	Baseline Performance		Five-Year Rate of Change	
Cognitive	T-MCI-M	P-MCI-M	T-MCI-M	P-MCI-M
Mini-Mental State Exam	**.88**	**1.18**	**1.15**	**1.75**
Trail Making Test Part B	**1.13**	**1.67**	**1.07**	**1.65**
Digit Symbol Coding Recall	**.88**	**1.04**	.25	.31
Functional				
Activities of Daily Living	**.48**	**.78**	.10	.10
Everyday Problems Test	**1.43**	**1.97**	.14	.14

Note: [a] Values represent absolute raw score differences converted to Z score metric by standardizing relative to Year 1 Control group (Figure 1) baseline performance. In all cases where results were statistically reliable, the MCI groups performed more poorly at baseline and had a steeper trajectory of decline over time relative to the Control groups. Significant values ($p < .05$) are shown in bold.

Table 29. Summary of Cognitive Status Group Differences in Baseline Performance and Five Year Rate of Change across Cognitive and Functional Outcome Measures when MCI Groups Sub-Classified by Specific Domain of Cognitive Impairment (Specificity variation)

Outcome Measure	Standardized Absolute Difference [a] Relative to Controls			
	Baseline Performance		Five-Year Rate of Change	
Cognitive	T-aMCI	P-aMCI	T-aMCI	P-aMCI
Mini-Mental State Exam	**.83**	**1.91**	.77	**2.22**
Trail Making Test Part B	**1.01**	**2.14**	**1.83**	**6.39**
Digit Symbol Coding Recall	**.98**	**1.71**	.17	.55
Functional				
Activities of Daily Living	.34	**.77**	.01	**1.05**
Everyday Problems Test	**1.30**	**1.72**	.31	.36

Note: [a] Values represent absolute raw score differences converted to Z score metric by standardizing relative to Year 1 Control group (Figure 1) baseline performance. In all cases where results were statistically reliable, the MCI groups performed more poorly at baseline and had a steeper trajectory of decline over time relative to the Control groups. Significant values ($p < .05$) are shown in bold.

Hypothesis 3b posited that the magnitude of discrepancies in baseline performance and rate of change over time between Persistent MCI versus Control groups would exceed discrepancies between Traditional MCI versus Controls groups. Because there is no appropriate statistical technique to test this hypothesis, these results were evaluated using descriptive comparison of parameter estimates. This hypothesis was generally supported. That is, the difference in baseline performance between the Persistent aMCI group and Controls was larger than the baseline performance difference between the Traditional aMCI group and Controls on all of the cognitive and functional measures. The difference in rate of change between the Persistent aMCI group and Controls was larger than the difference in rate of change between the Traditional aMCI group and Controls on most measures, with the exception of the Trails B measure under the Duration variation, the ADL measure under the Duration variation and the Pervasiveness variation, and the EPT measures under the Pervasiveness variation.

Discussion

Overview

The aim of the current study was to evaluate a proposed revision to existing methods for identifying individuals at elevated risk of dementia. The term MCI is used to describe older adults who show cognitive impairment relative to their peers, but do not meet full criteria for any dementia syndrome. Although some argue that MCI can be considered a precursor to dementia, evidence for this has been limited. Prior studies have shown that, at the group level, those identified as MCI show elevated rates of conversion to dementia compared to those without cognitive impairment. However, at the individual

level, heterogeneity of outcomes is common for those with MCI, with some individuals going on to develop dementia, some remaining stable for long periods, and some reverting to unimpaired status.

One potential source of instability in the MCI designation is a failure to account for normal variation in individual performance across a range of cognitive measures. Contemporary classification procedures for MCI ("Traditional MCI") typically require evidence of cognitive impairment on only a single test at a single occasion of measurement. Studies in normative neuropsychology have shown that, given a multi-test battery of cognitive measures, the probability that a carefully screened, neurologically normal individual will show isolated impairment on a single measure is high. Proposed explanations for isolated "impaired" scores in normal individuals include transient factors such as fatigue, attentional lapses, and/or loss of motivation, or stable features such as long-standing isolated weakness.

To address this overlooked limitation in Traditional MCI classification criteria, the current study proposed a modified classification algorithm for MCI which required evidence of *persistent* cognitive impairment across multiple occasions of measurement ("Persistent MCI"). This novel Persistent MCI classification was designed to reduce the proportion of individuals classified as "accidental MCI" (de Rotrou et al., 2005) based on an impaired score on a single psychometric measure secondary to transient factors. Using data from a prospective five-year longitudinal study of cognitive change in community-dwelling older adults, this modified Persistent MCI classification scheme was evaluated relative to Traditional MCI criteria in terms of prevalence rates, group differences in

background characteristics, and baseline and five-year cognitive and functional outcomes.

A total of five variations of the Persistent and Traditional MCI inclusion criteria were evaluated, given the current lack of consensus as to specific inclusion/exclusion criteria for MCI in the extant literature. First, a "Basic" set of inclusion/exclusion criteria was proposed based on a relatively liberal existing classification scheme (Levy, 1994). In this Basic variation, Traditional MCI was operationalized as performance 1.0 SD below normative mean on any one of five cognitive measures at study Year 1. Persistent MCI was operationalized as performance 1.0 SD below normative mean on any one of five cognitive measures at study Year 1 and Year 2. To test the robustness of the Persistent versus Traditional MCI classification, these criteria were varied systematically as follows: Under the first variation ("Severity"), a more stringent normative cutoff for inclusion in the MCI groups (i.e., from 1 SD to 1.5 SD) was applied. Under the second variation ("Duration"), the duration of cognitive impairment required for inclusion in a Persistent MCI group was increased (i.e., from two to three consecutive assessments). Under the third variation ("Pervasiveness"), the MCI groups were sub-classified according to the pervasiveness of cognitive impairment (i.e., impaired performance on single test versus multiple tests). Under the fourth variation ("Specificity"), the MCI groups were sub-classified according to the specific cognitive domain of observed impairment (i.e., amnestic versus nonamnestic impairment).

Objective 1: Prevalence of Traditional MCI versus Persistent MCI

The first study objective was to replicate findings from a single prior study which employed a Persistent MCI classification scheme in a community sample (i.e., Collie et al., 2002). These authors found that, although 20% of participants met their criteria for Traditional MCI (i.e., memory impairment on one assessment), only 13% met their criteria for Persistent MCI classification (i.e., memory impairment on three semi-annual assessments). Put another way, approximately one third of individuals classified as MCI at baseline did not show persistent evidence of cognitive impairment across three occasions. Thus, Hypothesis 1 specified that base rates of impaired performance on any one cognitive test at any one time point in the current study (Traditional MCI) would exceed base rates of persistent impaired performance observed on consecutive assessment sessions (Persistent MCI). This discrepancy was expected to hold across the five proposed operational definitions of Traditional and Persistent MCI.

Key Findings

Study Objective 1 findings were consistent with this hypothesis. That is, regardless of specific variation of Persistent versus Traditional MCI classification scheme under consideration, the prevalence of Traditional MCI exceeded that of the Persistent MCI. The proportion of individuals who met criteria for MCI at baseline (i.e., met criteria for Traditional MCI) but did not meet criteria for Persistent MCI ranged from 33-67% of the group, depending on the specific classification scheme under consideration. These values tended to exceed the 35% reported by Collie and colleagues, as well as other previously reported estimates which ranged from 15-44% across epidemiological and community samples. Results of the current study are consistent with these prior

findings in that the rates of instability of classification appear to be highly susceptible to the specific inclusion/exclusion criteria for MCI under consideration.

These data were further explored by investigating the overall directionality of instability in each variation of MCI under consideration. For example, in the Basic MCI classification scheme, there was a significant difference in the overall proportion of individuals who showed impairment at Year 1, but not Year 2 compared with the proportion of individuals who showed impairment at Year 2, but not Year 1 such that a greater proportion of individuals showed less impairment at follow-up testing than showed greater impairment. This finding was generally supported in terms of trend across the four remaining MCI classification scheme variations. Results were statistically significant under the Basic and Duration variations. Under the Pervasiveness variation, this finding was significant for the Control versus MCI-Single comparison, but not for either the Control versus MCI-Multiple comparison or the MCI-Single versus MCI-Multiple comparison. Results were not significant under either the Severity or the Specificity conditions. Of note, investigation of Year 2 to Year 3 instability in the Duration variation revealed equal rates of conversion and reversion.

Retest Effects

This observed global trend across all five variations of MCI classification schemes evaluated in the present study towards reduced impairment at follow-up testing is intriguing and somewhat unexpected given the commonly held assumption of generally declining (or at best, stable) cognitive abilities in aging. One key candidate contributor to this finding, and likely to prior findings of instability of MCI classification, is retest or practice effects (i.e., improved test scores at follow-up measurement occasions secondary

to factors such as prior experience with test material and reduced anxiety in the assessment situation).

Recently, a number of authors (e.g., Ferrer, Salthouse, Stewart, & Schwartz, 2004; Rabbitt, Diggle, Smith, Holland, & Innes, 2001; Ronnlund, Nyberg, Backman, & Nilsson, 2005; Salthouse & Tucker-Drob, 2008; Thorvaldsson, Hofer, Berg, & Johansson, 2006) have employed a range of novel statistical and methodological approaches targeted towards dissociating any presumably positive effect of prior testing from the presumably negative effects of age-related maturation in longitudinal studies of cognitive aging. This work has provided evidence for a hypothesis that retest effects may account for a substantial part of the historical discrepancy between the larger apparent age-related decline in cognitive ability evident in cross-sectional as compared to longitudinal studies. As Salthouse (2009) has recently argued, while it is commonly accepted that cohort effects (i.e., differences in performance as a function of birth cohort, with a trend for better performance in more recently-born cohorts) may lead to an *over-estimation* of age-related decline in cross-sectional studies, recent advances in studying retest effects, as well as converging evidence from animal and neurobiological studies, suggest that retest effects may lead to an *under-estimation* of age-related decline in longitudinal studies.

This line of research has clear implications for longitudinal studies of MCI and likely accounts for the trend towards net reduced cognitive impairment at follow-up testing observed in the present study. This includes the present study finding of equal rates of conversion and reversion between the Year 2 and Year 3 classification in the Duration variation as positive benefit from retest effects has been shown to diminish after

the initial follow-up testing session (Ivnik et al., 1999). Future research could employ recently developed statistical approaches to investigate the implications of retest effects more directly in MCI conversion studies in datasets such as the one under consideration the present study.

As Ferrer and colleagues highlight (2004), benefit from retest may be an individual characteristic. Indeed, a failure to achieve normative benefit from prior exposure to testing material may indicate an elevated risk of impending cognitive decline, despite an apparent stability of cognitive test scores over time. One prior study has provided some preliminary evidence in support of this hypothesis, illustrating that a previously identified MCI group, compared to healthy controls, showed significantly reduced retest benefit across four serial computerized testing batteries repeated within a single day (Darby, Maruff, Collie, & McStephen, 2002). Longitudinal investigation of these issues warrants future study.

The substantial rate of reversion within the current and prior MCI studies supports the need for continued efforts to refine diagnostic criteria for MCI. The present study modification (i.e., requiring evidence of persistent impairment for MCI inclusion) represents one approach to this endeavor. Study findings support the feasibility, as well as provide some evidence for the efficacy of this approach for enhancing identification of those at greater risk of decline. Other modifications of MCI criteria could also be considered. One option might include disregarding baseline testing entirely to minimize the classification of those individuals who score poorly on initial assessment solely due to transient factors such as test anxiety associated with the novel assessment scenario. Another option, commonly employed in clinical assessment, is to compare domain-

specific cognitive test scores to measures such as the North American Adult Reading Test (Blair & Spreen, 1989) or the Wechsler Test of Adult Reading (The Psychological Corporation, 2001) which can be used to provide estimates of premorbid intellectual functioning. This method can be used to distinguish those who may be scoring poorly on current testing secondary to stable low baseline ability from those whose current poor performance likely represents cognitive decline. These, and other related modifications, may warrant future study.

Clinical Considerations

Clinical implications of Objective 1 results are somewhat constrained to the degree that the study evaluation procedures represent a less detailed assessment than would be typical of a full clinical neuropsychological evaluation. However, prior studies employing detailed assessment comparable to typical clinical evaluations have documented significant rates of reversion from MCI to unimpaired status across repeat assessment similar to reversion rates reported in epidemiological literature and the present study (de Jager & Budge, 2005; Devanand et al., 1997; Loewenstein et al., 2007), suggesting that there may be a considerable degree of overlap in factors that underlie classification instability in both the clinical and the research setting.

Replication of previously documented high rates of reversion from MCI to unimpaired status support the cautious application of the MCI concept in the clinical setting. The implications of an MCI diagnosis for an individual patient are potentially significant. Receiving a label understood to mean "at-risk" for dementia may lead to understandable reactions of grief and fear, and possibly adjustment-related depressive symptoms, which may cause undue harm to a patient given the substantial probability

that an individual who presents with MCI on an initial assessment will show cognitively normal performance on repeat evaluation. On the other hand, for an individual patient on a pre-dementia course, this label may offer validation of a perceived memory decline and afford this patient an opportunity to participate proactively in treatment decisions and disease management.

It is interesting to note that, across the five variations of MCI classification criteria employed in the present study, there were inconsistent but frequent findings of significant group differences between the Traditional and Persistent MCI groups and Controls in terms of self-rated memory and health. This finding, in the context of the relatively small net cognitive and functional decline observed in the sample and discussed below in Objectives 2 and 3, raises the possibility that the MCI groups in the present study may include older adults whose self-perceived suboptimal health and cognitive status may reflect psychological factors such as negative affect which may be associated with suboptimal performance on cognitive testing (Lezak et al., 2004) and consequent classification into an MCI group. It is noted, however, that the groups did not differ significantly on self-rated depression scores, though there was a non-significant trend to this effect.

An alternate interpretation of the above finding is that a number of individuals in the MCI groups had preserved insight into a declining trajectory of memory ability with a possible pre-dementia course. Although prior iterations of diagnostic criteria for MCI (i.e., Petersen et al., 1999) have specified preserved insight into cognitive decline as a diagnostic criterion; this is no longer the prevailing view (c.f., Winblad et al., 2004) as a

number of studies have shown declining insight into cognitive and functional status to be predictive of greater risk of decline (Albert et al., 1999; Tabert et al., 2002).

Aside from the group differences in self-rated memory and health, the Control groups also had significantly higher levels of education than the Persistent MCI group in the Basic, Severity, and Pervasiveness variations (with a trend to this effect across remaining variations) and the Traditional MCI group in the Severity variation. Because the psychometric cut-offs for cognitive impairment employed in the MCI classification algorithm were based on age- and education-referenced normative data designed to reduce bias towards identifying those as impaired whose poor performance on testing may reflect low educational attainment as opposed to cognitive decline, this finding is somewhat surprising. The concept of "cognitive reserve" may offer some insight into this finding. Recent evidence suggests that factors such as higher occupational and educational attainment, as well as increased engagement in intellectual, social, and physical activities, are associated with slower cognitive decline in older adults and may reduce the clinical manifestations of underlying neurodegenerative pathology by facilitating adaptive behaviors that may compensate for circumscribed cognitive decline (e.g., Scarmeas & Stern, 2004; Stern, 2006).

Taken together, clinical implications of the observed correlates of cognitive status group differences in Objective 1 offer some limited clues to the clinician who may be interested in clinical features which could inform impressions regarding whether an individual who presents with MCI on initial evaluation is likely to show persistent impairment on re-evaluation. That is, although the observed between-group differences were generally small and do not support confident prognostication at the individual level,

findings suggest that higher educational attainment may be associated with reduced risk of classification as Persistent MCI, while greater self-reported complaints regarding memory or health status may be associated with increased risk of classification as Persistent MCI.

Objective 2: External Validity of Persistent MCI Classification

In Study Objective 2, the external validity of the Basic Persistent MCI classification scheme was evaluated using five-year cognitive and functional outcome data. Hierarchical linear modeling was used to explore cognitive status group differences in baseline and longitudinal trajectory of performance on three cognitive and two functional outcome measures. Cognitive measures included tests of global cognitive status (Mini-Mental State Exam; MMSE), executive functioning (Trail Making Test Part B; Trails B), and visual recall (Digit Symbol Coding Recall). Functional measures included an assessment of global functional status (Activities of Daily Living; ADLs) and a measure of applied problem solving (Everyday Problems Test; EPT).

Key Findings

Two main hypotheses were specified in this section. First, as previous studies have demonstrated that those with Traditional MCI go on to develop dementia, a disorder characterized by impairment in cognitive and functional status, at higher rates relative to controls, Hypothesis 2a specified that both the Traditional MCI and Persistent MCI groups would show lower levels of performance and greater decline in both cognitive and functional status over five years relative to Controls.

In terms of cognitive status group differences in baseline performance, this hypothesis was almost universally supported. That is, for all outcome measures (with the exception of the ADL measure), the Traditional MCI group showed worse performance at baseline than did their respective Control group. Similarly, for all outcome measures, the Persistent MCI group showed poorer performance at baseline than did their respective Control group. It is noted that this finding is likely attributable to some degree to the small to medium correlations observed between the cognitive measures used in the MCI classification algorithm and the cognitive and functional outcome measures at baseline (see Table 3).

In terms of cognitive status group differences on rate of change over time, the findings were restricted to cognitive measures. The Traditional MCI group showed a greater rate of decline on the MMSE compared to Controls. The Persistent MCI group showed a greater rate of decline on the MMSE and Trails B measure compared to Controls. It is noted that, at the sample level, the amount of observed decline in the MMSE, Trails B, ADL, and EPT measures, while statistically significant, was of marginal clinical significance. Statistically significant decline was not observed on the Coding Recall measure. Retest effects, discussed above, may account in part for the relatively modest longitudinal decline across the sample.

To the extent that the Persistent MCI classification improves upon limitations in Traditional MCI classification, thereby more accurately capturing those at risk of dementia, Hypothesis 2b proposed that the magnitude of difference in baseline level of performance and longitudinal trajectory of decline in cognitive and functional status between those classified as Persistent MCI and Controls would exceed the magnitude of

difference in baseline level of performance and longitudinal trajectory of decline in cognitive and functional status between those classified as Traditional MCI and Controls. Descriptive comparisons showed this hypothesis to be universally supported. Implications of these findings are discussed below alongside Objective 3 results.

Objective 3: Robustness of Persistent MCI Classification

In Study Objective 3, the four previously outlined variations of the Persistent MCI classification scheme were employed to evaluate the robustness of the Persistent MCI versus Traditional MCI distinction under inclusion criteria specifying varying Severity, Duration, Pervasiveness, and Specificity of cognitive impairment required for inclusion in an MCI group. Hypotheses paralleled those outlined in Objective 2.

Key Findings

Results for analyses associated with Hypothesis 3a generally paralleled those for Hypothesis 2a outlined above. In terms of group differences in baseline performance, the Traditional MCI groups showed worse performance at baseline than did their respective Control group on all cognitive and functional outcome measures (with the exception of the MMSE in the Severity variation and the ADL in the Duration and Specificity variations). Similarly, the Persistent MCI groups showed poorer performance at baseline than did their respective Control group on all cognitive and functional outcome measures.

In terms of group differences in rate of change over time, significant findings were more frequent among the cognitive measures. Under each MCI inclusion variation, both the Traditional MCI and the Persistent MCI groups showed a greater rate of decline on the MMSE compared to Controls (except the Specificity variation where only the

Persistent MCI group showed greater rate of decline relative to Controls). Under the Severity, Pervasiveness, and Specificity variations, both the Traditional MCI and the Persistent MCI groups also showed a greater rate of decline on the Trails B measure compared to Controls. There were fewer differences associated with the functional outcome measures. Under the Duration variation, the Persistent MCI group showed greater decline relative to Controls on the EPT. Under the Specificity variation, the Persistent MCI showed greater decline relative to Controls on the ADL measure.

Results for Hypothesis 3b also generally paralleled those outlined in Objective 2. As expected, the descriptive comparison of the magnitude of differences in baseline level of performance and longitudinal trajectory of decline in cognitive and functional status showed that the magnitude of baseline and trajectory differences between those classified as Persistent MCI and Controls almost universally exceeded the magnitude of these differences between those classified as Traditional MCI and Controls.

Clinical Considerations

It is again noted that clinical implications of Objective 2 and 3 results are somewhat constrained to the degree that the current study evaluation procedures represent a less detailed assessment than would be typical of a full clinical neuropsychological evaluation. Nonetheless, findings of most prominent clinical relevance in Objective 2 and 3 include the evidence for a steeper trajectory of cognitive and functional decline for those identified as Persistent MCI compared to Traditional MCI (Hypotheses 2b and 3b), as well as the differing patterns of longitudinal change across the five MCI classification variations.

Though no appropriate statistical test is available to directly assess Hypothesis 2b and 3b, evidence from overall HLM model fit parameters consistently pointed towards best overall fit for Persistent MCI models compared to the Traditional MCI models and the simple Time in Study models. Moreover, descriptive comparison of specific model parameters revealed a near-universal pattern of greater discrepancy relative to Controls in terms of baseline status and trajectory of cognitive and functional decline in the Persistent MCI groups relative to Traditional MCI groups for each variation of the MCI classification evaluated, though these differences were generally small. The clinical significance of this overall trend should be noted, but not overstated. That is, the current findings argue that the clinician may wish to make note of the possibility of an increased risk of accelerated decline in the patient who presents with persistent cognitive impairment, but the current findings do not provide adequate data to allow prognostic interpretations at the level of the individual patient.

Qualitative examination of the overall pattern of findings across the five MCI classification variations under investigation in the present study yields further information of potential clinical relevance. Specifically, the relative difference in magnitude of decline in cognitive outcome measures (MMSE and Trails B) over time in the Persistent MCI group compared to Controls was largest under the Specificity variation of the MCI classification criteria. That is, compared to the Year 1 Control group, the Persistent Amnestic MCI group experienced a five-year decline in MMSE score of 2.22 standard deviation (SD) units and a decline in Trails B score of 6.39 SD units. In contrast, in the Basic MCI variation, the Persistent MCI group experienced a relative five-year decline in MMSE score of .93 SD units and a decline in Trails B score

of .76 SD units. Relative cognitive decline in the Persistent MCI group in the Duration, Severity, and Pervasiveness variations were intermediate to the Specificity and Basic variations, with the least severe decline in the Duration variation and the most severe decline in the Pervasiveness variation.

Plausible factors which may account for the different pattern of outcomes across MCI variations include the nature of the specific inclusion/exclusion criteria under each variation. The Basic variation, which had the highest base rate across the five conditions, represented the most liberal classification scheme in terms of (a) psychometric cut-off for impairment (i.e., 1.0 SD), (b) number of consecutive assessments required to establish persistent impairment (i.e., two assessments), and (c) number of tests considered in decision rules for establishing presence of impairment (i.e., performance below cut-off on any one of five benchmark cognitive tests). In contrast, the Specificity variation, which had the lowest base rate across the five conditions, employed a more strict decision rule in regards to the number of tests considered in establishing presence of impairment (i.e., performance below cut-off specifically in the benchmark memory test).

It is plausible that the liberal Basic variation led to the highest rates of false positives, while the Specificity scheme, which maps most closely to diagnostic criteria for Alzheimer's Disease, as well as to the pre-dementia MCI diagnostic criteria championed by some authors (Petersen, 2004), led to improved accuracy in identifying those most likely to show impending cognitive decline. The utility of the Specificity variation is further supported by the evidence of greatest functional decline in the Persistent Amnestic MCI group in the Specificity variation relative to Controls, compared to the smaller (and generally non-significant) rates of functional decline in the Persistent

MCI groups under the Basic, Severity, Duration, and Pervasiveness variation. It is also noted that the rates of reversion from cognitively impaired to non-impaired from Year 1 to Year 2 were highest in the Specificity variation (69%), and lowest in the Basic variation (33%), revealing an apparent trade-off between classification stringency and stability.

Taken together, clinical implications of these findings suggest that an individual who shows persistent memory impairment across serial assessments may be at greatest risk of decline. However, an individual who shows memory impairment at a single initial assessment may have a high likelihood of failing to show memory impairment at a follow-up session. Cautious application of this finding is warranted as the Persistent Amnestic MCI group in the Specificity condition was small in the current study (4% of the sample). Replication of these findings would be beneficial to further inform clinical practice.

Study Limitations

Limitations of the present study include the nature of the study sample, benchmark cognitive status evaluation procedures, choice of the five specific variations in inclusion/exclusion criteria for MCI classification schemes, decision rules within the MCI classification algorithms, and the selection of outcome measures.

Characteristics of the current sample which constrain generalizability of study findings include the relatively healthy, well-educated, high functioning nature of the all-volunteer community sample. It is likely that the current participant group undersampled those older adults with more significant health and cognitive concerns which may have

precluded study participation (including those conditions specifically identified as exclusion criteria, but also those conditions not specifically listed that would likely decrease interest in voluntary research participation). The use of local (age and education adjusted) normative data constrained the impact of this specific limitation on the present study regarding prevalence estimates of cognitive impairment. By using local normative data to compare the present study participants to a sample of demographically similar prior research participants, the ability to detect those who are showing cognitive performance below expected levels was enhanced.

Characteristics of the annual cognitive evaluation used to determine MCI status also represent a limitation of the present study. As detailed above, MCI status in the present study was determined based on performance across group-administered psychometric measures of perceptual speed, reasoning, episodic memory, verbal fluency, and vocabulary. This procedure differed from a more detailed, carefully controlled clinical evaluation in a number of ways. Clinical evaluation is typically performed in an individual face-to-face encounter which includes a clinical interview and history in addition to a tailored individually-administered battery of psychometric measures. Norm-referenced cognitive test scores would then be interpreted by a clinician in light of relevant history and behavioral observations before arriving at a clinical impression regarding the presence, nature, severity, and presumed etiology of any observed cognitive impairment. The current evaluation procedures, including group administration of the cognitive test battery, did not allow for a nuanced interpretation of test findings relative to individual patient characteristics.

Further, the test battery employed in the current evaluation was constrained to include only measures amenable to group administration. In some cases, such as the episodic memory measure, this meant using a paper and pencil version of a test which is typically administered orally in the clinical setting. These relatively minor adaptations of typical clinical assessment procedures may have had some implications for individual level performance. The group-testing format and the overall limitations inherent in longitudinal research studies also constrained the scope of the evaluation. Some measures typically administered in the clinical setting (e.g., measures of naming, visuoconstruction) were not included in the group test battery due to a lack of amenability to the group testing format.

The specific inclusion and exclusion criteria of the five MCI classification variations also represent an overall study limitation. These five specific MCI variations were chosen from a wide range of possible variations. Efforts were made to choose variations which captured common current research criteria for MCI; however, the relatively small sample size, combined with the relative infrequency of memory impairment in the current sample, constrained the ability to evaluate potential permutations of relevance. For example, there are a large number of studies which have employed a 1.5 SD impairment cutoff on a memory measure to define an Amnestic MCI group. Evaluation of such a condition with the Persistency modification was not possible in the current sample, because very few individuals in the current sample would meet criteria for Persistent MCI under such a variation.

Within each specific variation of the Traditional and Persistent MCI classification algorithms, a number of decision rules were applied which may have had some influence

on the overall pattern of findings. For example, the inclusion of an Unstable group under each of the five variations of Persistent MCI is a potential issue. It could be argued that a more parsimonious approach would be to forgo creation of an Unstable group and instead classify all those individuals who do not show persistent impairment as Controls. Given that the cognitive and functional performance and trajectory of decline in the Unstable groups tended to be intermediate to that of the Persistent MCI group and Controls, it is possible that merging the Unstable and the Control groups would have resulted in reduced detection of differences between the Persistent MCI group and Controls. The consistently intermediate nature of the Unstable group, however, could be used to argue in favor of its inclusion as a separate entity. That is, the overall pattern of findings is consistent with a continuum of cognitive risk; those in the Unstable group may well turn out to have rates of conversion to dementia intermediate to those of the Control and Persistent MCI groups. Future research could address this issue as well as the related issue of whether those who convert to Unstable from a Control group (i.e., show an apparent cognitive decline) have a different trajectory of outcome from those who convert to an Unstable from an MCI group (i.e., show an apparent cognitive improvement).

The final study limitation relates to the choice of outcome measures in the present project. Although it could be considered a strength of the study that there were multiple cognitive and functional outcome measures employed, the available measures represent a small sample of the possible outcome indicators that may be of relevance to a clinician. For example, a more detailed neuropsychological battery, incorporating tests of cognitive domains known to show decline in Alzheimer's disease (e.g., verbal recall, naming,

visuoconstruction), may have yielded informative findings. The existing outcome measures tapped important domains of functioning (e.g., global cognitive status, executive functioning, episodic visual recall, global functional status, everyday problem solving), but some of the specific measures were relatively gross indicators of functioning (i.e., Coding Recall, ADL) and this may have constrained ability to detect meaningful longitudinal change on these measures. The Trails B measure was also limited in that the available follow-up data spanned only a four-year study interval, while all other measures spanned a six-year interval.

Outcome variables of potential interest that were not considered in the present study due to limitations in available follow-up data include, most prominently, conversion to confirmed dementia diagnosis. A related alternative could be to use attrition as an outcome variable, as individuals who drop-out of longitudinal studies tend to have worse cognitive and health outcomes than those who remain in studies (e.g., MacDonald, Hultsch, & Dixon, 2003; Singer, Verhaeghen, Ghisletta, Lindenberger, & Baltes, 2003; Sliwinski, Hofer, Hall, Buschke, & Lipton, 2003). Verbal memory recall change and informant-rated functional status are also potential outcome variables of high relevance, as these tend to be areas of prominent clinical decline in the course of Alzheimer's disease and related dementias (Albert, 2008).

Conclusions and Future Directions

The present study was designed to address key limitations in current MCI classification procedures which tend to rely on single occasion assessment (Traditional MCI) by evaluating an alternate operational definition of MCI (and four variations of this

definition) requiring evidence of persistent cognitive impairment over multiple testing sessions (Persistent MCI). Results offer a replication and extension of a single prior study documenting substantial rates of reversion from Traditional MCI to unimpaired status at follow up testing. Retest effects, which may be generally overlooked in both the clinical setting and in prior studies of MCI, are posited to be a substantial contributor to this finding. Moreover, results suggest that, regardless of specific MCI variation under consideration, Persistent MCI groups show a poorer trajectory of cognitive, and to a lesser extent, functional decline over a five-year follow-up interval. There was some evidence that a variation of the Persistent MCI classification which mandated persistent memory impairment as an inclusion criteria achieved optimal prediction of cognitive and functional decline. Globally, results support the utility of the Persistent MCI concept as an avenue to enhance detection of dementia risk and suggest a number of avenues for future study, including replication and extension of the present findings using alternate outcome measures, as well as more formally accounting for instability of classification related to re-test effects in both the clinical setting and in future MCI studies.

References

Albert, M. (2008). Neuropsychology of the development of Alzheimer's disease. In F. I. M. Craik & T. A. Salthouse (Eds.), *Handbook of Aging and Cognition* (3 ed.). New York, NY: Psychology Press.

Albert, S. M., Michaels, K., Padilla, M., Pelton, G., Bell, K., Marder, K., et al. (1999). Functional significance of mild cognitive impairment in elderly patients without a dementia diagnosis. *American Journal of Geriatric Psychiatry, 7*(3), 213-220.

APA. (1994). *Diagnostic & Statistical Manual - IV*. Washington, DC: American Psychiatric Association.

Blackford, R. C., & la Rue, A. (1989). Criteria for diagnosing age-associated memory impairment: Proposed improvements from the field. *Developmental Neuropsychology, 5*(4), 295-306.

Blair, J. R., & Spreen, O. (1989). Predicting premorbid IQ: A revision of the National Adult Reading Test. *The Clinical Neuropsychologist, 3*, 129-136.

Brandt, J., Spencer, M., & Folstein, M. (1988). The Telephone Interview for Cognitive Status. *Neuropsychiatry, Neuropsychology, & Behavioral Neurology, 1*(2), 111-117.

Brooks, B. L., Iverson, G. L., Holdnack, J. A., & Feldman, H. H. (2008). Potential for misclassification of mild cognitive impairment: A study of memory scores on the Wechsler Memory Scale-III in healthy older adults. *Journal of the International Neuropsychological Society, 14*(3), 463-478.

Brooks, B. L., Iverson, G. L., & White, T. (2007). Substantial risk of 'accidental MCI' in healthy older adults: Base rates of low memory scores in neuropsychological

assessment. *Journal of the International Neuropsychological Society, 13*(3), 490-500.

Bruscoli, M., & Lovestone, S. (2004). Is MCI really just early dementia? A systematic review of conversion studies. *International Psychogeriatrics, 16*(2), 129-140.

Busse, A., Hensel, A., Guhne, U., Angermeyer, M. C., & Riedel-Heller, S. G. (2006). Mild cognitive impairment: long-term course of four clinical subtypes. *Neurology, 67*(12), 2176-2185.

Collie, A., Maruff, P., & Currie, J. (2002). Behavioral characterization of mild cognitive impairment. *Journal of Clinical & Experimental Neuropsychology, 24*(6), 720-733.

Crook, T., Bartus, R. T., Ferris, S. H., & Whitehouse, P. (1986). Age-associated memory impairment: Proposed diagnostic criteria and measures of clinical change: Report of a National Institute of Mental Health work group. *Developmental Neuropsychology, 2*(4), 261-276.

Darby, D., Maruff, P., Collie, A., & McStephen, M. (2002). Mild cognitive impairment can be detected by multiple assessments in a single day. *Neurology, 59*(7), 1042-1046.

Davis, H. S., & Rockwood, K. (2004). Conceptualization of mild cognitive impairment: A review. *International Journal of Geriatric Psychiatry, 19*(4), 313-319.

de Jager, C. A., & Budge, M. M. (2005). Stability and predictability of the classification of mild cognitive impairment as assessed by episodic memory test performance over time. *Neurocase, 11*(1), 72-79.

de Rotrou, J., Wenisch, E., Chausson, C., Dray, F., Faucounau, V., & Rigaud, A. S. (2005). Accidental MCI in healthy subjects: A prospective longitudinal study. *European Journal of Neurology, 12*(11), 879-885.

Devanand, D. P., Folz, M., Gorlyn, M., & Moeller, J. R. (1997). Questionable dementia: Clinical course and predictors of outcome. *Journal of the American Geriatrics Society, 45*(3), 321-328.

Dixon, R. A., & de Frias, C. M. (2004). The Victoria Longitudinal Study: From characterizing cognitive aging to illustrating changes in memory compensation. *Aging, Neuropsychology, & Cognition, 11*, 346-376.

Dubois, B., Feldman, H. H., Jacova, C., Dekosky, S. T., Barberger-Gateau, P., Cummings, J., et al. (2007). Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *Lancet Neurology, 6*(8), 734-746.

Ebly, E. M., Hogan, D. B., & Parhad, I. M. (1995). Cognitive impairment in the nondemented elderly: Results from the Canadian Study of Health and Aging. *Archives of Neurology, 52*(6), 612-619.

Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.

Ferrer, E., Salthouse, T. A., Stewart, W. F., & Schwartz, B. S. (2004). Modeling age and retest processes in longitudinal studies of cognitive abilities. *Psychology and Aging, 19*(2), 243-259.

Fisk, J. D., Merry, H. R., & Rockwood, K. (2003). Variations in case definition affect prevalence but not outcomes of mild cognitive impairment. *Neurology, 61*(9), 1179-1184.

Flicker, C., Ferris, S. H., & Reisberg, B. (1991). Mild cognitive impairment in the elderly: Predictors of dementia. *Neurology, 41*(7), 1006-1009.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*(3), 189-198.

Ganguli, M., Dodge, H. H., Shen, C., & DeKosky, S. T. (2004). Mild cognitive impairment, amnestic type: An epidemiologic study. *Neurology, 63*(1), 115-121.

Gilewski, M. J., Zelinski, E. M., & Schaie, K. W. (1990). The Memory Functioning Questionnaire for assessment of memory complaints in adulthood and old age. *Psychology and Aging, 5*(4), 482-490.

Goodglass, H., & Kaplan, E. (1979). Assessment of cognitive deficit in the brain-injured patient. In M. S. Gazzaniga (Ed.), *Handbook of behavioral neurobiology: Vol. 2. Neuropsychology* (pp. 3-22). New York: Plenum Press.

Graham, J. E., Rockwood, K., Beattie, B. L., Eastwood, R., Gauthier, S., Tuokko, H., et al. (1997). Prevalence and severity of cognitive impairment with and without dementia in an elderly population. *The Lancet, 349*(9068), 1793-1796.

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtiss, G. (1993). *Wisconsin card sorting test manual: Revised and expanded*. Odessa, FL: Psychological Assessment Resources.

Heaton, R. K., Grant, I., & Matthews, C. G. (1991). *Comprehensive norms for an expanded Halstead-Reitan battery: Demographic corrections, research findings, and clinical applications.* Odessa, FL: Psychological Assessment Resources.

Hertzog, C., Van Alstine, J., Usala, P. D., & Hultsch, D. F. (1990). Measurement properties of the Center for Epidemiological Studies Depression Scale (CES-D) in older populations. *Psychological Assessment, 2*(1), 64-72.

Hultsch, D. F., Hertzog, C., & Dixon, R. A. (1990). Ability correlates of memory performance in adulthood and aging. *Psychology and Aging, 5*(3), 356-368.

Hultsch, D. F., Hertzog, C., Dixon, R. A., & Small, B. J. (1998). *Memory change in the aged*. Cambridge: Cambridge University Press.

Ivnik, R. J., Smith, G. E., Lucas, J. A., Petersen, R. C., Boeve, B. F., Kokmen, E., et al. (1999). Testing normal older people three or four times at 1- to 2-year intervals: defining normal variance. *Neuropsychology, 13*(1), 121-127.

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *The Boston Naming Test* (2nd ed.). Philadelphia: Lea & Febiger.

Kral, V. A. (1962). Senescent forgetfulness: benign and malignant. *Canadian Medical Association Journal, 86*, 257-260.

Larrieu, S., Letenneur, L., Orgogozo, J. M., Fabrigoule, C., Amieva, H., Le Carret, N., et al. (2002). Incidence and outcome of mild cognitive impairment in a population-based prospective cohort. *Neurology, 59*(10), 1594-1599.

Levy, R. (1994). Aging-associated cognitive decline. *International Psychogeriatrics, 6*(1), 63-68.

Lezak, M. D. (1995). *Neuropsychological Assessment* (3rd ed.). New York: Oxford University Press.

Lezak, M. D., Howieson, D. B., & Loring, D. W. (2004). *Neuropsychological Assessment* (4th ed.). New York: Oxford University Press.

Loewenstein, D. A., Acevedo, A., Agron, J., & Duara, R. (2007). Stability of neurocognitive impairment in different subtypes of mild cognitive impairment. *Dementia And Geriatric Cognitive Disorders, 23*(2), 82-86.

Luis, C. A., Loewenstein, D. A., Acevedo, A., Barker, W. W., & Duara, R. (2003). Mild cognitive impairment: Directions for future research. *Neurology, 61*(4), 438-444.

MacDonald, S. W. S., Hultsch, D. F., & Dixon, R. A. (2003). Performance variability is related to change in cognition: Evidence from the Victoria Longitudinal Study. *Psychology and Aging, 18*(3), 510-523.

McKhann, G., Drachman, D., Folstein, M. F., Katzman, R., Price, D., & Stadlan, E. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA work group under the auspices of Department of Health and Human Services Task Force. *Neurology, 34*, 939-944.

Palmer, B. W., Boone, K. B., Lesser, I. M., & Wohl, M. A. (1998). Base rates of "impaired" neuropsychological test performance among healthy older adults. *Archives of Clinical Neuropsychology, 13*(6), 503-511.

Petersen, R. C. (2003). Conceptual overview. In R. C. Petersen (Ed.), *Mild cognitive impairment: Aging to Alzheimer's disease.* (pp. 1-14). New York, NY, US: Oxford University Press.

Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine, 256*(3), 183-194.

Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., & Kokmen, E. (1999). Mild cognitive impairment: Clinical characterization and outcome. *Archives of Neurology, 56*(3), 303-308.

Rabbitt, P., Diggle, P., Smith, D., Holland, F., & Innes, L. M. (2001). Identifying and separating the effects of practice and of cognitive ageing during a large longitudinal study of elderly community residents. *Neuropsychologia, 39*(5), 532-543.

Radloff, L. S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1*(3), 385-401.

Raudenbush, S. W., & Bryk, A. S. (2002). Applications in the study of individual change. In *Hierarchical Linear Models* (2 ed., pp. 160-205). Thousand Oaks: Sage Publications.

Reitan, R. M., & Wolfson, D. (1985). *The Halsted-Reitan Neuropsychological Test Battery*. Tuscon, AZ: Neuropsychological Press.

Ritchie, K., Artero, S., & Touchon, J. (2001). Classification criteria for mild cognitive impairment: A population-based validation study. *Neurology, 56*(1), 37-42.

Rodgers, W. L., & Miller, B. (1997). A comparative analysis of ADL questions in surveys of older people. *Journals of Gerontology: Series B: Psychological Sciences & Social Sciences, 52*, 21-36.

Ronnlund, M., Nyberg, L., Backman, L., & Nilsson, L. (2005). Stability, Growth, and Decline in Adult Life Span Development of Declarative Memory: Cross-Sectional and Longitudinal Data From a Population-Based Study. *Psychology and Aging, 20*(1), 3-18.

Salthouse, T. A. (2009). When does age-related cognitive decline begin? *Neurobiology of Aging*, in press.

Salthouse, T. A., & Tucker-Drob, E. M. (2008). Implications of short-term retest effects for the interpretation of longitudinal change. *Neuropsychology, 22*(6), 800-811.

Scarmeas, N., & Stern, Y. (2004). Cognitive reserve: implications for diagnosis and prevention of Alzheimer's disease. *Current Neurology and Neuroscience Reports, 4*(5), 374-380.

Singer, T., Verhaeghen, P., Ghisletta, P., Lindenberger, U., & Baltes, P. B. (2003). The fate of cognition in very old age: Six-year longitudinal findings in the Berlin Aging Study (BASE). *Psychology and Aging, 18*(2), 318-331.

Sliwinski, M. J., Hofer, S. M., Hall, C., Buschke, H., & Lipton, R. B. (2003). Modeling memory decline in older adults: The importance of preclinical dementia, attrition, and chronological age: Special Section: Applied Longitudinal Methods in Aging Research Section Editors, Leah L. Light and Christopher Hertzog. *Psychology and Aging, 18*(4), 658-671.

Smith, G. E., Petersen, R. C., Parisi, J. E., & Ivnik, R. J. (1996). Definition, course, and outcome of mild cognitive impairment. *Aging, Neuropsychology, and Cognition, 3*(2), 141-147.

Stern, R. A., & White, T. (2003). *Neuropsychological Assessment Battery (NAB)*. Lutz, FL.

Stern, Y. (2006). Cognitive reserve and Alzheimer disease. *Alzheimer Disease and Associated Disorders, 20*(2), 112-117.

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary*. New York: Oxford University Press.

Stuss, D. T., Kaplan, E. F., Benson, D. F., Weir, W. S., Chiulla, S., & Sarazin, F. F. (1982). Evidence for involvement of orbitofrontal cortex in memory functions: An interference effect. *Journal of Comparative and Physiological Psychology, 96*(6), 913-925.

Tabert, M. H., Albert, S. M., Borukhova-Milov, L., Camacho, Y., Pelton, G., Liu, X., et al. (2002). Functional deficits in patients with mild cognitive impairment: Prediction of AD. *Neurology, 58*(5), 758-764.

The Psychological Corporation. (2001). *Wechsler Test of Adult Reading*. San Antoinio, TX: Harcourt Assessment.

Thorvaldsson, V., Hofer, S. M., Berg, S., & Johansson, B. (2006). Effects of Repeated Testing in a Longitudinal Age-Homogeneous Study of Cognitive Aging. *Journals of Gerontology: Series B: Psychological Sciences and Social Sciences*(6), P348-p354.

Thurstone, T. G. (1962). *Primary mental abilities: grades 9±12 (1962 revision)*. Chicago, IL: Science Research Associates.

Tuokko, H., Frerichs, R., Graham, J., Rockwood, K., Kristjansson, B., Fisk, J., et al. (2003). Five-year follow-up of cognitive impairment with no dementia. *Archives of Neurology, 60*(4), 577-582.

Tuokko, H., & McDowell, I. (2006). An overview of mild cognitive impairment. In H. Tuokko & D. F. Hultsch (Eds.), *Mild cognitive impairment: International perspectives*. New York: Psychology Press.

Warrington, E. K. (1984). *Recognition Memory Test*. Windsor, UK: NFER-Nelson.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale - Revised*. San Antonio: The Psychological Corporation.

Wechsler, D. (1987). *Wechsler Memory Scale - Revised*. New York: Psychological Corporation.

Wechsler, D. (1997). *Wechsler Memory Scale - Third Edition*. San Antonio, TX: The Psychological Corporation.

WHO. (1993). *The ICD-10 classification of mental and behavioral disorders: diagnostic criteria for research*. Geneva: World Health Organization.

Willis, S. L., & Marsiske, M. (1993). *Manual for the Everyday Problems Test*. University Park: Pennsylvania State University.

Winblad, B., Palmer, K., Kivipelto, M., Jelic, V., Fratiglioni, L., Wahlund, L. O., et al. (2004). Mild cognitive impairment--beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. *Journal of Internal Medicine, 256*(3), 240-246.

Zaudig, M. (1992). A new systematic method of measurement and diagnosis of 'mild cognitive impairment' and dementia according to ICD-10 and DSM-III--R criteria. *International Psychogeriatrics, 4*(2), 203-219.