PREDICTING CONVERSION FROM MILD COGNITIVE IMPAIRMENT TO ALZHEIMER’S DISEASE USING PARTIALLY ORDERED MODELS OF NEUROPSYCHOLOGICAL MEASUREMENTS

By

YAN YANG

Submitted in partial fulfillment of the requirements

For the degree of Master of Science

Clinical Research Scholars Program

School of Medicine

CASE WESTERN RESERVE UNIVERSITY

August, 2011
We hereby approve the thesis/dissertation of Yan Yang candidate
for the Master of Science (M.S.) degree *

(Signed) Douglas Einstadter, MD (chair of the committee)

Alan J. Lerner, MD

Curtis Tatsuoka, Ph.D

(date) June 30, 2011

*We also certify that written approval has been obtained for any proprietary material contained therein.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>2</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>2</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>4</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>5</td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>6</td>
</tr>
<tr>
<td>1.1 Alzheimer’s disease</td>
<td>7</td>
</tr>
<tr>
<td>1.2 Mild Cognitive Impairment</td>
<td>9</td>
</tr>
<tr>
<td>1.3 Neuropsychological examination in the predicting progression of mild</td>
<td>12</td>
</tr>
<tr>
<td>cognitive impairment to Alzheimer’s disease</td>
<td></td>
</tr>
<tr>
<td>1.4 Partially ordered sets (posets) model in predicting progression of</td>
<td>16</td>
</tr>
<tr>
<td>mental illness</td>
<td></td>
</tr>
<tr>
<td>2. SUBJECTS AND METHODS</td>
<td>17</td>
</tr>
<tr>
<td>2.1 Subjects and study design</td>
<td>17</td>
</tr>
<tr>
<td>2.2 Clinic Neuropsychological measurement and APOE data</td>
<td>19</td>
</tr>
<tr>
<td>2.3 Poset Models and statistical Methods</td>
<td>20</td>
</tr>
<tr>
<td>2.3.1 Partially ordered sets as models of NP functioning</td>
<td>20</td>
</tr>
<tr>
<td>2.3.2 Linking NP functions to NP measures and the generation of partially ordered models</td>
<td>23</td>
</tr>
<tr>
<td>3. RESULTS</td>
<td>33</td>
</tr>
<tr>
<td>4. DISCUSSION</td>
<td>41</td>
</tr>
<tr>
<td>5. APPENDIX</td>
<td>45</td>
</tr>
</tbody>
</table>
5. REFERENCES
List of Tables:

Table 1: Demographic of MCI subjects with baseline functional and neuropsychological measures ..........................................................25

Table 2: Selected ADCS NP measures and their relation to cognitive functions.................................................................................26

Table 3: The poset consists of a collection of cognitive states that are associated with profiles of functionality ..................................................27

Table 4: Conversion of MCI to AD with/without APOE e4 allele in 24 months....35

Table 5: The effect of episodic memory level 2 on conversion of AD..............35

Table 6: Episodic memory level 2 by diagnosis Alzheimer’s disease stratified by APOE e4 status .................................................................37

Table 7: Episodic memory level 3 functioning by AD diagnosis by stratified by APOE e4 status .................................................................39

Table 8: Decreased Neuropsychological measures of MCI subjects at baseline after 24 months ........................................................................40
List of Figures:

Figure 1: Sample Poset model for cognitive functioning ..........................21

Figure 2: Largest Posterior Probability of State Membership Values ............21

Figure 3: Sum of the largest and Second largest Posterior probability ..........22

Figure 4: Poset representation of ADCS NP functioning profiles ................23

Figure 5: ROC curve shown with all cognitive function ..............................30

Figure 6: ROC curve shown without medium episodic memory level 2 and high episodic memory level 3 cognitive function .................................31
Acknowledgements

I would like to thank Dr. Douglas Einstadter, Dr. Alan Lerner and Dr. Curtis Tatsuoka for serving on my committee. I thank them for their mentoring and support of this study. Without their guidance, I would not be able to complete the study. I would like to thank Dr. Robert Miller and Dr. Anthony Furlan for their generous support. Without their support the study could not have been performed. My thanks also go to Mrs. Natalie Milone and Dr. James Spilsbury for their consistent help during the period of my CRSP studying.

My thanks also go to my husband and my dear daughter for their spirit supporting and nurturing me while I spent numerous hours at home to complete this study. I thank all my friends for their supports and understanding as well.
Predicting Conversion from Mild Cognitive Impairment to Alzheimer’s Disease using Partially Ordered Models of Neuropsychological Measurements

Abstract

By

YAN YANG

Background: Mild cognitive impairment (MCI) is a risk factor for conversion to Alzheimer’s disease (AD).

Objectives: To identify predictors of the conversion of MCI to AD.

Methods: Data from Alzheimer’s Disease Cooperative study (ADCS) MCI patients. Cognitive functionality profiles were identified using partially ordered sets models (posets) for each subjects, and linked to AD conversion, stratified by APOE-e4 status.

Results: Low episodic memory (LEM) (RR=3.7), APOE-e4 (RR2.7) and the higher percentage rate of LEM and APOE-e4 were strongly associated with conversion of MCI to AD. The effect of LEM on AD was modified by APOE-e4: LEM alone RR+4.5; LEM and PAPOE-e4 combined RR=2.8.

Conclusion: Among MCI subjects, LEM is a strong predictor of conversion to AD.
APOE-e4 modifies the effect of LEM. Posets allows finer stratification of MCI patients and may be help in defining more heterogeneity of cognitive functionality among MCI for early-stage studies.

**Keywords:** Alzheimer’s disease; Mild Cognitive Impairment, neuropsychological functioning, conversion prediction; Partially Ordered Sets.
1. INTRODUCTION

According to the United Nations (UN, 2007) and the United States Centers for Disease Control and Prevention (CDC, 2003), the number of elderly in the population is increasing worldwide due to the combination of declining fertility and increasing life expectancies. As result of this trend, there has been a significant increasing number of people suffering from age-related forms of dementia such as Alzheimer’s disease (AD). In the United States alone the number of individuals with AD is approximately 5.3 million (Hebert et al., 2003), and this number is expected to increase dramatically in the near future due to the number of aging baby boomers (CDC, 2007, USA; (Kinsella et al., 2009). The increasing population of AD has many effects on both the individual level and on society as a whole. AD, it is a major causes of physical disability, institutionalization, and decreased quality of life and significantly shortens life expectancy among the elderly (Qiu et al., 2009). The marked increase in the number of dementia patients will also result in an increased economic burden and demand on social systems and health care-related institutions.

In the last decade, focus has increasingly shifted to accurate detection of the earliest phase of AD. Mild cognitive impairment (MCI) has been applied to represent a transitional state between cognitive changes of healthy aging and very early dementia. MCI is becoming increasingly recognized as a risk factor for developing AD. Numerous epidemiological studies have documented the accelerated rate of progression of MCI to AD subjects, and certain identified predictor variables of neuropsychological and genetic information that are critical
for the conversion of MCI to AD. However, studies have also suggested that not all MCI patients will progress to AD (Petersen, 2004) and hence there is a need to develop a more accurate, precise and informative method to identify those with MCI who are likely to convert to AD. Early identification of MCI patients who will will not convert to AD is essential for timely administration of pharmacologic and therapeutic interventions, as well as to potentially slow the progression to AD.

In this study, we have applied an established methodology - Partially Ordered Sets (posets) Models to analyze the data of patients with MCI based on cognitive/neuropsychological examination who participate in the Alzheimer’s Disease Cooperative study (ADCS). Based on our results, we suggest that a posets model can provide a more precise understanding of how specific combinations of cognitive functions associate with the progression of MCI to AD.

1.1 Alzheimer’s disease

Alzheimer’s disease (AD) is a complex, chronically progressive and devastating neurodegenerative disease. AD is the most common form of late-life dementia. With increasing age, the incidence and prevalence of dementia increase markedly. The prevalence of clinically manifest AD is about 2% at the age of 65 years but increases to about 30% at the age of 85 years (Wimo et al., 1997, Winblad et al., 1997). AD presents a spectrum of significant clinical symptom that includes deficits in the ability to form recent memories, executive dysfunction, impairment of cognition and other abnormal behavioral symptoms. There are two basic
classifications of AD: familial AD and sporadic late onset AD (LOAD).
Familial AD has a relatively early age of onset, typically before the age of 65 and it accounts for about 5% of all AD cases. Familial AD is nearly always related to genetic mutation caused by a single autosomal dominant mutation, such as APP (chromosome 21), PS1 (chromosome 14), or PS2 (chromosome 1) with almost 100% penetrance. Nearly 95% of the patients with AD suffer from LOAD. The onset of LOAD is typically after the age of 65. A number of basic biomedical research studies have been undertaken with the goal of determining the mechanism of biochemical pathway leading to AD pathology and with a focus on AD pathological features: Aβ deposits, neurofibrillary tangles and neuronal cell death (Neve and Robakis, 1998, Selkoe, 2002). Biomedical research has shown that the allele e4 of apolipoprotein E (APOE) is a risk factor for the development of AD (Farrer et al., 1997, Lautenschlager et al., 1999). But the association between the APOE e4 allele and MCI has shown inconsistent results. Some studies have shown that APOE e4 accelerates MCI conversion rates (Forstl et al., 1995, Zill et al., 2001, Traykov et al., 2002), while others have found no association between APOE e4 and MCI progression (Bartres-Faz et al., 2001, Collie et al., 2002, Devanand et al., 2005).
Clinical diagnosis of early and mild clinical stage AD is challenging since dementia symptoms are not yet fully expressed. The gold standard criteria of diagnosis for AD are neuro-pathological findings in the brain, characterized by abnormal deposits of Aβ peptide and neurofibrillary tangles and the loss of neuronal numbers. The Aβ consists of an abnormal cleavage product (A-beta) of a
membrane protein known as APP. Neurofibrillary tangles are the aggregates of hyperphosphorylated microtuble associated protein named tau. The density of neuronal cells is significantly reduced in various brain areas, such as frontal, entorhinal, hippocampal cortex and other limbic areas (Lyness et al., 2003, Zarow et al., 2003) (Whitehouse et al., 1982), (Zweig et al., 1988, Chen et al., 2000, Yang et al., 2001). Studies have shown evidences that post-mitotic neuronal cell death is associated with various pathological mechanisms, such as Aβ toxicity, pathological tau formed tangle induced neuronal death, as well as cell cycle reentry related neuronal cell death in both MCI and late stage AD (Yang et al., 2001, Yang et al., 2003). To better understand the mechanism of AD pathology, several transgenic mouse models of AD that mimic human AD with Aβ or neurofibrillary tangles have been generated through recreation of the genetic changes found in familial AD (Lamb, 1995, Holcomb et al., 1998; Hsiao et al., 1996, Sturchler-Pierrat et al., 1997, Oddo et al., 2003). These APP transgenic mouse models often exhibit an age-related development of diffuse and neuritic plaques, with plaque severity burdens equivalent to those found in advanced cases of AD. Many of these models have also been shown to have significant memory deficits (Hsiao et al., 1996, Holcomb et al., 1998, Oddo et al., 2003)These rodent models have been proven to be a valuable resource in the exploration and design of disease therapies (Schenk, 2002).

1.2 Mild Cognitive Impairment
Mild cognitive impairment (MCI) is a transitional state between the cognitive

The clinical criteria for diagnosis of MCI is: 1) evidence of memory impairment (memory complaint, preferably corroborated by a caregiver); 2) preservation of general cognitive function; 3) normal activities of daily living; 4) abnormal memory on cognitive testing (generally performance ≥1.5 SD below age and education-adjusted scores); and 5) the absence of dementia (base on clinical judgment). MCI has been classified into two main subtypes: amnestic MCI (aMCI) and non-amnestic MCI (naMCI). aMCI is primarily characterized by impaired memory with relatively unimpaired or less-impaired functioning in other cognitive domains. aMCI impairment has a primary memory component, either alone (single domain) or in conjunction with other cognitive-domain impairments (multiple domain), but it is insufficient severity to constitute dementia (Petersen et al., 2001a, Petersen et al., 2001b, Lopez et al., 2003, Ganguli et al., 2004, Petersen, 2004). Non-amnestic MCI (naMCI) is characterized by impairment in one or more cognitive domains other than memory, including language, attention, executive functioning, etc. naMCI can present initially as a primary impairment in other cognitive domains including language, visuospatial or visuoperceptual abilities, executive function, or even affect (Mapstone et al., 2003).

Previous research has shown that the rate of progression to clinically diagnosable Alzheimer’s disease is 10% to 15% per year among persons who meet the criteria for the amnestic form of MCI, compared to a rate of 1% to 2% per year among
normal elderly persons (Petersen et al., 1999). Within 6 years, about 80% of those people who meet the criteria or amnestic MCI will progress to Alzheimer’s disease. The presence of one or more apolipoprotein (APOE) ε4 alleles is associated with a more rapid rate of progression (Petersen et al., 1995, Tierney et al., 1996a, Tierney et al., 1996b; (Bowen et al., 1997, Devanand et al., 1997, Petersen et al., 1999, Petersen et al., 2001b). Epidemiological studies (Fratiglioni et al., 1992, Petersen, 2004) suggest that amnestic MCI patients (such as patients with scores >1.5 SDs below age-adjusted norms on standard memory tests) with or without deficits in other cognitive domains, constitute subtypes that are most likely to convert to AD. Non-amnestic single or multiple-domain subtypes are more likely to convert to other dementias, such as vascular dementia.

It has been recognized however that significant heterogeneity exist among MCI subgroups. For instance, some have used the term amnestic multi-domain MCI, which reflects that some MCI subjects have cognitive deficits beyond memory-related ones. However, even within this subgroup, there is still more precise description is needed, as there is no differentiation as to what other cognitive functions are impaired. The presence of multiple domain MCI makes early diagnosis more complicated, as there are numerous disorders that may cause subtle cognitive deficits in multiple domains.

Although many investigations have studied different technologies for the detection of MCI, such as using biomarkers (e.g. CSF-tau, Aβ1-42), primary MRI imaging and neuropsychological measurements (Ewers et al., 2010a, Ewers et al., 2010b,
Ewers et al., 2010c, Teipel et al., 2010). A variety of measurements have been individually linked to decline in mild cognitive impairment (MCI), but the identification of optimal markers for predicting disease progression remains unresolved. Data have been reported in predicting progression of different types of MCI to dementia. Studies suggest that MCI with multiple domains has a higher risk of progression to dementia, others have reported that pure aMCI is more likely to progress to AD (Yaffe et al., 2006). Rozzini L., et al has reported that the Alzheimer Disease assessment Scale-Cognitive Subscale (ADAS-Cog) is a useful tool for screening participants at risk of developing AD among populations with aMCI (Rozzini et al., 2008a, Rozzini et al., 2008b). Recent studies by Chapman et al., (Chapman et al., 2010, Chapman et al., 2011) have also analyzed neuropsychological testing in MCI, using a multivariate linear model to predict conversion rate to AD in a longitudinal study, Chapman et al., found that episodic memory, speeded executive functioning, recognition memory (false and true positives), visuospatial memory processing speed, and visuospatial episodic memory together strongly predicted the conversion to AD.

To summary among MCI subgroups, there is increasing evidence recognizing the presence of heterogeneity. Even within MCI subgroups, more precise classification of cognitive functions is still needed. Episodic memory is the first and most severely affected cognitive domain in Alzheimer's disease (AD), and it is also the key early marker in amnestic MCI.

1.3 Neuropsychological (NP) examination in predicting the progression of MCI to AD
Increasing interest has been devoted to recognizing and preventing the progression of MCI to AD, as MCI is considered a prodromal state of AD. Treatments proposed for AD, such as anti-oxidants (Vitamin E) and cholinesterase inhibitors, may also have beneficial effects on improving cognition or delaying progression in MCI (Irizarry and Hyman, 2001, Petersen et al., 2005), (Grundman, 2000) (Grundman and Delaney, 2002, Petersen, 2003). However, detecting treatment effects that alter progression in clinical studies based on MCI has been difficult (Andrieu et al., 2009). Thus, a precise clinic diagnosis of MCI is critical.

Although a variety of neuropsychological (NP) tests have been applied to predict the progression of MCI to AD, the neurocognitive deficits that define the natural history of MCI to AD remain poorly understood. Clinical NP examinations are usually employed to determine whether a patient has neurocognitive deficits and, where present, to determine whether deficits selectively affect certain neurocognitive function relative to others. The latter process involves expert interpretation of the profile of NP test scores as well as patterns of responding within particular tests. Some patterns have diagnostic significance in that they are associated with focal damage to particular brain regions, or to degenerative processes having a well-documented course/sequence of cognitive deterioration. Cognitive profiles may also have rehabilitative value, where cognitive strengths may be engaged to offset weaknesses, thereby improving the patient’s functionality. A number of studies have applied neuropsychological tests to predict the conversion rate of MCI to AD. A major challenge is that conventional NP
measures are often polyfactorial, meaning the observed performance level on any single task is influenced by multiple cognitive functions and domains. Generally, single tests cannot be used to measure a single cognitive operation. Hence, it is difficult to pinpoint exactly for which cognitive operations a low scoring subject may have poor functionality.

Although there has been an increasing number of published studies examining NP test performance using statistical methods in mental illness, one reason for the lack of precision in these cognitive characterizations is that NP response data are complex making analysis difficult. In practice, it can be difficult to assess in isolation the integrity of a particular cognitive function, as most NP measures require tapping into several cognitive functions in order to perform well. It is often not possible to design single tests that require one function to the exclusion of all others. This is particularly true for executive functions, which are viewed as controlling other functions. Hence, for a subject who performs poorly on a given measure, it may be difficult to pinpoint exactly which of the associated functions are impaired. A major limitation of scale-based approaches is exactly this assumption of a direct correspondence between a measure and an associated function. An observed poor performance will penalize the subscale score for the associated function, even if it is another of the functions that is impaired. Linking performance on NP tests to specific cognitive functions can thus be problematic.

One of the limitation of previous studies predicting progression from MCI to AD is
that they have generally focused on only one or two of the following three risk categories: 1) demographic factors such as age, sex, and education; 2) Neurocognitive performance; 3) biologic factors, such as the presence of the APOE4 allele. Another limitation to many previous studies is the heterogeneity in how the pre-dementia diagnosis is defined and not using formal criteria for determining MCI. MCI has multiple heterogeneous attributes, including clinical presentation, etiology, and outcome (Mariani et al., 2007, Monastero et al., 2007, Palmer et al., 2007), therefore the detection of treatment effects that alter progression in clinical studies in MCI samples is difficult (Andrieu et al., 2009).

MCI subjects are also heterogeneous in terms of conversion outcomes, making the study of conversion to AD complicated. Hence, it would be useful to better understand how this heterogeneity arises, and to carefully characterize the multiple factors in neuropsychological measurement, by which MCI subjects differ in terms of progression outcomes.

Studies from several groups have shown several neuropsychological measures that may serve as predictors of MCI conversion: a) percent from immediate to delayed recall on the Buschke Selective Reminding Test (SRT) (Buschke and Fuld, 1974); b) Wechsler Memory Scale (WMS) visual reproduction subtest (WMSVR); c) performance on the Wechsler Adult Intelligence Scale–Revised (WAIS-R) Digit Symbol Test (Wechsler D. Wechsler Adult Intelligence Scale–Revised, 1981); d) confrontational naming on the Boston Naming Test (BNT) e) and category fluency on the Animal Naming Test (ANT) (Jacobs et al., 1995, Tierney et al., 1996a, Devanand et al., 1997). ADAS-cog (Alzheimer Disease
Assessment Scale-cognitive subscale) was also designed to measure the severity of the most important symptoms of Alzheimer’s disease (AD). Its subscale of ADAS-cog is the most popular cognitive testing instrument used in clinical trials. It consists of 11 tasks measuring the disturbances of memory, language, praxis, attention and other cognitive abilities, which are often referred to as the core symptoms of AD. A study by Rozzini et al (Rozzini et al., 2008b) has reported using ADAS-cog for screening participants at risk of developing AD among aMCI populations suggesting aMCI who progressed to AD were characterized at baseline by greater cognitive impairment. The ADAS-Cog subscale is a useful and brief cognitive assessment tool to screen a MCI participants converting to AD. Other NP measurements are linked in the Appendix.

1.4 Partially ordered sets (posets) model in predicting progression of mental illness

In this study, we report on an application of statistical cognitive modeling methods that can systematically, accurately, and efficiently diagnose cognitive functioning in a manner that more closely approaches that of an expert neuropsychologist than currently available methods. The methods we will describe are based on finite partially ordered sets (posets) as classification models (Tatsuoka, et al, 2002, 2003). Posets models are useful for many statistic applications, including cognitive modeling since the cognitive deficit profile associated with a given state (a group patients share the same pattern of attribute strengths and weaknesses or same cognitive deficit profile) may differ from that of another state both qualitatively and
also in terms of its overall severity. Posets are comprised of states, into which cases are classified, that are associated with distinct patterns of attribute strengths and weaknesses. Posets can represent these inherent relationships between states. Research studies have applied this approach in large-scale clinical studies such as in schizophrenia (Jaeger et al., 2006a, Jaeger et al., 2006b). These studies suggest that the classification of cases into cognitive deficit profiles is more likely than previous methods to reveal valid neurocognitive treatment targets for novel interventions.

Models based on partially ordered sets (posets) can be used to address the complexities that arise in NP assessment data analysis, and can more fully exploit the information provided by NP batteries (see Tatsuoka, 2002, Tatsuoka & Ferguson, 2003); (Jaeger et al., 2006a, Jaeger et al., 2006b); (Lerner et al., 2009). These models generate detailed profiles of cognitive functioning, and establish a direct link to NP response patterns and specific cognitive functionalities. This leads to more precise cognitive subgroups, making the associations between cognition and outcomes such as AD conversion clearer.

2. Subjects and methods:

2.1 Subjects and study design

Of the total 769 MCI patients who were recruited for Alzheimer’s Disease Cooperative Study (ADCS) (http://www.adcs.org). We have obtained data for 513 MCI with who followed up for 24 months. All the obtained data were de-identified
without any private health information. Because of this, Institutional Review Board (IRB) approval for this study is not required.

The ADCS study was conducted between March 1999 and January 2004. Participants were between of 55 to 90 years of age and were enrolled from 69 ADCS sites in United States and Canada (Grundman et al., 2004). Patients with MCI were evaluated in this study. This study has conducted a 36-month randomized drug trail consisting of three parallel drug arms: donepezil, vitamin E, and placebo. To be eligible for study treatment, all participants had to meet the operational criteria for MCI, including (1) memory complaint, corroborated by an close relatives; (2) abnormal memory function, documented by delayed recall of one paragraph from the Logical Memory II subtest of the Wechsler Memory Scale-Revised (cutoff scores: ≤ 8 for ≥16 years of education; ≤ 4 for 8 to 15 years of education; and ≤ 2 for 0 to 7 years of education; (3) normal general cognitive function, as determined by a clinician’s judgment based on a structured interview with the patient and an informant (Clinical Dementia Rating [CDR]) and a Mini-Mental State Examination (MMSE) score greater than or equal to 24 (Folstein et al., 1975); (4) no or minimal impairment in activities of daily living (ADLs), as determined by a clinical interview with the patient and informant; and (5) not sufficiently impaired, cognitively and functionally, to meet National Institute of Neurological and Communicative Disorders and Stroke/Alzheimer’s Disease and Related Disorders Association (NINCDS/ADRDA) criteria for AD (McKhann et al., 1984), as judged by an experienced AD research clinician. In addition, all patients
had to be free of cerebral vascular diseases. All patients were assessed using the MMSE, the ADAS-cog; the global CDR and CDR sum of boxes, the ADCS/MCI-ADL Scale, a modification of the ADCS-ADL Scale, Boston Naming Test, New York University paragraph recall test, a neuropsychological battery and other examination details see Grundman et al., (Grundman et al., 2004). Participants were excluded who had a history of significant cerebral vascular disease, depression, CNS infarct, infection or focal lesions of clinical significance on CT or MRI scan and medical diseases, or psychiatric disorders that could interfere with study participation. Details of the study design were described in (Petersen et al., 2005, Fleisher et al., 2007).

2.2 Clinical Neuropsychological measurement and APOE data

513 MCI participants assessed at baseline and followed up for 24 months were used in current study. Clinical neuropsychological evaluations were performed at 6, 12, 18 and up to 24 months. A variety of neurocognitive measures were tested and the data were collected. The following neuropsychological measurements were included: Folstein Mini-Mental State Examination (MMSE) (Folstein et al., 1975), Alzheimer's Disease Assessment Scale (ADAS) – Cognitive subscale (ADAS-cog) (Rosen et al., 1984, Mohs, 1996); Delayed Recall of the ADAS 10-Word List Recall (Mohs, 1996), the New York University (NYU) Paragraph Recall Test (Immediate and Delayed) (Kluger et al., 1999), the Symbol Digit Modalities Test, Category Fluency Test (Monsch et al., 1992), a number cancellation test (Mohs et al., 1997), the Boston Naming Test (10-picture version), the Digits Backwards Test, clock
drawing, and a maze tracing task (Mohs et al., 1997). In addition to neurocognitive evaluations, all participants received baseline and follow-up assessments of overall dementia severity and functional status with the Clinical Dementia Rating (CDR) (Morris, 1993), ADCS MCI-Activities of Daily Living (ADL) Scale (Galasko et al., 1997), and the Global Deterioration Scale (GDS) (Reisberg et al., 1982). For detailed description of the NP measurements, see Appendix.

2.3 Poset Models and statistical Methods

2.3.1 Partially ordered sets as models of NP functioning

Partially ordered set models (posets) are useful models that recognize the polyfactorial nature of widely used NP measures. Poset models consist of a discrete collection of states, each of which is associated with a profile that describes functionality levels for the range of cognitive functions being tested. A partial ordering can be determined through comparison of the respective functionality levels associated to states, with a first state considered “greater” than a second state if associated functionality levels for the first state are at least as high across all functions. Posets allow for the possibility that one subject does not necessarily dominate another in terms of functionality levels. Hence, these models can describe a complex range of response patterns. From these profiles, subgroups can be identified based on relative functionality levels with specific cognitive functions. Poset have several advantages over conventional statistical methods for handling large numbers of poly-factorial neuropsychological test variables as they can, in effect, mimic the expert judgment of a clinical neuropsychologist for
each case in a large sample. They are efficient since valid conclusions can be
drawn based on relatively few measures and classification of large samples can be
accomplished rapidly. A limitation of the poset model applied here is that it is
restricted to a selected set of attributes, while additional important distinctions
remain to be tested. A simplified sample of poset model is shown in Figure 1.

Figure 1: Sample Poset model for cognitive functioning

Both A & B

A only

B only

Neither A or B

Figure 2. Largest Posterior Probability of State Membership Values. Plot indicates the largest
posterior probability value after classification to a state. The plot above indicates the frequency
as to which states had a subject classified to it.
Details of posets on methodological implementation and model validation are given in the Appendix or see Tatsuoka et al, 2002 and Tatsuoka and Ferguson, 2003.

As for a brief methodological overview, the objective in testing is to classify each subject to a cognitive profile, or state, within a poset model based on observed responses. Each measure is associated with the functions that are required in order to perform well, and this provides a link between measures and functions. Bayesian classification is adopted, with prior probabilities assigned to each state reflecting prior belief about which profile a subject possesses. Response distributions are estimated for each NP measure, depending on whether or not a subject has all the required functionality to perform well. These respective response distributions represent the relative tendencies of subjects in these groupings to give certain responses. It is expected that those having all associated functionalities should have relatively higher probabilities of performing well, while those not having all associated functionalities should have relatively higher probabilities of performing poorly.

Ideally, these distributions for a measure should be highly discriminatory with Figure 3. Sum of the largest and second largest Posterior probability. Sum of first and second largest posterior probability values in classification and shown that the mass is between two states.
respect to one another, so that responses are informative in terms of identifying a subject’s profile. This happens when the likelihood of response for much of the possible observed values is much more likely for one of the groups versus the other. Once a subject’s responses are observed, Bayes rule is used to formally weigh the relative empirical evidence. Statistical learning is embodied in updated posterior probabilities of state membership. Ideally, one state in the model will have probability close to 1, while other states will have probabilities near 0, indicating that with near certainty, a subject’s cognitive profile is known. As long as models are correctly specified, this situation will indeed arise with a sufficient amount of testing (Tatsuoka and Ferguson, 2003).

2.3.2 Linking NP functions to NP measures and the generation of partially ordered models

Figure 4. Poset representation of ADCS NP functioning profiles.
The ADCS NP measures included in the analysis, and the corresponding required cognitive functions are listed in Table 2. Identified functions are episodic memory, word fluency, cognitive flexibility, perceptual motor speed, and attention. Selection of this subset from the NP battery in the ADCS was based on the type of functions being tested by the measures, and statistical criteria such as discriminatory properties. These were identified through expert neuropsychologist opinion, and provide the linkage between function(s) and measures. Given this reliance on expert opinion, data analytic validation is important. See the Appendix for details.

Given the specifications of cognitive requirements, a poset model and cognitive functionality profiles corresponding to each of the states can be generated. This poset is illustrated in Figure 4 as a Hasse diagram. The 35 states represent cognitive states that are statistically distinguishable from the battery of 10 NP measures, in the sense that for every state, there exists at least one NP measure that has differential response distributions from each of the other states. These states and associated cognitive profiles can be identified algorithmically (see Tatsuoka, 1996).
Table 1: Demographic of MCI subjects with baseline functional and neuropsychological measures

<table>
<thead>
<tr>
<th></th>
<th>Progressors</th>
<th>Noprogressors</th>
<th>P Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Participants</td>
<td>212</td>
<td>327</td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>74.87 ± 6.64</td>
<td>71.52 ± 7.41</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>Sex (% female)</td>
<td>46.7</td>
<td>40.67</td>
<td>0.182</td>
</tr>
<tr>
<td>Education (years)</td>
<td>14.47 ± 3.10</td>
<td>14.97 ± 2.84</td>
<td>0.04*</td>
</tr>
<tr>
<td>APOE e4 (% positive)</td>
<td>161 (75.94%)</td>
<td>143 (43.73%)</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>CDR-SB</td>
<td>2.20 ± 0.79</td>
<td>1.60 ± 0.71</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>GDS</td>
<td>2.87 ± 0.61</td>
<td>2.62 ± 0.56</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>ADL</td>
<td>44.04 ± 5.47</td>
<td>47.16 ± 3.82</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>ADAS-cog total score</td>
<td>13.97 ± 4.07</td>
<td>9.45 ± 3.65</td>
<td>&lt;0.001*</td>
</tr>
</tbody>
</table>

* = Significant; CDR-SB = Clinical Diagnostic Rating Scale Sum of Boxes; GDS = Global Deterioration Scale; ADL = Alzheimer’s Disease Cooperative Study Activities of Daily Living Scale; ADAS-cog = Alzheimer’s Disease Assessment Scale-Cognitive subscale. Modified from Fleisher et al., 2007.
Table 2: Selected ADCS NP measures and their relation to cognitive functions

<table>
<thead>
<tr>
<th>x</th>
<th>Medium Episodic Memory</th>
<th>High Episodic Memory</th>
<th>Working Memory</th>
<th>Word Fluency</th>
<th>Selective Attn</th>
<th>Cognitive Flex</th>
<th>Perceptual Motor Speed</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADAS-delayed recall subscale</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADAS-word recognition subscale</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NYU paragraph</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logical Memory Delayed Recall</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boston Naming Test</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category Fluency (average of Vegetable and Animal)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number Cancellation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maze Task Time</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digit Backward</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WAIS-R Digit Symbol Substitution</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Functionality levels were classified for episodic memory, working memory, word fluency, attention, cognitive flexibility, and perceptual motor speed. MCI subjects were grouped according to cognitive profiles at baseline, and respective observed conversion rates to AD within two years were calculated. Additional genetic analyses incorporating APOE e4 allele presence/absence were performed.

Figure 4 shows how states are associated with cognitive profiles. Functionality is generally denoted either as relatively high, or relatively low, in relation to the population of interest. “Lower” functionality with respect to a function refers to relatively lower functionality with respect to these MCI subjects. State 35, the
bottom state, represents lower functionality with all the functions. In terms of the partial ordering, note for instance, from Figure 2 and Table 3 below, that States 2 and 9 are incomparable in the sense that one of the associated profiles does not dominate the other. For State 2, functionality is relatively low only for working memory, while for State 4, it is low for perceptual motor speed and working memory.

Table 3: The poset consists of a collection of cognitive states that are associated with profiles of functionality **

<table>
<thead>
<tr>
<th></th>
<th>Function</th>
<th>Medium Episodic Memory</th>
<th>High Episodic Memory</th>
<th>Word Fluency</th>
<th>Selective Attn</th>
<th>Cognitive Flex</th>
<th>Perceptual Motor Speed</th>
<th>Working Memory</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>21</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>26</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>29</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>31</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>35</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

✓ = has functionality; * = Undetermined; ** = Relative to MCI and AD only

For episodic memory, we identified three levels of proficiency, and denote them as
level 1 through 3. This was done in order to better represent the varying levels required across the measures since we have known that the episodic memory is a critical and early symptom in the early stage of AD. Level 3 is assumed to be the highest level of functionality in episodic memory requirement for the associated measures, level 2 the second highest, etc. Also, assume that if a subject has high functionality with level 3, the subject also has high functionality with levels 1 and 2. A frame of reference for interpreting how the levels of episodic memory differ can be inferred by comparing the cognitive requirements of measures respectively associated with each of the levels.

For the present model, functionality levels can for the most part always be statistically distinguished from the NP test battery under consideration. There are exceptions that, in states 14, 29, 31, high episodic memory and cognitive flexibility, such as for when State is the true state of a subject, functionality level cannot be determined. Cognitive flexibility is an executive function, which is generally tested in conjunction with other functions, such as seen in Table 2. Identifying functionality levels for such functions is more problematic, as they can be masked by poor functionality with respect to other functions.

As an example of how high episodic memory is confounded for state 31, note that given its functionality profile, a subject in State 31 is only expected to perform well on ADAS-Cog categorical selective attention and perceptual motor speed (see Table 3). For selective attention and perceptual motor speed, the two measures that tap into cognitive flexibility, note that expected performances are poor since a subject will have lower level functioning with both word fluency and perceptual
motor speed, which are respectively associated with those measures. Indeed, poor performances are expected regardless of the subject’s functionality with cognitive flexibility. Hence, it cannot be determined from the model assumptions if a subject has higher or lower functioning with cognitive flexibility.

Information about functioning can be summarized through calculation of probabilities that a subject is at a given functionality level after poset-based classification has been conducted. To determine the probability that a subject has a particular functionality level for a function, the posterior probabilities of each of the states associated with that functionality level are summed. Such probabilities can serve as a basis for identifying subgroups by functionality profiles, and generating cutoff values as a basis for grouping. An exception to this approach here is for cognitive flexibility, where some states are not associated with either high or low functionality, due to confounding. Derivation of probabilities of functionality levels for cognitive flexibility is thus conducted with more attention. More details are seen in the Appendix.

For illustration, consider the simple poset model in Figure 1, and suppose that the posterior probabilities of state membership are equally divided for a subject, 0.5 each, between State 2 (“A only”) and State 1 (“Both A and B”). Note then that the probability that the subject has a high level of functioning with A is actually 1.0, while for B it is 0.5. Hence, we are quite sure that the subject has a high level for A, even though there is still uncertainty about B.

For more precise determination of cut-off values for these probabilities associated
with cognitive functionality for predicting conversion, we have also applied ROC curve analysis to the samples. ROC curves given below indicate the diagnostic properties of variables associated with subjects having high functionality (probability values for being at high functioning with respect to the function). We have noticed that episodic memory level 2 (MEDEP) has the best tradeoff between sensitivity and specificity for predicting conversion to AD at 24 months from MCI.

<table>
<thead>
<tr>
<th>Test Result Variable(s)</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medium episodic memory</td>
<td>.789</td>
</tr>
<tr>
<td>High episodic memory</td>
<td>.756</td>
</tr>
<tr>
<td>Word fluency</td>
<td>.584</td>
</tr>
<tr>
<td>Selected attention</td>
<td>.727</td>
</tr>
<tr>
<td>Cognitive flexibility</td>
<td>.519</td>
</tr>
<tr>
<td>Perceptual motor speed</td>
<td>.481</td>
</tr>
<tr>
<td>Working memory</td>
<td>.496</td>
</tr>
</tbody>
</table>

The test result variable(s): MEDEP, HIEP, WF, ATTN, COGFLEX, PERCSPD, WM has at least one tie between the positive actual state group and the negative actual state group. Statistics may be biased.

Figure 5. ROC curve shown with all cognitive function.
This can be determined in part by largest area under the curve in Figure 5. Thus, we will focus on using episodic memory as a basis for prediction in our model. ROC Curve only applied to those subjects with lower functioning with medium episodic memory level 2 (the easier gradation level of episodic memory associated with ADAS-Cog Delayed recall and word recognition). Objective is to see if further diagnostic information is available with variables (attribute probabilities) associated with the other functions. A higher area under curve indicates better diagnostic properties above and beyond the information provided when the episodic memory level 2 variable is low (less than 0.30). Examining the

Figure 6: ROC curve shown without medium episodic memory level 1 and high episodic memory level 2 cognitive function.
ROC Curve missing episodic memory level 2 and 3, the other variables are seem not significant effect in conjunction with being low with medium episodic memory.

It should be noted that we had difficulty in analysis of the measurement of processing speed and cognitive flexibility, as displayed in the graph for maze time.

### Area Under the Curve

<table>
<thead>
<tr>
<th>Test Result Variable(s)</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word Fluency</td>
<td>.536</td>
</tr>
<tr>
<td>Select attention</td>
<td>.511</td>
</tr>
<tr>
<td>Cognitive Flexibility</td>
<td>.528</td>
</tr>
<tr>
<td>Perceptual motor speed</td>
<td>.418</td>
</tr>
<tr>
<td>Working memory</td>
<td>.514</td>
</tr>
</tbody>
</table>

The test result variable(s): WF, ATTN, COGFLEX, PERCSPD, WM has at least one tie between the positive actual state group and the negative actual state group. Statistics may be biased.

(see in Appendix), where the probability distributions have classified with all associated functions versus those who do not have all associated functions (see
attribute table in Table 2) are not well distinguished. This is not good for the classification, and so we could not get precise information about who has or does not have a relatively high level 3 of functioning with respect to those functions.

For better represent the varying levels required across the measures, and the importance note that if a measure requires high episodic memory level 3, then it necessarily requires a medium level of episodic memory level 2 as well. Similarly, if a subject is classified as having a high level of proficiency, then necessarily he or she will also have a proficiency at a medium level 2 as well. A frame of reference for interpreting how the level 2 of medium episodic memory differs from the high episodic memory level 3 can be inferred by comparing the cognitive requirements of measures respectively associated with each of the levels.

Results:

Analysis of ADCS data: heterogeneity of conversion outcomes by cognitive and genetic profile
MCI data from ADCS database, 513 amnestic MCI subjects at baseline were evaluated with neuropsychological tests and genetic APOE information was obtained and followed up to 24-month to evaluate the conversion to AD. Among these MCI subjects, there were 273 males (53.3%) and 240 females 46.7%, the average mean age is 73.2 years (SD=7.03) and mean education 14.72 years (SD=2.97, range 7-20). More details of demographic of ADCS data found in Table 1.
Since all the data were analyzed among MCI subjects, the neuropsychological function profiles indicate relative functionality levels among MCI subjects, and hence allow for measurement of differential cognitive functioning within MCI in order to predict conversion of AD. Normal subjects were not included, as the focus is on distinguishing changes and differences among MCI subjects. We expected that it would be shown that the significantly different conversion rates to AD after 24 months from MCI baseline were found across different cognitive functionality groupings by posets within MCI subjects.

In Table 4, we have tested whether there is a significant effect of the presence of an APOE e4 allele in conversion rates among MCI subjects. Among these subjects, with at least one APOE e4 allele, the conversion rate to AD is significant higher than those that do not have APOE e4 allele. Within the group of MCI subjects without the presence of APOE e4, 85.9% MCI patients do not progress to AD, and hence only 14.1% MCI subjects convert to AD during 24 months follow-up. In the group of MCI with the presence of APOE e4 allele, the conversion rate to AD is significant higher (37.9%) than those MCI patients without APOE e4 allele (14.1%), statistic significant P<0.001. But there are still 62.1% MCI patients with the presence of APOE e4 who did not progress to AD. We also found that, in general MCI subjects, without presence of APOE e4, 85.9% MCI patients do not convert to AD, which is higher than the MCI patients with APOE e4 allele (62.1%). With APOE e4 alone, relative risk (RR) for conversion to AD has been calculated is 2.68 (95% CI 1.9 – 3.9), p value < 0.001. These results suggest that the presence
of an APOE e4 allele is in fact associated with the conversion rate to AD after follow up 24 months.

Table 4: Conversion of MCI to AD with/without APOE e4 allele in 24 months.

**APOE e4 by diagnosis of Alzheimer’s disease**

<table>
<thead>
<tr>
<th>Apoe e4</th>
<th>Convert to AD</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>No</td>
<td>207</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>85.9%</td>
<td>14.1%</td>
</tr>
<tr>
<td>Yes</td>
<td>169</td>
<td>103</td>
</tr>
<tr>
<td></td>
<td>62.1%</td>
<td>37.9%</td>
</tr>
<tr>
<td>Total</td>
<td>376</td>
<td>137</td>
</tr>
<tr>
<td></td>
<td>73.3%</td>
<td>26.7%</td>
</tr>
</tbody>
</table>

Fisher’s Exact Test for “No APOE e4” and “at least one allele APOE” table: two-sided p-value = 0.001.

Table 5: The effect of episodic memory level 2 on conversion of AD

**Episodic memory level 2 by diagnosis Alzheimer disease**

<table>
<thead>
<tr>
<th>Episodic memory level 2</th>
<th>Convert to AD</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>No</td>
<td>269</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>84.9%</td>
<td>15.1%</td>
</tr>
<tr>
<td>Yes</td>
<td>58</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td>43.9%</td>
<td>56.1%</td>
</tr>
<tr>
<td>Total</td>
<td>327</td>
<td>122</td>
</tr>
<tr>
<td></td>
<td>72.8%</td>
<td>27.2%</td>
</tr>
</tbody>
</table>

Fisher’s Exact Test for “No episodic memory level 1” and “have episodic memory level 1” table: two-sided p-value = 0.001.

As we have known, episodic memory is a characteristic feature in early cognitive
impairment. As described above, episodic memory levels by posets were analyzed to define whether there are affects on conversion of MCI to AD. Among MCI patients classified with episodic memory level 2, 74 out of 132 MCI patient (56.1%) convert to AD significant higher compared to MCI patients with no episodic memory level 2 (15.1%). We have also found that relative risk of episodic memory level 2 for conversion to AD is 3.7 (95% CI 2.7 ± 5.0), p < 0.001. These findings suggest that the low functionality of episodic memory level 2 is associated with the conversion rates of MCI to AD.

Further more, in Table 6, we found that episodic memory level 2 functioning is associated with conversion rate to AD among MCI patients with or without the presence of APOE e4. MCI subjects with at least one APOE e4 allele for who have relative low episodic memory level 2 functioning, 60 out of 92 (65.2% ± 6.3%, 95% CI) converted to AD after 24 months, which is significant higher than the overall MCI to AD at the rate of 95 out of 243 (39.1% ± 6.8%, 95% CI) in this sample. Those with relatively lower functionality episodic memory level 2 with the APOE e4 allele is present (65.2%) differ significantly in conversion rates compared with those have lower episodic memory functionality level 2 with absence of APOE e4 (35.0%). Among those with episodic memory level 2 with or without APOE e4 the RR is 2.8 (95% CI 2 ±3.9) or 4.5 (95% CI 2.1 ± 9.2). Mantel Haenzel adjusted RR for episodic memory level 2, adjusting for APOE e 4) is 3.08 (95% CI 2.4 ± 3.4).

These findings suggest the low episode memory functioning level 2 is as strong (or a slightly stronger ) predictor of conversion to AD as is presence of APOE e4. More
interestingly, APOE e4 is an effect modifier of the episodic memory level 2 to AD relationship. The presence of APOE e4 diminishes the RR for conversion associated with episodic memory (RR goes from 4.5 to 2.8). Respective two-sided p-value of Fisher’s Exact Tests for “episodic memory level 1” without APOE e4 allele and with at least one allele, the p value is <0.001.

Table 6. Episodic memory level 2 by diagnosis Alzheimer’s disease stratified by APOE e4 status

<table>
<thead>
<tr>
<th></th>
<th>Convert to AD</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>No</td>
<td>Yes</td>
<td>Total</td>
<td></td>
</tr>
<tr>
<td>No APOE 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Episodic memory level 2</td>
<td></td>
<td>No</td>
<td>153</td>
<td>13</td>
<td>166</td>
</tr>
<tr>
<td></td>
<td></td>
<td>%</td>
<td>92.2%</td>
<td>7.8%</td>
<td>100.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Yes</td>
<td>26</td>
<td>14</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td></td>
<td>%</td>
<td>65.0%</td>
<td>35.0%</td>
<td>100.0%</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>Count</td>
<td>179</td>
<td>27</td>
<td>206</td>
</tr>
<tr>
<td></td>
<td></td>
<td>%</td>
<td>86.9%</td>
<td>13.1%</td>
<td>100.0%</td>
</tr>
<tr>
<td>At least one APOE 4 allele</td>
<td></td>
<td>No</td>
<td>116</td>
<td>35</td>
<td>151</td>
</tr>
<tr>
<td>Episodic memory level 2</td>
<td></td>
<td>%</td>
<td>76.8%</td>
<td>23.2%</td>
<td>100.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Yes</td>
<td>32</td>
<td>60</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td></td>
<td>%</td>
<td>34.8%</td>
<td>65.2%</td>
<td>100.0%</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>Count</td>
<td>148</td>
<td>95</td>
<td>243</td>
</tr>
<tr>
<td></td>
<td></td>
<td>%</td>
<td>60.9%</td>
<td>39.1%</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

Fisher’s Exact Test for “episodic memory level 2” and “no presence of APOE allele” table: two-sided p-value < 0.001. And Fisher’s Exact test for “episodic memory level 2” and “at least one allele APOE” table: two-sided p-value <0.001.

Amnestic multi-domain MCI, identified with poset models

An advantage of the posets approach is the ability to provide classification to
profiles of cognitive functioning across a range of functions. Notions of subgrouping, such as amnestic multi-domain MCI, can be characterized more precisely by identifying specific functions that are relatively impaired along with episodic memory.

In the ADCS data, an interesting finding has been found when grouping subjects with high functionality for both of episodic memory level 3 with the presence/absence APOE e4 allele. In Table 7, patients have no both of episodic memory level 3 functioning and APOE e4 allele are shown 23 out of 136 MCI convert to AD (conversion rate is 16.9% ± 7.2%, 95% CI), which is higher than those having relative high episodic memory level 3 functioning but no APOE e4 allele (4 out of 70) MCI subjects convert to AD (5.7% ± 7.6%, 95% CI). The conversion rate is significantly higher in the subjects without episodic memory level 3 functioning and no APOE e4 allele (Fisher's Exact test, two-sided p-value = 0.028). Even in the patients with the presence of at least one APOE e4 allele, subjects has episodic memory level 3 functioning (5.3% ± 6.7%, 95% CI), the conversion rate (17.5% ± 5.3%) is significant lower than the subjects have no high episodic memory level 3 functioning and not carried APOE e4 allele (45.7% ± 6.5%, 95% CI). These findings suggest that the high functionality of episodic memory level 3 is associated in decreasing the progression rates of MCI to AD. Our results suggest that the conversion rate is related to high functionality of episodic memory level 3 in presence/absence of APOE e4 allele.
Table 7: Episodic memory level 3 functioning by AD diagnosis by stratified by APOE e4 status

<table>
<thead>
<tr>
<th>APOE4</th>
<th>Episodic memory level</th>
<th>Convert to AD</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No</td>
<td>No</td>
<td>Count</td>
<td>Yes</td>
<td>Total</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>APOE4</td>
<td>No</td>
<td>113</td>
<td>23</td>
<td>136</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>%</td>
<td>83.10%</td>
<td>16.9%</td>
<td>100.0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>Count</td>
<td>66</td>
<td>4</td>
<td>70</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>%</td>
<td></td>
<td>94.30%</td>
<td>5.7%</td>
<td>100.0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>Count</td>
<td>179</td>
<td>27</td>
<td>206</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>%</td>
<td></td>
<td>86.9%</td>
<td>13.1%</td>
<td>100.0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>At least one allele</td>
<td>Episodic memory level</td>
<td>No</td>
<td>101</td>
<td>85</td>
<td>186</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>APOE4</td>
<td>%</td>
<td>54.3%</td>
<td>45.7%</td>
<td>100.0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Yes</td>
<td>47</td>
<td>10</td>
<td>57</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>%</td>
<td></td>
<td>82.5%</td>
<td>17.5%</td>
<td>100.0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>Count</td>
<td>148</td>
<td>95</td>
<td>243</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>%</td>
<td></td>
<td>60.9%</td>
<td>39.1%</td>
<td>100.0%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fisher’s Exact Test two-sided p-value = 0.028 for No APOE e4; two-sided p-value = 0.001 for with subjects that have at least one APOE4 allele.

In addition, to compare the scores changes of ADAS-cog, CDR sum of boxes, ADL (activity of daily living) and Beck depression inventory from MCI baseline to 24 months base on whether or not medium episodic memory is high or not. If patients have low medium episodic memory functionality, the mean difference of ADAS-cog changes from 5.93 (SD = 8.5) to -1.52 (SD = 7.3) (significant two tailed t-test p value = 0.001) after 24 months. The difference of the CDR sum of boxes is decreased from 1.79 (SD =1.99 ) to 0.29 (SD = 1.64) (significant two tailed t-test p value = 0.001). Also, the total scores of mean of ADL is reduced from -6.75 (SD = 11.08) to –2.5 (SD = 11.48), significant two tailed t-test p value is 0.001. However, Beck depression score is decreased from 1.15 (SD = 7.47) to -0.18 (SD = 7.9), not
statistically significant two tailed t-test $P=0.091$. Details are seen in table 8. These findings provide further validation of progression risk differs depending on cognitive profile at baseline.

Table 8: Decreased Neuropsychological measures of MCI subjects at baseline after 24 months.

<table>
<thead>
<tr>
<th>episodic memory level 2</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADAS difference</td>
<td>NO</td>
<td>-1.5244</td>
<td>7.38786</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>5.9328</td>
<td>8.54814</td>
</tr>
<tr>
<td>CDR sum of Boxes</td>
<td>No</td>
<td>0.2881</td>
<td>1.63862</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>1.7873</td>
<td>1.99188</td>
</tr>
<tr>
<td>Activities of Daily Living difference</td>
<td>No</td>
<td>-2.4970</td>
<td>11.48167</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>-6.7463</td>
<td>11.07777</td>
</tr>
<tr>
<td>Becks depression</td>
<td>No</td>
<td>-0.1799</td>
<td>7.96828</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>1.1493</td>
<td>7.47377</td>
</tr>
</tbody>
</table>

**Discussion**

The major concerns have been increasing in how to precisely to detect and to predict the progression of MCI to AD. Because of heterogeneity among MCI subjects, more precise detection methods based on neuropsychological measures are needed. In order to develop a more accurate method to predict MCI progression, we used partially ordered set models to directly address the complexities inherent in neuropsychological assessment data analysis, to more
fully utilize the information provided by neuropsychological batteries and cognitive functionality.

In the current study, our results indicate that episodic memory at a level requiring a relatively high demand, among the functions considered in this study, is the strongest single cognitive function linked to conversion from MCI to AD. These findings confirm findings from previous studies (Tierney et al., 2005), (Tabert et al., 2006), (Blacker et al., 2007), and (Landau et al., 2010). Second, our findings clearly show the importance of the APOE e4 allele in affecting risk for AD, which supports the similar findings of others (Aggarwal et al., 2005, Landau et al., 2010). The combination of both functionality episodic memory and genetic APOE information has the strongest association with conversion rate to AD.

For amnestic multi-domain MCI, one advantage of the proposed posets methods is that more specific parsing of neuropsychological functioning is possible. It appears from our analyses that low episodic memory functioning, in conjunction with APOE e4, has a stronger link to AD progression risk than other cognitive functions considered in conjunction with episodic memory. Poset models thus appear helpful in precisely stratifying the episodic memory risk cognitively.

These findings may be helpful in guiding the design, of more focused and efficient neuropsychological batteries that are tailored for early-stage studies. For instance, we conducted a classification analysis based on the subset of the measures that exclusively tap into episodic memory and attention while also considering APOE genotype. Similar findings in terms of conversion rates by
functional groupings arose. For instance, with the NP battery, and using similar cutoff values, for those with relative lower functioning of episodic memory level 2 and an APOE e4 allele, the conversion rates is higher within two years. However, among those with high episodic memory level 3 functioning, even in the presence of APOE e4 allele, AD conversion rates within two years remain lower compared to those with comparing to the lower episodic memory level 2. In addition, from MCI baseline to 24 months, we found a significant decreased total score for ADAS-cog, CDR sum of boxes and ADL, suggesting that use of posets may provide a more focused battery that is useful for early stage AD studies.

If these findings can be validated further, then in combination with other laboratory markers, they may be useful as a basis for screening in broad settings where costs, invasiveness concerns, and technological access constraints could hinder use of imaging methods and/or procurement of CSF (De Meyer et al., 2010). Based on an individual’s NP test results, risk assessments could be made based on classification results using poset models. Note that from Table 5, the observed specificity of the low episodic memory level 2 criterion among MCI, even if used alone in diagnosis regardless of APOE e4 status, the conversion rate in converter group are in the “Not low” group. It may be that for non-AD subjects who currently fall under the MCI categorization, lack of AD risk can be established quickly with fair accuracy.

According to these data, we believe it is important to recognize and treat early in subjects with MCI because this co-morbidity can be associated with worse cognitive performance, mild extrapyramidal signs, and functional disability. It
reveals that neuropsychologic measures are highly prevalent in MCI. Due to the highly heterogeneous data on neuropsychological measures in MCI to date, suggesting standardized MCI criteria and behaviors instruments are required to evaluate the prognostic role of NPS in MCI. These findings may be helpful in guiding the design of future, more focused and efficient NP batteries that are tailored for early-stage studies.

However, in ADCS data set, we could not duplicate the findings that we found in previous ADNI study here (Tatsuoka et al., 2011, unpublished data). One reason probably due to how/what the neuropsychological measurements were used in these two studies, for example, ADNI study has used Trails A and B, instead of using Maze test in ADCS study. As seen in our study, the density estimates for the Maze test is not well distinguished, and the test has poor statistical discrimination in our classification system. We had also difficulty with NYU Delayed recall, instead, we used Logical Memory delayed recall and it had better discrimination in ADCS data. Also, an ordered correspondence between Boston naming and categorical fluency was not clear. Note that the functions involved with Boston naming (word fluency and attention) are also required for Categorical fluency. We thus expected that poor performance on Boston naming would indicate that one would do poorly with categorical fluency, and that good performance on categorical fluency would lead to good performance on Boston naming. But that is not the case, unlike in ADNI.
Finally, advantage of poset models classified cognitive function profiles into states can be used in the areas of AD-related research and other cognitive-impaired neurodegenerative diseases. The poset models can also be applied in continuing to validate these findings in other longitudinal data sets, combining these results in prediction with other variables such as imaging and CSF biomarkers (De Meyer et al., 2010, Landau et al., 2010), clinical trial design, and modeling the longitudinal course of cognitive change over a longer time course. This information may be useful as a basis for screening in broad settings where costs, invasiveness concerns, and technological access constraints could hinder use of imaging methods and/or procurement of CSF (De Meyer et al., 2010).
Appendix:

**Description of Cognitive Measures:**

**ADAS-cog:** Alzheimer’s disease assessment scale-cognitive subscale (11 and 13 sub-items; Mild Cognitive Impairment); the original ADAS-cog is a psychometric scale consisting of 11 items that evaluate selected aspects of memory, orientation, attention, language, reasoning and carrying out instructions, it ranges from 0 to 70 and for all its variants higher scores indicate poorer function.

**BNT:** Boston Naming Test; that assesses the ability to name pictures of objects through spontaneous responses and need for various types of cueing.

**Category Fluency test:** quickly producing words in a specified category (e.g. animals, fruits, vegetables) is the major ability required by this test.

**Delayed word list recall:** a test of memory

**Digit-Backward test:** a test of executive function

**Digit Cancellation Task:** this task is essentially a measure of short-term memory and reaction time

**Maze test:** a nonverbal test of performance intelligence consisting in a graded set of paper forms on which the subject traces the way from a starting point to an exit.

**NYU PR:** New York University Paragraph Recall; a test of memory

**Symbol Digit Modalities Task:** a test-to-test executive function

**Verbal Fluency category:** a test designed to measure the speed and flexibility of verbal thought processes
**WMS-R:** Wechsler Memory Scale-Revised; a test that yields information about various kinds of memory and learning processes and provides a comprehensive assessment of memory

**Global and Clinic Measures**

**CDR:** Clinical Dementia Rating (scale); The Clinical Dementia Rating Scale (CDR), a comprehensive structured interviews based on worksheets. CDR assesses dementia severity by staging, and includes cognitive, functional and social domains in the overall staging

**CDR-SB:** Clinical Dementia Rating–Sum of the Boxes; Its variant, the CDR sum of the boxes (CDR-SB), is a more quantitative numerical rating (it merely sums up the scores of the six individual domains of the CDR)

**CGIC-MCI:** a clinical global impression of change scale for use with patients with aMCI based on semi-structured interviews from both the subject and the informant

**GDS:** Global Deterioration Scale; the Global Deterioration Scale (GDS) which rates seven stages of dementia, with higher scores rating poorer cognition, assessing the phenomenological global progression of AD cognitively, functionally and behaviourally without structured interview

**MMSE:** Mini Mental State Examination; The MMSE is a brief, quantitative measure of cognitive status in adults

**ADCS-ADL:** Alzheimer’s disease Cooperative Study Activities Living

**ADCS-ADL-MCI:** Alzheimer’s disease Cooperative Study Activities Living adapted to MCI
**Beck Depression Inventory**: a test to assess the depression

**HAM-D**: Hamilton Rating Scale for Depression

### Statistical framework and data analysis

The statistical framework follows as in Tatsuoka (2002), where further details are given. Also see Jaeger et al. (Jaeger et al., 2006a, Jaeger et al., 2006b). Briefly, a Bayesian approach to classifying subjects to a state will be adopted, so that prior probabilities of state membership are assigned for each test subject. For the present study, a uniform, non-informative prior probability was assigned, with each state being viewed as equally likely to be the true one prior to updating the probabilities of state membership through observed test scores. While not done in the present study, an advantage of the Bayesian approach is that pertinent background information or expert opinion can be incorporated into a subject’s state membership prior probability specifications, if it is believed a priori that certain states are more likely to be the true cognitive state than others.

### Estimation of the distributions of responses for the measures by profile

For each measure, it will be assumed that there are two response probability distributions for observing a score on the measure. These distributions respectively describe probabilities of performing in certain ranges of scores, and depend on the cognitive functionality of a subject. One distribution will be associated with subjects in states that possess the functionality levels with respect to the functions associated with the measure. The other distribution will represent the response
distribution for subjects in states that do not have all the associated functionality levels. An advantage to only specifying two response distributions per measure is that these models are parsimonious, an important consideration in latent class settings with moderate sample size. Classification is still quite efficient in this setting, in terms of the exponential rate of convergence of the posterior probability of the “true” state to one (Tatsuoka and Ferguson, 2003). Conditional independence of response distributions between measures is assumed.

Based on baseline ADCS data for MCI and early AD, response distributions were estimated for each NP measure in the analysis. In addition to increasing sample size, combining data from both diagnostic groups allows for the modeling of a range of scores, which in turn allows for the capturing of change in MCI subjects over a two-year period. The Bayesian statistical estimation procedures for the multinomial distributions follow as in Tatsuoka (2002) and Jaeger et al. (Jaeger et al., 2006a, Jaeger et al., 2006b), with Dirichlet conjugate priors used with Markov Chain Monte Carlo with stationary convergence attainment assessed as in Geweke (1992). In the estimation of response distributions, prior specifications were uniformly distributed and the same for all the category responses. A fairly large prior variance for these probability parameters also was assigned, with Dirichlet parameters all set to the value 2. Non-informative priors were selected in the estimation of the normal mixture distributions as well as detailed by Ishwaran and James (2002). Sensitivity analysis for priors was conducted as well, in that some modifications to prior specifications were made, to see if classification results differ. Since MCI and early AD subjects in ADCS combine to a baseline sample
size of 513, estimation results were found not to be sensitive to prior specifications.

Finally, while the normal mixture models can indeed be more precise, many parameters are involved in estimating mixtures, which can lead to variability. Other measures were not modeled with normal mixtures, for instance because classification performance was not improved, in terms of decisiveness or variability.

**Data-Analytic Validation**

Data-analytic validation of model fit and of the cognitive specifications associating functions to measures is essential to providing reliable and accurate results, as the cognitive processes underlying the assessment responses are latent and complex. In general, the data-analytic tools for assessing model fit and validating cognitive specifications involve analyzing NP test distribution estimates and patterns in the classification results (see Tatsuoka, 2002).

Note that the estimates of probabilities for an observation belonging to certain intervals in most instances reflect the order structure of the cognitive model. In other words, subjects in a state associated with functionality levels necessary for performance on an NP test have higher estimated probabilities of performing well than subjects in states without all the required functionality levels associated with the NP test. Conversely, then, a subject without all the functionality levels associated for a given NP test has a relatively higher estimated probability of performing poorly. These estimation results indicate that the response behavior of subjects is consistent with model specifications.
Classification results also can indicate how well a model fits. Ideally, classification for each subject results in posterior probability mass for state membership concentrating on one state; in other words, the probability after observing responses from a subject for state membership is near 1 for one state, and 0 for all others. This indicates that response behavior is consistent across measures towards one state, as would be expected if the model were correctly specified. In this analysis, for the most part, posterior probability mass settled on one or two states, as reflected in Figure 3, indicating good model fit to the response data. However, this did not always occur for all subjects, perhaps due to several possible reasons. These include possible issues with reliability of the NP tests, their construct validity in relation to the specified functions, or the limitations of model fit in terms of how the specified functions adequately describe performance. This may hold true for instance for the Maze test, which involves several functions. Still, we think that a main reason is the lack of replication of measurement in the battery. While profiles can be statistically distinguished through one measure, for many profiles, only one measure is not sufficient to decisively classify between two profiles, particularly when mid-range scores are observed.

**Determining probabilities of functionality under confounding**

Confounding of functionality profiles, and of functionality levels of specific functions in particular, occurs when they are not distinguishable statistically. This occurs when more than one profile of functionality levels is expected to give the same responses across all measures from a given battery. In other words, associated response distributions for subjects with those profiles are all the same. In such
cases, profiles confounded with each other can be viewed as an equivalence class, and the corresponding state in the poset that represents this equivalence class will have undetermined information for functionality levels that are in contradiction among profiles in the class. More discussion is on this phenomenon is given in Tatsuoka (2002). Determining the probability of functionality levels for cognitive flexibility, which are confounded for certain states, was done as follows. The proportion of corresponding profiles in each of the equivalence class associated with the functionality level of interest can be determined, and these proportions are weighted (multiplied) by the posterior probability that a subject belongs to the corresponding state. Note that for states with no confounded profiles, this proportion is either 0 or 1. These products are summed for all states, to obtain the posterior probability of a subject having the functionality level of interest. Note that for this approach, the assumption here is that profiles that are confounded are equally likely to be true. Although not done here, as it did not appear that cognitive flexibility was a statistically significant marker for conversion, the inferential impact of this assumption can be assessed by comparing results derived from differing assumptions, to assess sensitivity.

**Alzheimer’s Disease Cooperative Study (ADCS)**

The study data used in the preparation of this article were obtained from the Alzheimer’s Disease Cooperative Study (ADCS) database. The ADCS was formed in 1991 as a cooperative agreement between the National Institute on Aging (NIA) and the University of California San Diego. It included 69 sites in United State and Canada. The ADCS is a major initiative for Alzheimer’s disease clinic studies in the
Federal government, addressing treatment for both cognitive and behavioral symptoms. This is part of the NIA Neuroscience and Neuropsychology of Aging Program’s effort to facilitate the discovery, development and testing of new drugs for the treatment of AD and also is part of the Alzheimer’s Disease Prevention Initiative. ADCS is the result of efforts of many co-investigators from academic institutions, and subjects have been recruited from over 69 sites across the U.S. and Canada.
REFERENCES:


Ishwaran H and James LF. Approximate Dirichlet process computing in finite normal mixtures: smoothing and prior information. J Comp Graph 2002; 97, 1154-1166.


Lerner A, Tseng HY, Jaeger J, Tatsuoka C. Predicting conversion from MCI to AD using partially ordered models of cognitive functioning, Alzheimer’s Disease Neuroimaging Initiative Meetings 2009, Seattle.


predicting progress for amnestic mild cognitive impairment to Alzheimer disease.


