American Journal of Alzheimer's Disease & Other Dementias®
Volume 35: 1-8

DOI: 10.1177/1533317520918719
journals.sagepub.com/home/aja

Longitudinal Sensitivity of Alzheimer's Disease Severity Staging

Julia S. Benoit, PhD[1], Wenyaw Chan, PhD[2], Linda Piller, MD, MPH[3], and Rachelle Doody, MD, PhD[4]

Abstract

Understanding Alzheimer's disease (AD) dynamics is essential in diagnosis and measuring progression for clinical decision-making; however, clinical instruments are imperfect at classifying true disease stages. This research evaluates sensitivity and determinants of AD stage changes longitudinally using current classifications of "mild," "moderate," and "severe" AD, using Mini-Mental State Examination (MMSE), Alzheimer's Disease Assessment Scale–Cognitive subscale (ADAS-Cog), and the Clinical Dementia Rating–Sum of Boxes (CDR-SB) thresholds. Age and pre-progression rate were significant determinants of AD progression using MMSE alone to stage AD, and pre-progression was found to impact disease progression with CDR-SB. Sensitivity of these instruments for identifying clinical stages of AD to correctly staging a "moderate" level of disease severity for outcomes MMSE, CDR-SB, and ADAS-Cog was 92%, 78%, and 92%, respectively. This research derives longitudinal sensitivity of clinical instruments used to stage AD useful for clinical decision-making. The MMSE and ADAS-Cog provided adequate sensitivity to classify AD stages.

Keywords

Alzheimer's disease, disease progression, continuous-time Markov models, longitudinal study

Introduction

Alzheimer's disease (AD), a major public health problem, affects over 5 million elder individuals and is among the leading causes of US mortality.[1] Due to the nature and duration of the disease, it is costly to patients, caregivers, and the health-care system. Caregivers often encounter higher levels of emotional stress, and their health, employment, and financial security are negatively impacted. While the elderly patients continue to live longer, without curative treatment, prevalence is expected to double by 2050. Clinically, understanding the timing of AD progression and prediction of AD stage duration, especially the mild stage, would help tailor treatment plans that parallel the progression, thus improving disease management.[2-4]

Although methods for predicting progression rates have been studied,[2,5] disease onset and progression is difficult to quantify through studying AD natural history alone, as AD is a complex disease with complex pathogenesis comprised of social, environmental, and genetic factors. One challenge is monitoring change using clinical outcomes. Identifying and validating neuropsychological measures to predict changes in AD severity is important in linking the underlying disease process to observed clinical symptoms, monitoring response to potential therapies, and assessing potential predictability of biomarkers, all requiring sophisticated longitudinal statistical models, in which few studies have utilized. Discrete-time Markov chains have been used to understand the natural history of AD[6] along with other longitudinal techniques,[7] but research thus far does not account exhaustively for unequal duration between visits or lapses between actual stage changes. Baseline risk factors associated with long-term outcomes have been studied,[8] but just how risk factors impact the actual dynamics of the process have not been elucidated.

Clinically, AD is often staged as "mild," "moderate,"or "severe" using thresholds of well-validated and reliable

[1] Texas Institute for Measurement Evaluation and Statistics (TIMES), University of Houston, TX, USA
[2] Department of Biostatistics and Data Science, School of Public Health, University of Texas Health Science Center, Houston, TX, USA
[3] Department of Epidemiology, Human Genetics, and Environmental Sciences, School of Public Health, University of Texas Health Science Center, Houston, TX, USA
[4] PD Neuroscience, Roche Pharma, Basel, Switzerland

Corresponding Author:
Julia S. Benoit, Texas Institute for Measurement Evaluation and Statistics (TIMES), University of Houston, TX, USA.
Email: julia.benoit@times.uh.edu

neuropsychological scales such as the Mini-Mental State Examination (MMSE),[9] Alzheimer's Disease Assessment Scale-cognitive Subscale (ADAS-Cog),[10] and/or Clinical Dementia Rating–Sum of Boxes (CDR-SB),[11] among others. Imaging biomarkers indicating existing levels of severe cognitive impairments alongside a positive amyloid positron emission tomography scan[12,13] and/or other biomarkers such as cerebrospinal fluid amyloid Beta and tau,[14-16] each potentially costly and burdensome, are used for AD diagnosis, giving confidence that the diagnosis of dementia is caused by Alzheimer's pathology. Alzheimer's disease staging, however, is a separate process, of which there is no easily obtainable neuropsychological-based "gold standard" instruments; thus, examining cut-point sensitivities is challenging. Nevertheless, though subject to misclassification, these instruments remain the primary assessments used to classify AD severity. Studies have examined MMSE sensitivity and specificity of predicting AD onset or progression[17-20] using receiver operating characteristic curves and growth mixture models, among others. Although these studies assess misclassification post-analysis, they lack statistical model development to allow for misclassification of severity.

Operative use of the CDR-SB to stage Alzheimer's dementia has been established and cross-validated to classify dementia severity into 6 groupings, ranging from "normal" (CDR-SB: 0) to "severe dementia" (CDR-SB: 16.0-18.0).[21-23] This cross-validation was not used in determining change in disease over time. The ADAS-Cog, an end point used in AD clinical trials, has been widely used to assess progression by examining change from baseline to measure improvement[24] but has not been used to stage AD severity.[25] Benge and others,[26] through item response theory analysis methods, found that scores 14 and 64 of the raw score provided information on lower and higher levels of cognitive dysfunction.

Few studies have examined sensitivity of MMSE and others for staging AD without a "gold standard." Luo et al[27] studied the sensitivity of baseline MMSE, ADAS-Cog, CDR-SB, and CDR using binomial regression framework. To the best of our knowledge, sensitivity of tertiary disease stages based on cognitive or functional instruments has not been studied in a longitudinal setting.

For research pertaining to monitoring AD and patient responses in clinical studies, cognitive measures will continue to provide useful assessments of AD progression.[28] Clear staging criteria using accepted instruments will improve comparability. Sensitivity of AD severity changes should be considered due to the continuous nature of the AD process. The purpose of this study is to apply a recently developed continuous-time Markov chain (CTMC) model estimation technique to estimate the probability of misclassification of AD severity and the determinants (eg, sex, age, years of education, pre-progression rate [PPR]) of AD stage changes over time using current instruments (MMSE, ADAS-Cog, and CDR-SB). Namely, this research aims to analyze the progression of AD stage changes using 3 different instruments and measure the sensitivity of each in the absence of a "gold standard."

Methods

This research has been conducted according to the World Medical Association Declaration of Helsinki. Approval of this study has been provided by Baylor College of Medicine–Alzheimer's Disease and Memory Disorders Center (IRB# H-9095).

Study Population

A prospective cohort of patients with probable AD[29] from the Baylor Alzheimer's Disease and Memory Disorders Center collected from January 1990 to September 2011 were examined. Patients were self-referred or referred to the center and evaluated using history and physical examinations, laboratory testing, neuropsychological instruments, and imaging[30] after providing written informed consent, which also included consent by the patient(s) or a legally authorized representative for patient information and images to be published. Patient sociodemographic information such as age, sex and years of education, medical history, and estimates of symptom duration[31] was collected at baseline. Further details of the study design and outcome diagnosis have been described elsewhere.[30]

Disease Severity

Patients underwent neuropsychological testing at baseline and annually or on an as-needed basis for medication management. The MMSE (scored 0-30) focuses on memory, attention, and language and aids in identifying dementia progression and severity; lower scores indicate more severe dementia. The ADAS-Cog (scored 0-70) targets cognitive impairment in patients with AD; higher scores indicate worse cognitive impairment. The CDR-SB (scored 0-18) measures global performance; higher scores indicate higher levels of global impairment. These 3 measures of cognitive outcome and global impairment can be used to classify AD severity into mild, moderate, and severe stages. Standard cutoff points of MMSE were used to classify AD severity into mild (MMSE $\geq$ 20), moderate (10 $\leq$ MMSE $\leq$ 19), and severe (MMSE $\leq$ 9).[32] We applied staging presented by O'Bryant et al[21] to CDR-SB (mild: 0-9.0, moderate: 9.5-15.5, severe: 16.0-18.0) and by Benge et al[26] to ADAS-Cog (mild: 0-14, moderate: 15-63, severe: 64-70).

Pre-Progression Rate

Pre-progression rate, a baseline enrollment measure indicating decline from symptom onset to first clinic visit, is defined as the average decline of MMSE per year before the first physician visit and is calculated as (30 − baseline MMSE)/estimated symptom duration in years.[31] It has been proven to be predictive of cognitive performance over time[2] and used to classify patients as slow, intermediate, and rapid progressors.

Inclusion Criteria

Patients with probable AD with complete information on baseline covariates, a baseline pre-progression index, and at least 1

follow-up visit post-baseline (ie, at least one possible transition) were included in this study. Interobservation time was calculated as the duration between 2 consecutive observations.

Data Analysis

Baseline characteristics were compared using χ^2 tests for categorical variables and analysis of variance for continuous variables across severity levels for each outcome. For longitudinal modeling, true disease severity (mild, moderate, severe) of each individual is modeled as a 3-state CTMC with a transition rate matrix of unknown parameters $\{q_{ij} > 0, i \neq j; i,j = 1, 2, 3\}$, where i represents the state of the current process and j represents the state visited at the time of change. Defining the mild stage of severity as state 1, moderate as state 2, and severe as state 3 under CTMC framework, q_{12} refers to the rate of transitioning from the mild stage of severity to the moderate stage at the time of transition. We model each observed outcome MMSE, ADAS-Cog, and CDR-SB as a separate sequence of observed outcomes to estimate the transition rates of an individual currently in one stage of AD severity of transitioning into one of the other 2 stages as a function of demographic covariates to determine the disease process. Further details of this approach can be found in the study by Benoit et al.[33] The model assumes that the underlying disease stage at any particular time point is dependent upon the most recent stage and not the full history of transitions. Here, interest lies on the mean duration, $-q_{ii}^{-1} = \left(q_{ij} + q_{ij'}\right)^{-1}$, i, j, j$'$ = 1 to 3; $i \neq j, j'; j \neq j'$, spent in one of the 3 disease stages before exiting out. A log-link function is used to link the transition rates to a linear combination of the baseline covariates (age, education, gender, and PPR), that is,

$$\log q_{ij} = \beta_{0ij} + \sum_{r=1}^{p} \beta_r x_r = (\beta_{0ij}, \beta_1, \ldots, \beta_p)\underline{x}, \quad \text{for } i \neq j; i,j = 1,2,3 \text{ and } r = 1, \ldots, p. \tag{1}$$

This method also assumes that the covariates impact the dynamics of the transition, not on any specific hazard. The duration in i follows an exponential distribution with mean q_{ii}^{-1}. From equation 1, the mean can be expressed in multiplicative form of the covariates as: $\left(q_{ij} + q_{ij'}\right)^{-1} = \left(e^{\sum_{r=1}^{p} \beta_r x_r} \left(e^{\beta_{0ij} + \beta_{0ij'}}\right)\right)^{-1}$ and each $e^{-\beta_r}$ is the multiplicative impact on the mean duration of individuals of a specific value of age, sex, education level, and/or baseline PPR transitioning from one stage of severity to another, and $e^{\beta_{0ij}}$ is the intercept (weight) or the transition intensity in favor of moving from category i to j for any subject with covariates $x_1, x_2, \ldots, x_r$ equal to 0. Note that e^{β_r} is the transition intensity ratio for individuals whose $X_r = 1$ versus $X_r = 0$ and similar interpretation applies when X_r is continuous. Note also under this model and using CTMC framework, that probability that an individual in state i transitions to j can be shown to be calculated as:

$$\frac{q_{ij}}{\sum_{i=1,\, i\neq l}^{j} q_{ij} + q_{il}} = \frac{e^{\beta_{0ij}}}{e^{\beta_{0ij}} + e^{\beta_{0il}}}. \tag{2}$$

For estimating misclassification probabilities, although an individual's true severity or observed severity may transition from mild to severe over a period of time, we assumed that an individual who is truly in a mild severity state would not be misclassified as severe and that severe would not be misclassified at mild. As the statistical model was based on commonly used and studied staging assessments used to classify disease severity, it is unlikely that these scores would have such large errors. An explicit representation of the misclassification matrix can be found in the study by Benoit et al.[33] Misclassification is assumed independent of the disease process and due to the staging; however, in this model, there is an indirect adjustment of misclassification posed in the presence of explanatory variables. All statistical analyses and customized estimation procedures were written and conducted using SAS version 9.3 (SAS Institute Inc, Cary, North Carolina).

Table 1. Baseline Characteristics of the ADMDC Probable AD Cohort, Houston, Texas, 1990 to 2011 (N = 1091).

	Mean (SD)/n (%)	Range
Female	730 (67)	
Age, years	73.3 (8.6)	43-93
Education, years	13.9 (3.5)	0-29
Early exposure index, years	3.8 (2.2)	0.5-13
Follow-up time, years	3.2 (2.2)	0.8-13.9
Total number of visits	3.7 (1.9)	2-12
Pre-progression rate (PPR)[a]	3.4 (2.8)	0-22
Slow (PPR < 2)	360 (33%)	
Intermediate (PPR = 2-4)	505 (46%)	
Rapid (PPR ≥ 5)	226 (21%)	

Abbreviations: AD, Alzheimer's disease; ADMDC, Alzheimer's Disease and Memory Disorders Center; SD, standard deviation.
[a]Pre-progression rate at enrollment = (30 − baseline Mini-Mental State Examination)/estimated duration of symptoms in years.

Results

Among 1106 patients with probable AD evaluated from January 1990 to September 2011, 1091 had complete baseline characteristics. Baseline characteristics included age, gender, years of education, cumulative months of exposure to antidementia drugs, follow-up time, total number of visits, and PPR (Table 1) as well as the MMSE, CDR-SB, and ADAS-Cog baseline scores and disease severity along with a complete case description (Tables 2 and 3). At baseline, patients were on average 73 years old (43-93), had 14 years of education (0-29), and predominantly female (67%). The number of visits ranged from 2 to 12 with a median of 3. After exclusion and inclusion criteria were applied, longitudinal analysis included the following

Table 2. Baseline Cognitive and Functional Measures and Their Severity Distributions Among the ADMDC Probable AD Cohort, Houston, Texas, 1990 to 2011 (N = 1091).

	Mean (SD)	Range	Mild n (%)	Moderate, n (%)	Severe, n (%)
MMSE (n = 1091)	19.8 (6.6)	0-30	651 (59.7)	339 (31.1)	101 (9.2)
CDR-SB (n = 1023)	6.5 (4.1)	0.5-18	809 (79.1)	168 (16.4)	46 (4.5)
ADAS-Cog (n = 884)	23.6 (12.5)	0-68	232 (26.2)	647 (73.2)	5 (6)

Abbreviations: AD, Alzheimer's disease; ADAS-Cog, Alzheimer's Disease Assessment Scale–Cognitive; ADMDC, Alzheimer's Disease and Memory Disorders Center; CDR-SB, Clinical Dementia Rating Scale–Sum of Boxes; MMSE: Mini-Mental State Examination; SD, standard deviation.

Table 3. Complete Case Distribution of Baseline Cognitive and Functional Measures and Their Severity Among the ADMDC Probable AD Cohort, Houston, Texas, 1990 to 2011 (N = 871).

	Mean (SD)	Range	Mild n (%)	Moderate, n (%)	Severe, n (%)
MMSE	20.7 (5.9)	0-30	564 (64.8)	262 (30.1)	45 (5.2)
CDR-SB	5.9 (3.7)	0.5-18	728 (83.6)	123 (14.1)	20 (2.3)
ADAS-Cog	23.6 (12.5)	0-68	228 (26.2)	638 (73.3)	5 (6)

Abbreviations: AD, Alzheimer's disease; ADAS-Cog, Alzheimer's Disease Assessment Scale–Cognitive; ADMDC, Alzheimer's Disease and Memory Disorders Center; CDR-SB, Clinical Dementia Rating Scale–Sum of Boxes; MMSE: Mini-Mental State Examination; SD, standard deviation.

sample sizes for each outcome: ADAS-Cog (n = 847), MMSE (n = 972), and CDR-SB (n = 982).

Baseline MMSE, CDR-SB, and ADAS-Cog mean (standard deviation) score was 19.8 (6.6), 6.5 (4.1), and 23.6 (12.5), respectively (Table 2). Increased baseline PPR was associated with baseline severity for all outcome measures as was education level (Table 4). Baseline AD severity varied by gender for all outcome measures. Age was associated with baseline severity for CDR-SB ($P = .0256$).

Sensitivity of Cognitive and Functional Measures in Staging AD

Overall, sensitivity to correctly staging a moderate level of disease severity for outcomes MMSE, CDR-SB, and ADAS-Cog was 92%, 78%, and 92%, respectively (Table 5). Moderately staged patients when classified by ADAS-Cog and/or MMSE were ≥2 times more likely to be incorrectly staged as mild rather than severe. Incorrectly classifying moderate stage to mild or severe was estimated as about equally likely (11%) using the CDR-SB.

Disease Progression

Mini-Mental State Examination. From the results in Table 6, the intensity of transitioning from mild to moderate severity is $e^{\hat{\beta}_{012}} = 0.97$ (95% confidence interval [CI]: 0.44-2.15); the intensity of transitioning from moderate to severe stages is $e^{\hat{\beta}_{023}} = 0.71$ (95% CI: 0.32-1.57); or to mild is $e^{\hat{\beta}_{021}} = 0.05$ (95% CI: 0.02-0.14). At the time of disease stage change, a patient with moderate disease severity has probability of .93 of moving to severe status: (0.7/[0.7+0.05] = .93) (from Equation 2). On average, patients with a baseline PPR of 1 point per year higher have their mean time to stage change decreased by a multiplicative factor of 1/1.07, whereas increased age had an increased association.

Clinical Dementia Rating–Sum of Boxes. Using the same interpretation as above, when modeling disease stage changes based on CDR-SB, the intensity of transitioning from mild to moderate severity is $e^{\hat{\beta}_{012}} = 0.34$ (95% CI: 0.15-0.74), and the intensity of transitioning from moderate to severe stages is $e^{\hat{\beta}_{023}} = 0.24$ (95% CI: 0.11-0.54). At the time of disease stage change, a patient with moderate stage of disease severity has a probability of .89 of moving to severe status. On average, patients with a baseline PPR of 1 point per year higher have their mean time to stage change decreased by a multiplicative factor of 1/1.08.

Alzheimer's Disease Assessment Scale–Cognitive subscale. The base transition intensity of moving from the mild to moderate stage of severity is 0.24 (95% CI: 0.03-1.80) based on the ADAS-Cog model, and the intensity associated with movement from moderate to severe is 0.01 (95% CI: 0.00-0.07). At the time of disease stage change, a patient with moderate moves to severe with a probability of .83. On average, patients with an enrollment PPR of 1 point per year higher have their mean time to stage change decreased by a multiplicative factor of 1/1.08.

Patient PPR at enrollment had a statistically significant multiplicative impact on the rate of movement of AD stage changes classified using MMSE and CDR-SB, and age had a slight multiplicative reduction in progression when disease status was classified using MMSE.

Discussion

The purpose of this article was 2-fold: to examine the validity of neuropsychological assessments (MMSE, CDR-SB, and ADAS-Cog) used to stage AD disease severity and to characterize the dynamic characteristics of AD progression over time to aid in clinical staging of AD.

Our results indicate that MMSE best classifies moderate stage of severity, with 7% misclassified as mild. This could suggest that misclassification rates may rely on the neuropsychological measures. The MMSE has been criticized by some as being insensitive in distinguishing mild cognitive impairment (sensitivity/specificity reported <80% depending on criteria).[21] However, diagnostic screening studies typically use cross-sectional data to compute sensitivity and specificity and also report crude estimates.[27] At most, an adjusted categorical model might be used. This study takes advantage of clinical

Table 4. Distribution of Baseline Covariates Among AD Stages of Severity Defined Using Cognitive and Functional Instruments Among the ADMDC Probable AD Cohort, Houston, Texas, 1990 to 2011 (N = 1091).

	Mild Mean (SD)/n (%)	Moderate Mean (SD)/n (%)	Severe Mean (SD)/n (%)	*P* Value
MMSE				
Female	406 (62.4)	245 (72.3)	79 (78.2)	.0003
Age, years	73 (8.1)	73 (9.3)	72 (9.2)	.1152
Education, years	14.5 (3.2)	13.1 (365)	13.0 (4.1)	<.0001
Pre-progression rate	2.4 (2.0)	4.6 (3.0)	5.7 (3.)	<.0001
Slow	324 (50)	36 (10.6)	0 (0)	
Intermediate	261 (40)	197 (58)	47 (46.5)	<.0001
Rapid	66 (10)	106 (31.2)	54 (53.5)	
CDR-SB				
Female	519 (64)	131 (78)	36 (78)	.0006
Age, years	73.3 (8.2)	74.8 (9.5)	73.9 (9.9)	.0259
Education, years	14.2 (3.4)	13.1 (3.6)	13.3 (4.0)	<.0001
Pre-progression rate	2.9 (2.5)	4.5 (3.0)	4.9 (2.8)	<.0001
Slow	323 (39)	19 (11)	1 (2)	
Intermediate	357 (44)	93 (55)	30 (65)	<.0001
Rapid	129 (16)	56 (33)	15 (33)	
ADAS-Cog				
Female	139 (59.9)	448 (69.2)	5 (100)	.0101
Age, years	73 (7.5)	74 (7.9)	74.7 (9.3)	.1017
Education, years	15.4 (3.1)	14.2 (3.3)	13.4 (3.5)	.0001
Pre-progression rate	1.9 (1.8)	3.5 (2.5)	4.2 (1.1)	<.0001
Slow	135 (58.2)	182 (28.1)	0 (0)	
Intermediate	79 (34.0)	334 (51.6)	3 (60)	<.0001
Rapid	18 (7.8)	131 (20.3)	2 (40)	

Abbreviations: AD, Alzheimer's disease; ADAS-Cog, Alzheimer's Disease Assessment Scale–Cognitive; ADMDC, Alzheimer's Disease and Memory Disorders Center; CDR-SB, Clinical Dementia Rating Scale–Sum of Boxes; MMSE: Mini-Mental State Examination; SD, standard deviation.

Table 5. Misclassification Probabilities as "Mild" or "Severe" for Disease Severity Staged by Cognitive and Functional Measures Among the ADMDC Probable AD Cohort, Houston, Texas, 1990 to 2011.

	Estimate	Standard Error	*P* Value
MMSE			
Probability of misclassified as "mild"	0.078	0.008	<.0001
Probability misclassified as "severe"	0.003	0.005	<.0001
CDR-SB			
Probability of misclassified as "mild"	0.109	0.010	<.0001
Probability misclassified as "severe"	0.112	0.010	<.0001
ADAS-Cog			
Probability of misclassified as "mild"	0.055	0.005	<.0001
Probability misclassified as "severe"	0.023	0.004	<.0001

Abbreviations: AD, Alzheimer's disease; ADAS-Cog, Alzheimer's Disease Assessment Scale–Cognitive; ADMDC, Alzheimer's Disease and Memory Disorders Center; CDR-SB, Clinical Dementia Rating Scale–Sum of Boxes; MMSE: Mini-Mental State Examination.

longitudinal measurements of observed AD staging and baseline characteristics, examining the interplay of severity over time, determinants of these dynamic changes, and diagnostic accuracy of disease staging through the CTMC model.

The CDR-SB was estimated to misclassify $\sim 20\%$ of those that could be staged as moderate, which is higher than what might be expected for clinical utility, given the frequency in which the CDR-SB is used and reported reliability.[34] As this tool is used in both staging and diagnosis, this finding supports the literature that CDR-SB may be better served at distinguishing normal from dementia as a diagnostic tool as opposed to later staging of AD. For comprehensiveness, as we did think this number was higher than expected, we investigated the crude proportion of individuals who would have bordered the moderately severe threshold to explore the potential sensitivity prior to estimation. Using a 2-point error, 21% of observations fell within up to 2 units milder and none up to 2 units more severe. Considering 21% being possibly misclassified from moderate to mild or severe, the estimated 22% misclassification is reasonable.

With this rich cohort, we assessed thresholds of 3 neuropsychological instruments of cognitive and functional measures collected in the clinic. Although the MMSE has customarily been used to stage AD clinically, thresholds for classifications vary. Guidelines on CDR-SB and ADAS-Cog for classifying cognitive and AD disease severity, while believed to be sensitive to change, have not been supported by the literature. Not only does this research provide a comparison of sensitivity of staging disease but aims to corroborate diagnostic thresholds used to help clinicians identify, monitor, and treat patients at different levels of severity as well as to aid in patient selection in clinical trial enrollment for treatment development using routinely collected measures beyond the MMSE.

Table 6. Transition Intensities and Determinants of AD Stage Changes Using Functional and Cognitive Instruments Among the ADMDC Probable AD Cohort, Houston, Texas, 1990 to 2011.

	Transition Intensity/Transition Ratio[a]	95% CI
MMSE (n = 972)		
Base (transition intensity)		
$\exp\left(\hat{\beta}_{012}\right)$	0.97	0.44-2.15
$\exp\left(\hat{\beta}_{021}\right)$	0.05	0.02-0.14
$\exp\left(\hat{\beta}_{023}\right)$	0.71	0.32-1.57
Transition ratio		
Gender	1.08	0.92-1.28
Age, years	0.98	0.97-0.99
Education, years	1.01	0.98-1.03
Pre-progression	1.07	1.04-1.11
CDR-SB (n = 982)		
Base transition intensity		
$\exp\left(\hat{\beta}_{012}\right)$	0.34	0.15-0.74
$\exp\left(\hat{\beta}_{021}\right)$	0.03	0.01-0.08
$\exp\left(\hat{\beta}_{023}\right)$	0.24	0.11-0.54
Transition ratio		
Gender	1.13	0.96-1.34
Age, years	0.99	0.98-1.00
Education, years	1.00	0.98-1.02
Pre-progression	1.08	1.05-1.11
ADAS-Cog (n = 847)		
Base transition intensity		
$\exp\left(\hat{\beta}_{012}\right)$	0.24	0.03-1.80
$\exp\left(\hat{\beta}_{021}\right)$	0.00	0.00-0.02
$\exp\left(\hat{\beta}_{023}\right)$	0.01	0.00-0.07
Transition intensity		
Gender	1.05	0.73-1.51
Age, years	0.99	0.97-1.02
Education, years	1.03	0.97-1.09
Pre-progression	1.08	0.99-1.18

Abbreviations: AD, Alzheimer's disease; ADAS-Cog, Alzheimer's Disease Assessment Scale–Cognitive; ADMDC, Alzheimer's Disease and Memory Disorders Center; CDR-SB, Clinical Dementia Rating Scale–Sum of Boxes; CI, confidence interval; MMSE: Mini-Mental State Examination.

[a]Note that $e^{-\beta_r} = 1/e^{\beta_r}$ is the multiplicative impact on mean duration. Transition ratio interpreted for gender as rate of transition for males relative to females. For other variables, the estimated ratio corresponds to a 1-unit incremental change.

Age significantly impacted disease progression when MMSE was used to classify disease stages but not when severity was defined using CDR-SB. Furthermore, increased enrollment PPR was found to increase the rate of transition of disease when using MMSE and CDR-SB to stage AD severity but not when using ADAS-Cog. Our results also showed that years of education was associated with severity of AD at baseline in an unadjusted analysis; however, education level was not found to be a determinant of the disease process when analyzing the data longitudinally and with other covariates such as PPR. Similar findings have been reported and discussed previously when assessing education attainment related to AD progression[35] and cognitive decline[36] and support a passive cognitive reserve hypothesis described in the literature[37] that while education attainment delays onset of cognitive decline (due to higher cognitive performance), education does not slow the progression of AD as it relates to cognitive decline. Understanding how education indicates cognitive reserve will help clinicians to tailor treatment for the elderly patients and help researchers fine-tune progression models for clinical trial design and AD research.

The clinical utility of the progression results of this study is the estimation of a particular patient's mean duration in stages of severity. In other words, given a particular patient, we can determine the expected duration of stay in the mild stage prior to transitioning into the moderate stage given an individuals' PPR, age at onset (or first clinical visit), and so on. For example, for a 73-year-old female with an enrollment PPR of 3.5 MMSE points of worsening per year and 16 years of education, the average length of stay in mild stage of severity is estimated to be 3.1 years based on the MMSE. That same patient, if diagnosed using the CDR-SB or ADAS-Cog, would be expected to transition to the moderate stage after 3.1 and 2.9 years, respectively. Since the indications for antidementia drugs vary by disease stage, and since the need for nonpharmacologic and respite care varies by stage, these estimates are important for treatment and other care planning.

For purposes of application of our model, while the model allows estimability from mild to severe and severe to moderate states of disease, there were either 0 or not enough transitions of this nature in the data set; thus, the transition rate is near 0.

One might argue that the utilization of all 3 scores simultaneously would be optimal to study our research question and to make better clinical decisions. First, collection of clinical data varies in practice across clinicians as well as research projects. Second, this research allows us to understand how each staging criteria could be utilized prior to use in a multidimensional manner. Still, either extension of this method to model all 3 scores simultaneously or utilization of multiple measurements into a composite score could provide more information upon which to assess sensitivity of staging and finding determinants of the process itself. As the methods used in this article were recently developed and not part of standard statistical practices, future research is needed to incorporate these data in a multidimensional space.

This study has several limitations. First, it is a cohort of patients with probable AD and therefore difficult to compare cognitive functions to those who did not have the disease. Second, until 1995, the ADAS-Cog was not routinely conducted; thus, not all patients in the cohort were administered the ADAS-Cog. Further, analyses were limited to the measures collected in this cohort, which included the best known and most often used cognitive impairment screening tool. We

acknowledge that other well-validated, reliable measures are used to screen and identify possible mild cognitive impairment and probable dementia (eg, the Montreal Cognitive Assessment,[38] Telephone Interview for Cognitive Status,[39] among others)[40] A future study using our analytical methods in a large cohort with other outcome measures may be useful. Our results are based on the model adopted. There is no other model of a similar nature to use to study the goodness-of-fit when the data are potentially misclassified or incomplete.

From the methodological perspective, we have laid the groundwork for a unique approach at analyzing AD progression/regression useful in clinical decision-making. Methodologically, this is the first research of which we are aware that incorporates longitudinal data to estimate diagnostic accuracy (ie, sensitivity) of cognitive disease severity using an exact analytical form approach. By modeling the natural history of AD, the relationship between true stage changes and baseline covariates was evaluated. We examined transition intensities of AD stage change assuming unobservable underlying disease stages and estimated sensitivities of classifications of moderately diseased patients using established neuropsychological instruments as observed outcomes, which could lead to an improved determination of disease severity cutoff points.

Results of this study have important clinical and research-related ADAS-Cog implications in the AD population. Identifying stages of AD severity will allow researchers to measure disease progression and obtain more information regarding the sensitivity and specificity of neuropsychological measures. The models may be useful in evaluating pharmaceutical and behavioral intervention studies that aim to prevent onset or slow progression of AD. The ability to model the severity of AD is necessary in designing treatment plans which better manage the disease, therefore delaying the progression of disease and enhancing quality of life. Based on the estimates of misclassification, MMSE and ADAS-Cog are appropriate instruments to classify the moderate stage of disease.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Julia S. Benoit https://orcid.org/0000-0002-2734-5477

References

1. Thies W, Bleiler L. Alzheimer's disease facts and figures. *Alzheimer's Dement*. 2013;9(2):208-245.
2. Doody RS, Pavlik V, Massman P, Rountree S, Darby E, Chan W. Predicting progression of Alzheimer's disease. *Alzheimer's Res Ther*. 2010;2(1):2.
3. Baker E, Iqbal E, Johnston C, et al. Trajectories of dementia-related cognitive decline in a large mental health records derived patient cohort. *PloS one*. 2017;12(6):e0178562.
4. Samtani MN, Xu SX, Russu A, et al. Alzheimer's Disease Assessment Scale-Cognitive 11-item progression model in mild-to-moderate Alzheimer's disease trials of bapineuzumab. *Alzheimers Dement (N Y)*. 2015;1(3):157-169.
5. Doody RS, Massman P, Dunn JK. A method for estimating progression rates in Alzheimer disease. *Arch Neurol*. 2001;58(3):449-454.
6. Yu L, Griffith WS, Tyas SL, Snowdon DA, Kryscio RJ. A non-stationary Markov transition model for computing the relative risk of dementia before death. *Stat Med*. 2010;29(6):639-648.
7. Ito K, Corrigan B, Zhao Q, et al. Disease progression model for cognitive deterioration from Alzheimer's Disease Neuroimaging Initiative database. *Alzheimer's Dement*. 2011;7(2):151-160.
8. Rountree SD, Chan W, Pavlik VN, Darby EJ, Siddiqui S, Doody RS. Persistent treatment with cholinesterase inhibitors and/or memantine slows clinical progression of Alzheimer disease. *Alzheimers Res Ther*. 2009;1(2):7.
9. Folstein M, Folstein S, McHugh P. Mini-Mental State". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975;12(3):189-198.
10. Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. *Am J Psychiatry*. 1984;141(11):1356-1364.
11. Hughes CP, Berg L, Danziger W, Coben LA, Martin RL. A new clinical scale for the staging of dementia. *Brit J Psych*. 1982;140(6):566-572.
12. Wong DF, Rosenberg PB, Zhou Y, et al. In vivo imaging of amyloid deposition in Alzheimer disease using the radioligand 18F-AV-45 (florbetapir [corrected] F 18). *J Nucl Med*. 2010;51(6):913-920.
13. Clark CM, Pontecorvo MJ, Beach TG, et al. Cerebral PET with florbetapir compared with neuropathology at autopsy for detection of neuritic amyloid-β plaques: a prospective cohort study. *Lancet Neurology*. 2012;11(8):669-678.
14. Buerger K, Ewers M, Pirttilä T, et al. CSF phosphorylated tau protein correlates with neocortical neurofibrillary pathology in Alzheimer's disease. *Brain*. 2006;129(11):3035-3041.
15. Souza LCD, Sarazin M, Teixeira Júnior AL, Caramelli P, Santos AED, Dubois B. Biological markers of Alzheimer's disease. *Arq Neuropsiquiatr*. 2014;72(3):227-231.
16. Buchhave P, Blennow K, Zetterberg H, et al. Longitudinal study of CSF biomarkers in patients with Alzheimer's disease. *PLoS One*. 2009;4(7):e6294.
17. Mejia S, Gutiérrez LM, Villa AR, Ostrosky-Solis F. Cognition, functional status, education, and the diagnosis of dementia and mild cognitive impairment in Spanish-speaking elderly. *Appl Neuropsychol*. 2004;11(4):194-201.
18. Geerlings MI, Schmand B, Jonker C, Lindeboom J, Bouter LM. Education and incident Alzheimer's disease: a biased association due to selective attrition and use of a two-step diagnostic procedure? *Int J Epidemiol*. 1999;28(3):492-497.
19. Small BJ, Bäckman L. Longitudinal trajectories of cognitive change in preclinical Alzheimer's disease: a growth mixture modeling analysis. *Cortex*. 2007;43(7):826-834.

20. Devanand DP, Liu X, Tabert MH, et al. Combining early markers strongly predicts conversion from mild cognitive impairment to Alzheimer's disease. *Biol Psychiatry*. 2008;64(10):871-879.
21. de Jager CA, Schrijnemaekers AM, Honey TE, Budge MM. Detection of MCI in the clinic: evaluation of the sensitivity and specificity of a computerised test battery, the Hopkins verbal learning test and the MMSE. *Age Ageing*. 2009;38(4): 455-460.
22. O'Bryant SE, Waring SC, Cullum CM, et al. Staging dementia using Clinical Dementia Rating Scale Sum of Boxes scores: a Texas Alzheimer's Research Consortium Study. *Arch Neurol*. 2008;65(8):1091-1095.
23. O'Bryant SE, Lacritz LH, Hall J, et al. Validation of the new interpretive guidelines for the Clinical Dementia Rating Scale Sum of Boxes score in the National Alzheimer's Coordinating Center Database. *Arch Neurol*. 2010;67(6):746-749.
24. Rockwood K, Fay S, Gorman M, Carver D, Graham JE. The clinical meaningfulness of ADAS-Cog changes in Alzheimer's disease patients treated with donepezil in an open-label trial. *BMC Neurol*. 2007;7(1):26.
25. Birks JS. Cholinesterase inhibitors for Alzheimer's disease. *Cochrane Database Syst Rev*. 2006;5(1):CD005593.
26. Benge JF, Balsis S, Geraci L, Massman PJ, Doody RS. How well do the ADAS-Cog and its subscales measure cognitive dysfunction in Alzheimer's disease? *Dement Geriatr Cogn Disord*. 2009; 28(1):63-69.
27. Luo S, Chan W, Detry MA, Massman PJ, Doody RS. Binomial regression with a misclassified covariate and outcome. *Stat Methods Med Res*. 2016;25(1):101-117.
28. Brooks LG, Loewenstein DA. Assessing the progression of mild cognitive impairment to Alzheimer's disease: current trends and future directions. *Alzheimer's Res Ther*. 2010;2(5):28.
29. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. 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 on Alzheimer's Disease. *Neurology*. 1984;34(7):939-944.
30. Doody R, Pavlik V, Massman P, et al. Changing patient characteristics and survival experience in an Alzheimer's center patient cohort. *Dement Geriatr Cogn Disord*. 2005;20(2-3):198-208.
31. Doody RS, Dunn JK, Huang E, Azher S, Kataki M. A method for estimating duration of illness in Alzheimer's disease. *Dement Geriatr Cogn Disord*. 2004;17(1-2):1-4.
32. Wattmo C, Wallin ÅK, Minthon L. Progression of mild Alzheimer's disease: knowledge and prediction models required for future treatment strategies. *Alzheimer's Res Ther*. 2013;5(5):44.
33. Benoit JS, Chan W, Luo S, Yeh H, Doody R. A hidden Markov model approach to analyze longitudinal ternary outcomes when some observed states are possibly misclassified. *Stat Med*. 2016; 35(9):1549-1557.
34. de Jager CA, Schrijnemaekers AM, Honey TE, Budge MM. Detection of MCI in the clinic: evaluation of the sensitivity and specificity of a computerised test battery, the Hopkins Verbal Learning Test and the MMSE. *Age Ageing*. 2009;38(4):455-460.
35. Pavlik VN, Chan W, Darby E. Cohort effects in progression rate on cognitive and functional measures in an Alzheimer's disease clinical cohort. *J Alzheimer's Dis*. 2019;71(2):659-669.
36. Zahodne LB, Glymour MM, Sparks C, et al. Education does not slow cognitive decline with aging: 12-year evidence from the Victoria longitudinal Study. *J Int Neuropsychol Soci*. 2011; 17(6):1039-1046.
37. Stern Y. What is cognitive reserve? Theory and research application of the reserve concept. *J Int Neuropsychol Soci*. 2002;8(3): 448-460.
38. Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc*. 2005;53(4):695-699.
39. Brandt J, Spencer M, Folstein M. The telephone interview for cognitive status. *Neuropsychiatry Neuropsychol Behav Neurol*. 1988;1(2):111-117.
40. Chang YT, Chang CC, Lin HS, et al. Montreal cognitive assessment in assessing clinical severity and white matter hyperintensity in Alzheimer's disease with normal control comparison. *Acta Neurol Taiwan*. 2012;21(2):64-73.