ELSEVIER

Alzheimer's & Dementia ■ (2019) 1-9

Alzheimer's & Dementia

Featured Article

Artificially low mild cognitive impairment to normal reversion rate in the Alzheimer's Disease Neuroimaging Initiative

Kelsey R. Thomas[a,b], Joel S. Eppig[c], Alexandra J. Weigand[c], Emily C. Edmonds[a,b], Christina G. Wong[a,b], Amy J. Jak[a,b], Lisa Delano-Wood[a,b], Douglas R. Galasko[a,b,d], David P. Salmon[d], Steven D. Edland[e,f], Mark W. Bondi[a,b,*], for the Alzheimer's Disease Neuroimaging Initiative[1]

[a]*Veterans Affairs San Diego Healthcare System, San Diego, CA, USA*
[b]*Department of Psychiatry, University of California, San Diego, School of Medicine, La Jolla, CA, USA*
[c]*San Diego State University/University of California, San Diego Joint Doctoral Program in Clinical Psychology, San Diego, CA, USA*
[d]*Department of Neurosciences, University of California, San Diego, School of Medicine, La Jolla, CA, USA*
[e]*Department of Biostatistics, University of California, San Diego, La Jolla, CA, USA*
[f]*Department of Family and Preventative Medicine, University of California, San Diego, La Jolla, CA, USA*

Abstract

Introduction: We examined reasons for low mild cognitive impairment (MCI)-to-cognitively normal (CN) reversion rates in the Alzheimer's Disease Neuroimaging Initiative (ADNI).
Methods: CN and MCI participants were identified as remaining stable, progressing, or reverting at 1-year of follow-up (Year 1). Application of ADNI's MCI criteria at Year 1 in addition to Alzheimer's disease biomarkers by group were examined.
Results: The MCI-to-CN reversion rate was 3.0%. When specific components were examined, 22.5% of stable MCI participants had normal memory performance at Year 1 and their Alzheimer's disease biomarkers were consistent with the stable CN group. At Year 1, when all MCI criteria were not met, the more subjective Clinical Dementia Rating rather than objective memory measure appeared to drive continuation of the MCI diagnosis.
Discussion: Results demonstrate an artificially low 1-year MCI-to-CN reversion rate in ADNI-diagnosed participants. If the Logical Memory cutoffs had been consistently applied, the reversion rate would have been at least 21.8%.
Published by Elsevier Inc. on behalf of the Alzheimer's Association.

Keywords: Reversion; Mild cognitive impairment; Alzheimer's disease; Diagnostic criteria; Cerebrospinal fluid markers; Apolipoprotein E

Disclosure: D.R.G. serves as editor for *Alzheimer's Research and Therapy*, and as a paid consultant on Data Safety Monitoring Boards for Pfizer, Inc, Elan, Inc, and Balance Pharmaceuticals, Inc. D.P.S. serves as a consultant for Takeda Pharmaceuticals, Inc. S.D.E. is a paid consultant on Data Safety Monitoring Boards for Suven Life Science Limited and for Eli Lilly and Company. M.W.B. is a consulting editor for the *Journal of the International Neuropsychological Society*, and serves as a paid consultant for Novartis, Eisai, and Roche pharmaceutical companies. The other authors report no disclosures.

[1]Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

*Corresponding author. Tel.: +1-858-552-8585x2809; Fax: +1-858-642-1218.
E-mail address: mbondi@ucsd.edu

https://doi.org/10.1016/j.jalz.2018.10.008
1552-5260/Published by Elsevier Inc. on behalf of the Alzheimer's Association.

1. Introduction

Mild cognitive impairment (MCI) represents a transitional stage between normal cognitive aging and the onset of dementia, including Alzheimer's disease (AD) [1,2]. However, not all who are diagnosed with MCI exhibit progressive decline—many remain at the level of MCI and a significant portion (up to 30%–50%) revert to a cognitively normal (CN) state, depending on diagnostic criteria [3,4]. Recent meta-analyses found reversion rates of 18% [5] and 24% [6], with community-based studies containing a greater proportion of "reverters" (31%) than clinically based studies (14%) [6]. In stark contrast, 1-year reversion rate within the Alzheimer's Disease Neuroimaging Initiative (ADNI) was previously estimated at 2.2% [7].

ADNI's MCI diagnostic criteria are relatively consistent with the conventional criteria of Petersen et al. and Winblad et al. [7–9]. Specifically, ADNI relies on subjective cognitive complaint, global cognitive screening, a self-informed and study partner–informed global functioning interview, and one memory measure [7]. Alternative neuropsychological actuarial MCI criteria [10,11] have reliably identified a large proportion of individuals with "false-positive" MCI diagnoses who, despite receiving a diagnosis of MCI in ADNI, exhibit biomarker profiles, neuropsychological performance, and functional trajectories most consistent with CN participants [11–16]. The susceptibility of these diagnostic errors and the lack of clear guidelines for the application of the individual components of ADNI's MCI criteria at follow-up occasions suggest that the basis of the very low 1-year MCI-to-CN reversion rate warrants further investigation.

Therefore, we comprehensively examined (1) 1-year diagnostic stability and reversion rates in ADNI, (2) possible explanations for ADNI's very low reversion rate, and (3) how stable MCI participants who would have been better classified as reverters differed from other stable MCI participants in terms of biomarkers, cognition, depressive symptoms, and rates of progression to dementia. We hypothesized, based on previous work in ADNI [7], that (1) we would replicate the low MCI-to-CN reversion rates with a larger ADNI sample, (2) diagnoses at the 1-year follow-up examination (Year 1) would be driven more by clinical judgment (i.e., Clinical Dementia Rating [CDR]) than objective memory performance, and (3) those participants who were diagnosed as stable MCI despite having normal objective memory performance at Year 1 would have biomarker profiles and progression rates that are more similar to stable CN participants than stable MCI participants.

2. Methods

Data used in the preparation of this article were obtained from the ADNI database (adni.loni.usc.edu). ADNI was launched in 2003 as a public-private partnership. The primary goal of ADNI has been to test whether serial magnetic resonance imaging, positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. For up-to-date information on ADNI, see www.adni-info.org. This study was approved by the Institutional Review Boards at each of the participating institutions, and written informed consent was obtained from all participants or authorized representatives at each site.

2.1. Participants

Specific enrollment criteria for ADNI have been described elsewhere [7]. All nondemented ADNI participants from ADNI 1, ADNI GO, and ADNI 2 cohorts who completed a baseline and 1-year neuropsychological assessment were considered for analyses (total N = 1208). Table 1 shows the specific ADNI diagnostic criteria for distinguishing CN, MCI, and AD participants [7,17]. At baseline, all MCI participants were considered "amnestic" (single domain or multidomain) except two participants who were considered "nonamnestic" MCI (one progressed to AD at Year 1, one remained MCI at Year 1).

2.2. Procedure

2.2.1. Components of diagnostic criteria

For CN and MCI participants, cutoffs for the Mini-Mental State Examination (MMSE; general cognition), Logical Memory II (LM; memory), and CDR (global function) were examined in more detail to determine whether these individual components were applied at Year 1 and which components of the diagnostic criteria were prioritized when cutoffs for all three components were not met. For these analyses, the diagnostic criteria components/cutoffs were based on those shown in Table 1. For MCI participants, the early MCI (E-MCI) LM cutoffs (rather than the late MCI [L-MCI] cutoffs) were used because, depending on the phase of ADNI, the MCI LM cutoffs varied, and the E-MCI cutoffs are the most diagnostically inclusive (i.e., more MCI participants will meet the less-strict cutoffs than if the L-MCI cutoffs were used). Importantly, only the CDR component had fully independent cutoffs for CN and MCI participants because there was overlap between the CN and E-MCI LM component and complete overlap between CN and MCI on other criteria (e.g., MMSE) except for the subjective complaint.

2.2.2. Additional measures

A large subset of participants underwent a baseline lumbar puncture (CN, n = 289; MCI, n = 581). AD cerebrospinal fluid (CSF) marker positivity was examined by diagnostic group, and biomarkers were measured using Elecsys immunoassays. Cutoff scores proposed by Hansson et al. [18] and optimized for ADNI were used to determine biomarker positivity: <976.6 pg/mL for amyloid β

Table 1
Specific components of ADNI classification criteria to distinguish CN, early MCI, late MCI, and AD

Component	CN	Early MCI	Late MCI	AD
Subjective complaint	None, aside from those common to other normal subjects of that age range	Yes, by subject (verified by study partner) or study partner or clinician	Yes, by subject (verified by study partner) or study partner or clinician	Yes, by subject (verified by study partner) or study partner or clinician
MMSE*	≥24	≥24	≥24	20–26 (Inclusive)
LM	≥9 for 16+ y of education ≥5 for 8–15 y of education ≥3 for 0–7 y of education	9–11 for 16+ y of education 5–9 for 8–15 y of education 3–6 for 0–7 y of education	≤8 for 16+ y of education ≤4 for 8–15 y of education ≤2 for 0–7 y of education	≤8 for 16+ y of education ≤4 for 8–15 y of education ≤2 for 0–7 y of education
CDR	CDR = 0 Memory Box score must be 0	CDR = 0.5 Memory Box score of at least 0.5	CDR = 0.5 Memory Box score of at least 0.5	CDR = 0.5 or 1.0
General cognition and functional status	CN, based on the absence of significant impairment in cognitive functions or activities of daily living	General cognition and functional performance sufficiently preserved such that a diagnosis of AD cannot be made	General cognition and functional performance sufficiently preserved such that a diagnosis of AD cannot be made	NINCDS/ADRDA criteria for probable AD

Abbreviations: AD, Alzheimer's disease; ADNI, Alzheimer's Disease Neuroimaging Initiative; CDR, Clinical Dementia Rating; CN, cognitively normal; LM, Logical Memory; MCI, mild cognitive impairment; MMSE, Mini-Mental State Examination; NINCDS/ADRDA, National Institute of Neurological and Communication Disorders and Stroke/Alzheimer's Disease and Related Disorders Association.

NOTE. LM impairment was based on education-adjusted cutoffs on one paragraph of the LM II subscale of the Wechsler Memory Scale—Revised (maximum score of 25). Late MCI criteria were the only MCI criteria for ADNI 1 phase; early MCI criteria were only included in ADNI GO and ADNI 2 phases. This table was adapted from the procedure manuals for ADNI 1, ADNI GO, and ADNI 2 available at http://adni.loni.usc.edu/methods/documents/.

*MMSE exceptions may be made for subjects with <8 years of education at the discretion of the project director.

(1–42) (Aβ), >0.0251 pg/mL for the hyperphosphorylated-tau (p-tau)/Aβ ratio, and >0.27 pg/mL for the total-tau (t-tau)/Aβ ratio. Participants with at least one apolipoprotein E (*APOE*) ε4 allele were also considered at risk (i.e., ε4-positive). The MMSE was used to measure global cognition, and the Geriatric Depression Scale (GDS) was used to measure depressive symptoms. The Rey Auditory Verbal Learning Test (AVLT) delayed recall score was used as a second measure of memory, in addition to LM, in post hoc analyses.

2.3. Statistical analyses

Baseline demographic and clinical characteristics by group (CN and MCI) were compared using independent *t* tests or χ^2 tests. Group differences in AD biomarker positivity were examined using χ^2 tests. Group differences in LM were examined using a one-way analysis of variance, and MMSE and depressive symptoms were examined using Kruskal-Wallis tests; Bonferroni corrections were used for multiple comparisons. A Cox regression adjusting for age, sex, and education was used to determine the hazard ratio (HR) and 95% confidence interval (CI) of progression to dementia. In these analyses, time-to-dementia was the number of months from the baseline assessment to the assessment when the participant first met criteria for dementia. Participants who did not progress to dementia during their follow-up period (range of follow-up, 12–60 months) were censored at their last visit. Kaplan-Meier curves were used to show the rate of progression to dementia by group. A post hoc mixed between-within analysis of variance examining the change in memory scores (LM and AVLT) from baseline to Year 1 by group, with Bonferroni-corrected post hocs, assessed for differential practice effects/regression to the mean.

3. Results

3.1. Baseline characteristics

At baseline, there were 420 CN and 788 MCI participants based on ADNI diagnoses. CN and MCI diagnostic groups significantly differed from one another on all demographic, functional, neuropsychological, and biomarker data examined (Table 2).

3.2. Classification at Year 1

Examination of stability, progression, and reversion using ADNI's reported diagnoses showed that among participants classified as CN at baseline (N = 420), 403 (96.0%) remained CN, 17 (4.0%) progressed to MCI, and 0 (0%) progressed to dementia at Year 1. Of those who were classified as MCI at baseline (N = 788), 661 (83.9%) remained MCI, 24 (3.0%) reverted to CN, and 103 (13.1%) progressed to dementia at Year 1. Stability, progression, and reversion rates for longer follow-up durations (e.g., baseline to Year 2, baseline to Year 3, and so forth) for CN and MCI participants are included in Supplementary Tables 1 and 2.

3.3. Follow-up on ADNI criteria

3.3.1. Components of criteria

At baseline, 99.5% of CN participants met all three required components of the CN criteria that were examined. At Year 1, however, 14.5% of stable CN participants did not meet all three components of the criteria (1 component

Table 2
Baseline demographic and clinical characteristics (mean [SD] or %) by group

Variable	CN (N = 420)	MCI (N = 788)	*P* value
Age	74.70 (5.76)	73.18 (7.54)	<.001
Education	16.34 (2.67)	15.99 (2.81)	.036
Female (%)	48.9%	40.8%	.007
GDS	0.76 (1.08)	1.68 (1.43)	<.001
MMSE	29.07 (1.14)	27.61 (1.80)	<.001
CDR = 0.5 (%)	0.5%	99.7%	<.001
CDR Memory = 0.5 (%)	0.6%	96.9%	<.001
FAQ	0.22 (0.94)	3.14 (4.06)	<.001
Logical Memory II (raw)	13.16 (3.30)	5.68 (3.42)	<.001
AVLT delayed recall (raw)	7.56 (3.89)	3.90 (3.86)	<.001
Aβ	1352.57 (659.56)	1013.88 (553.69)	<.001
p-tau/Aβ	0.02 (0.02)	0.04 (0.03)	<.001
t-tau/Aβ	0.23 (0.18)	0.36 (0.27)	<.001
p-tau	21.70 (8.91)	27.81 (14.19)	<.001
t-tau	236.39 (87.09)	287.19 (126.30)	<.001
APOE ε4+ (%)	27.6%	51.1%	<.001

Abbreviations: Aβ, amyloid β; *APOE*, apolipoprotein E; AVLT, Rey Auditory Verbal Learning Test; CDR, Clinical Dementia Rating; CN, cognitively normal; FAQ, Functional Activities Questionnaire; GDS, Geriatric Depression Scale; MCI, mild cognitive impairment; MMSE, Mini-Mental State Examination; p-tau, hyperphosphorylated tau; SD, standard deviation; t-tau, total tau.

NOTE. For the cerebrospinal fluid markers, the CN group had n = 289 and MCI group had n = 581.

n = 4, 1.1%; 2 components n = 53, 13.5%). Table 3 shows the decomposition of the specific ADNI criteria components (MMSE, CDR, and LM) at Year 1.

Given the very low baseline-to-Year 1 reversion rate (3.0%), we examined specific components of ADNI's MCI criteria (MMSE, CDR, and LM) to determine if there was a subgroup of stable MCI participants who might have been better characterized as CN at Year 1 (i.e., possible reverters, not classified as such). At baseline, 99.0% of MCI participants met all three required components of the MCI criteria that were examined. At Year 1, however, about one-third (34.9%) of stable MCI participants did not meet all three components of the criteria (1 component n = 15, 2.3%; 2 components n = 215, 32.6%). Of the participants who did not meet all three criteria (N = 230), the component that was most frequently not met at Year 1 was LM (64.3%; 148 of 230 MCI participants). Had those participants who did not perform at less than the LM cutoffs at Year 1 (*MMSE + CDR only* group) been instead diagnosed as CN at Year 1, the ADNI criteria would have shown a MCI-to-CN reversion rate of 21.8% (see Fig. 1). This is a conservative estimate using the E-MCI criteria and not including the 15 participants (14 of whom did not meet the LM cutoffs) who only met one of the components of the MCI criteria at Year 1.

Table 3
Specific components of the criteria met at year 1 for stable CN and stable MCI groups

	Stable CN (N = 392)	Stable MCI (N = 659)
Components	N (%)	N (%)
1 Component (total)	4/1.1%	15/2.3%
MMSE only	0/0.0%	13/2.0%
CDR only	3/0.8%	1/0.2%
LM only	1/0.3%	1/0.2%
2 Components (total)	53/13.5%	215/32.6%
MMSE + LM	26/6.6%	32/4.9%
CDR + LM	0/0.0%	35/5.3%
MMSE + CDR	27/6.9%	148/22.5%
3 Components		
MMSE + CDR + LM	335/85.5%	429/65.1%

Abbreviations: ADNI, Alzheimer's Disease Neuroimaging Initiative; CDR, Clinical Dementia Rating; CN, cognitively normal; LM, Logical Memory II, Delayed Recall of one paragraph; MCI, mild cognitive impairment; MMSE, Mini-Mental State Examination.

NOTE. CN had 11 participants and MCI had two participants who had missing data for one or more elements of the criteria, but were given an ADNI diagnosis. The MCI MMSE + LM only group included 10 participants with CDR = 1.0 and 22 participants with CDR = 0.

3.3.2. Biomarker analyses

Because the LM component of the criteria was the most frequently unmet component at Year 1 among "stable" MCI participants, we examined AD-related CSF and genetic susceptibility markers and compared the *MMSE + CDR only* group (did not meet LM criteria at Year 1) to *stable CN* (CN at both baseline and Year 1) and *stable MCI* participants who met all three criteria at Year 1 (n = 429).

The χ^2 analyses indicated significant differences between stable CN, MMSE + CDR only, and stable MCI participants on proportions of participants who were positive for Aβ [$\chi^2(2) = 59.89$, $P < .001$], p-tau/Aβ ratio [$\chi^2(2) = 74.24$, $P < .001$], t-tau/Aβ ratio [$\chi^2(2) = 77.66$, $P < .001$], and at least one *APOE* ε4 allele [$\chi^2(2) = 57.88$, $P < .001$].

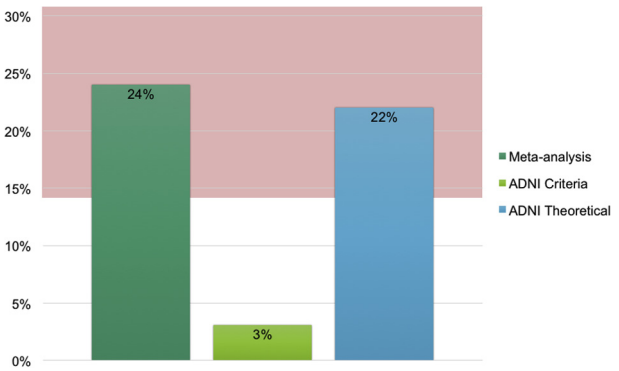


Fig. 1. Comparison of meta-analysis, ADNI-based diagnosis, and theoretical ADNI reversion at Year 1. Dark green bar represents the average reversion rate across studies from a recent meta-analysis [6]. Light green bar represents the baseline-to-Year 1 reversion rate based on the diagnoses in ADNI. Blue bar represents the theoretical ADNI reversion rate if the MMSE + CDR only group was classified as CN at Year 1 instead of MCI. Red box represents the 14% and 31% rate of reversion for clinical (14%) and community-based (31%) samples, respectively, found by a recent meta-analysis [6]. Abbreviation: ADNI, Alzheimer's Disease Neuroimaging Initiative.

Follow-up analyses revealed that only the stable MCI group differed from stable CN and MMSE + CDR only groups on AD biomarkers (all *P* values <.001), whereas the stable CN and MMSE + CDR only groups did not differ on any AD biomarkers (*P* values ranged from .202 to .992; see Fig. 2).

3.3.3. Cognition and depressive symptom analyses

There were significant group differences in LM [$F(2, 954) = 732.64$, $P < .001$], MMSE [*Kruskal-Wallis* $\chi^2(2) = 201.63$, $P < .001$], and GDS [*Kruskal-Wallis* $\chi^2(2) = 85.45$, $P < .001$] scores. LM performance at Year 1 did not differ between the stable CN (mean = 13.66, standard deviation [SD] = 4.01) and MMSE + CDR only (mean = 13.71, SD = 2.15) groups ($P > .99$), but both groups performed better than the stable MCI group (mean = 4.93, SD = 3.46, *P* values <.001). MMSE scores did not differ between the stable CN (mean = 28.96, SD = 1.30) and MMSE + CDR only (mean = 28.80, SD = 1.43) groups ($P = .879$). However, the stable MCI group had a lower MMSE score (mean = 27.37, SD = 1.78) than both the stable CN ($P < .001$) and MMSE + CDR only ($P < .001$) groups. GDS scores, however, differed between the stable CN (mean = 0.95, SD = 1.42) and both the MMSE + CDR only (mean = 1.66, SD = 1.83) and stable MCI (mean = 1.92, SD = 1.96) groups (both *P* values <.001), but the MMSE + CDR only and stable MCI groups did not differ from one another ($P = .263$).

3.3.4. Progression to dementia

Differential rates of progression to dementia were examined between groups. Compared with the stable CN group, Cox regressions showed that both the MMSE + CDR only group (HR = 5.98, 95% CI = 2.07–17.25, $P = .001$) and, to a much greater extent, the stable MCI group (HR = 36.15, 95% CI = 14.78–88.37; $P < .001$) had greater risk of progression to dementia. Compared with the MMSE + CDR only group, the stable MCI group had greater risk of progression to dementia (HR = 6.04, 95% CI = 3.26–11.19; $P < .001$). Kaplan-Meier curves and numbers of events/persons at risk are shown in Fig. 3.

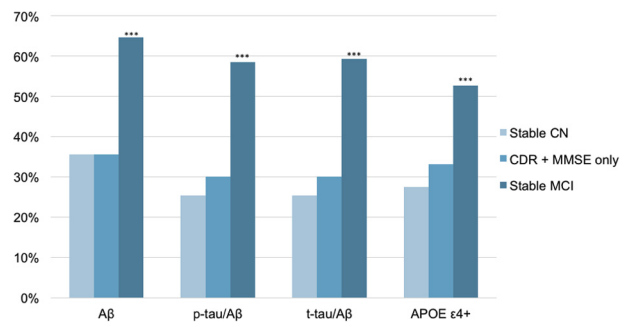


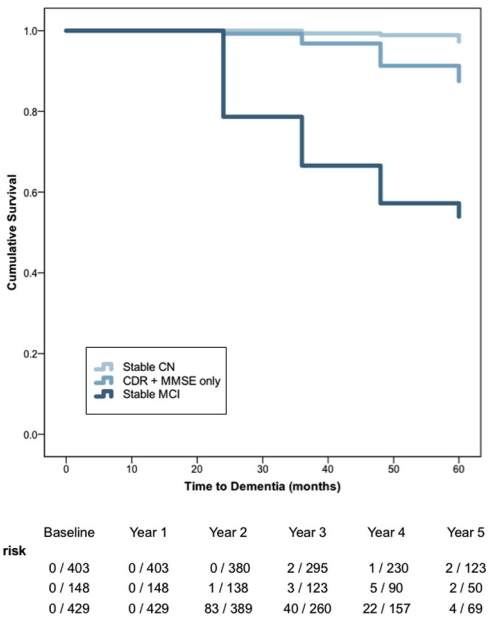
Fig. 2. Proportion of stable CN, MMSE + CDR only, and stable MCI groups with positive AD biomarkers. ***$P < .001$; For the CSF markers (Aβ, p-tau/Aβ, and t-tau/Aβ), the stable CN group had n = 276, CDR + MMSE only group had n = 110, and stable MCI group had n = 324. Abbreviations: Aβ, amyloid β; *APOE*, apolipoprotein E; CDR, Clinical Dementia Rating; CN, cognitively normal; MCI, mild cognitive impairment; MMSE, Mini-Mental State Examination; p-tau, hyperphosphorylated tau; t-tau, total tau.

3.4. Post hoc analysis of change in memory

Change in objective memory performance from baseline to Year 1 by group was examined. For LM, there was a significant group × time interaction [$F(2, 975) = 188.97$, $P < .001$, $\eta_p^2 = .279$] such that the stable MCI group declined and stable CN and MMSE + CDR only groups significantly improved from baseline to Year 1 (all *P* values < .01); however, the MMSE + CDR only group demonstrated accelerated improvement ($P < .001$) and no longer differed from the stable CN group at Year 1 (see Supplementary Fig. 1). For AVLT delayed recall, there was a small, but significant, group × time interaction [$F(2, 974) = 8.49$, $P < .001$, $\eta_p^2 = .017$], but the stable CN and MMSE + CDR only groups did not differ in AVLT performance and did not have a significant improvement from baseline to Year 1 (*P* values >.05); instead, the stable MCI group significantly declined (see Supplementary Fig. 2).

4. Discussion

The present study found that there was significant diagnostic stability of participants initially identified as CN and MCI across assessments, as would be generally expected over the course of a relatively short period of time. However,



Events / Number at risk	Baseline	Year 1	Year 2	Year 3	Year 4	Year 5
Stable CN	0 / 403	0 / 403	0 / 380	2 / 295	1 / 230	2 / 123
CDR + MMSE only	0 / 148	0 / 148	1 / 138	3 / 123	5 / 90	2 / 50
Stable MCI	0 / 429	0 / 429	83 / 389	40 / 260	22 / 157	4 / 69

Fig. 3. Kaplan-Meier curves for stable CN, MMSE + CDR only, and stable MCI group progression to dementia. Abbreviations: CDR, Clinical Dementia Rating; CN, cognitively normal; MCI, mild cognitive impairment; MMSE, Mini-Mental State Examination.

the reversion rate of 3% appears to be substantially lower than data reported in meta-analyses showing typical reversion rates of 18% [5] and 24% [6]. Furthermore, 3% is notably lower than the most conservative estimates of reversion in clinical samples of 8% [5] and 14% [6], despite the fact that these meta-analyses included several studies that used similar criteria to those used in ADNI (i.e., Mayo and Petersen et al. [8] or International Working Group [9] criteria).

This very low reversion rate appeared to be at least partly related to incomplete application of the components of ADNI's cognitive diagnostic criteria (MMSE, CDR, and LM) at Year 1, as only 65.1% of those considered stable MCI continued to meet the cutoffs for all three components of the criteria at Year 1. The largest portion of participants who did not meet all three components of the criteria at follow-up was a group who met the MMSE and CDR components, but not LM (22.5% of the stable MCI sample). These MMSE + CDR only participants performed above the MCI cutoffs on the LM component of the criteria. Because the MMSE cutoff was the same for both ADNI CN and MCI classifications, it appears that ADNI diagnosticians weighted the CDR more heavily than LM, despite the CDR's reliance on subjective reporting from the participant and study partner. ADNI's diagnostic reliance on the CDR is also embedded within these three operationalized components of the criteria; only the CDR has completely mutually exclusive categories for identifying CN versus MCI participants. As mentioned previously, the MMSE cutoffs do not discriminate between CN and MCI participants (the criteria completely overlap). For LM, there is considerable overlap between CN and E-MCI cutoffs for ADNI GO and ADNI 2. Thus, the diagnosis of individuals with objective test performance that falls in this "borderzone" (i.e., CN or E-MCI based on MMSE and LM scores) relies solely on the more subjective CDR ratings.

These findings are consistent with our previous work showing that about one-third of ADNI-defined MCI participants have cognitive functioning [12], AD biomarkers [12–14,16], and everyday functioning [15] that is more consistent with CN than MCI participants (i.e., false-positive MCI diagnosis). The present study adds to our previous work by showing that, when more weight is given to more subjective (CDR) than objective (LM) information in the diagnostic decision-making process, there is a greater propensity for diagnostic misclassification, including artificially high rates of MCI stability.

Inconsistencies between the subjective reporting of the participant or study partner on the CDR and objective memory performance on LM (i.e., over-reporting cognitive or functional difficulties on the CDR in the context of intact objective performance) may be at least partially related to depressive symptoms [19–21]. Despite normal performance on objective memory (LM) and global cognition (MMSE) tests at Year 1, participants in the MMSE + CDR only group reported more depressive symptoms than the stable CN group. However, the severity of their depressive symptoms did not differ from symptoms of the stable MCI group. Although the severity of the depressive symptoms reported by both the MMSE + CDR only and stable MCI groups were well below clinically significant levels, the difference in reported symptoms at Year 1 between the MMSE + CDR only and the stable CN groups may contribute to the inconsistency between LM performance and CDR score for the MMSE + CDR only group.

When the MMSE + CDR only group was compared with stable MCI participants who met all three components of the MCI criteria and stable CN participants, results showed that the MMSE + CDR only group had a lower proportion of positive AD biomarkers (CSF and genetic susceptibility) at baseline than the stable MCI group, and they did not differ from the stable CN participants on any of these biofluid and genetic susceptibility markers. A similar pattern of findings was found for objective memory performance and global cognition at Year 1, in that both the stable CN and MMSE + CDR only groups had better performance than the stable MCI group but did not differ from one another. These findings suggest that the MMSE + CDR only group would have been better characterized as CN (and with low biomarker positivities) and not MCI at Year 1. Had the MMSE + CDR only participants been diagnosed as CN at Year 1 follow-up (i.e., reverter), the ADNI reversion rate would have been at least 21.8%, which is consistent with two recent meta-analyses showing rates of reversion at 18% [5] and 24% [6].

When the rates of progression to dementia for more than 5 years were examined, the stable MCI participants progressed to dementia at a significantly faster rate than the stable CN or the MMSE + CDR only groups. Although the MMSE + CDR only group progressed to dementia faster than the stable CN group, this risk was much lower than that of the stable MCI group, consistent with the differences in biomarker positivity. Prior work suggests that MCI participants that revert to CN status ultimately progress to dementia at a faster rate than those initially diagnosed as CN [4]. The HR for risk of dementia in the MMSE + CDR only group (HR = 5.98) is consistent with previous studies examining the risk of progression to dementia in reverters compared with baseline normal participants (e.g., HRs of 6.6 in the Mayo study [4] and 6.4 in the Sydney Memory and Aging Study [22]).

The reasons for what appears to be a greater focus on the CDR than the LM performance at Year 1 follow-up are unclear, although bias (e.g., expectancy, confirmation, and anchoring) might be a possibility if the diagnosticians had awareness of a participant's prior year's diagnosis when deciding on their Year 1 diagnosis. This notion is consistent with past work that has also noted how bias may negatively impact diagnostic decisions if researchers are not blinded to

past diagnoses [3]. Furthermore, based on the protocol manuals available on the ADNI website (http://adni.loni.usc.edu/methods/documents/), the protocol appears to be vague regarding how to diagnose participants at follow-up visits and how to handle inconsistent results among the components of the classification criteria. An additional factor that likely made classification difficult was the inconsistency in the LM criteria between ADNI 1 and ADNI GO/ADNI 2. In ADNI GO and ADNI 2, LM was used to distinguish E-MCI from L-MCI. Thus, participants were classified as MCI using slightly different criteria between ADNI waves. The current investigation used the E-MCI criteria for all participants because this was more lenient (i.e., LM cutoff for E-MCI is more easily met than the cutoff for L-MCI). However, this approach provides the most conservative estimate of individuals that did not meet all three components of the MCI criteria, and it is likely that even fewer participants who were considered by ADNI to be stable MCI from baseline to Year 1 would have met the MCI LM criteria at Year 1 had the L-MCI criteria been applied.

The findings that a large proportion of ADNI MCI participants were not accurately identified as reverters and instead retained a potentially false-positive diagnosis of MCI could have significant clinical and research implications. Maintaining a diagnosis of MCI when objective testing suggests reversion may have clinical and research consequences. From a psychological perspective, being consistently diagnosed with MCI despite having normal memory performance has unknown, yet potentially detrimental effects. Prior work has shown that knowledge of one's own *APOE* ε4 allele status, for example, may have a negative effect on both subjective and objective memory performances relative to *APOE* ε4 carriers without knowledge of their genotype [23]. In addition, there are significant implications of including these potential reverters in analyses that examine the longitudinal trajectories of AD biomarkers. We found the MMSE + CDR only group to have very similar AD biomarker profiles as stable CN participants. Therefore, participants who are mislabeled as "stable MCI" but do not actually meet criteria for MCI may be negatively impacting our understanding of AD biomarker trajectories or washing out effects of biomarkers as predictors of progression to dementia.

One potential difficulty in relying on objective memory performance in serial assessment is the accounting for possible practice effects. Previous work has identified significant practice effects of memory tests within ADNI [24], so we speculated that perhaps ADNI diagnosticians considered the potential for practice effects when diagnosing MCI at Year 1 despite normal LM performance. However, examination of the baseline-to-Year 1 change in LM and AVLT memory measures showed that stable CN and MMSE + CDR only groups did not differ at Year 1 on either measure and did not differ on AVLT performance at baseline. Together, these data suggest that the MMSE + CDR only change in performance from baseline to Year 1 is more likely regression to the mean than an elevated practice effect given the much smaller practice effect (likely not because of ceiling effects) of the stable CN group on LM and the similar performances at both baseline and Year 1 on the AVLT. Notably, these findings add to the literature that highlights the importance of using multiple objective neuropsychological measures in a diagnosis of MCI [10,25]. There has been considerable work showing that, within a neurologically normal population, the proportion of individuals with at least one impaired score is high [26–28], again emphasizing the importance on balancing reliability of impairment through the use of more than one test in a domain.

It is important to note that our study focused on the 1 year of follow-up for examining diagnostic stability. Future work examining the stability, reversion, and progression rates over longer follow-up periods in more detail will expand on the present findings. In addition, future work should examine the precision with which other large aging studies adhere to their diagnostic criteria and how diagnostic decisions are made when the criteria are mixed. This type of detailed inspection of diagnostic criteria may be particularly relevant to clinical trials because false positive errors and artificial diagnostic stability could mask potential results [29].

There is a small subset of ADNI participants with autopsy data, and it would be ideal to determine that the MMSE + CDR only group did not go on to develop AD pathology, although autopsy confirmation of individuals who were CN or MCI at baseline is not without issues. Most older adults with normal cognition or mild cognitive changes will live many years past the date of an initial observation, and autopsies will capture new pathology that may have developed after that baseline observation. However, reliance on CSF biomarkers is also not without limitations. Although CSF $A\beta_{1\text{-}42}$ and tau-Aβ ratios (t-tau/Aβ, p-tau/Aβ) have shown strong concordance with Aβ PET and level of cognitive impairment [18,30], CSF AD biomarkers are limited in that there is no way to determine regional patterns that may be evident on PET. Furthermore, CSF $A\beta_{1\text{-}42}$ levels have been shown to be decreased in other neurodegenerative processes (i.e., Lewy body dementia) [31,32]; similarly, CSF tau levels have been shown to be increased in non-AD pathologies (e.g., Creutzfeldt-Jakob disease, stroke) [33]. Despite these limitations, the ability to incorporate CSF AD markers that correspond to their baseline assessment is a relative strength of the ADNI study.

The present study, consistent with prior ADNI work, showed a very low MCI-to-CN reversion rate over 1 year. It also demonstrated that when all features of the criteria were not met, the more subjective CDR was weighted more heavily than the objective memory measure (LM) in the diagnostic decision-making process, resulting in an

artificially low reversion rate. These findings are further supported through biofluid and genetic markers, progression rates, and memory changes, which all suggest that baseline MCI participants with normal memory scores at Year 1, despite a CDR score of 0.5, would have likely been better classified as CN at Year 1. Our future directions for this work include examination of diagnostic stability, reversion, and progression rates in ADNI using MCI criteria that are based on actuarial neuropsychological test performances [10,11]. The use of objective neuropsychological scores in a way that balances sensitivity and reliability (i.e., two impaired cognitive scores within a cognitive domain, rather than one impaired LM score) may provide a method to more precisely identify stable MCI participants and differentiate them from true reverters who are CN at follow-up.

Acknowledgments

This work was supported by NIH grants R01 AG049810 (M.W.B.), K24 AG026431 (M.W.B.), and P50 AG005131 (D.R.G., D.P.S., S.D.E.), the Alzheimer's Association (AARF-17-528918 to K.R.T. and AARG-17-500358 to E.C.E.), and the U.S. Department of Veterans Affairs Clinical Sciences Research and Development Service (Career Development Award-2 1IK2 CX001415-01A1 to E.C.E.). Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc; Cogstate; Eisai Inc; Elan Pharmaceuticals, Inc; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc; Fujirebio; GE Healthcare; IXICO Ltd; Janssen Alzheimer Immunotherapy Research & Development, LLC; Johnson & Johnson Pharmaceutical Research & Development LLC; Lumosity; Lundbeck; Merck & Co, Inc; Meso Scale Diagnostics, LLC; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jalz.2018.10.008.

RESEARCH IN CONTEXT

1. Systematic review: The authors reviewed studies (e.g., using PubMed) related to mild cognitive impairment (MCI) diagnostic stability and reversion. MCI-to-normal reversion rates vary widely, depending on the study and MCI criteria. The reversion rate previously reported in Alzheimer's Disease Neuroimaging Initiative (ADNI) was lower than all other studies included in a recent meta-analysis.
2. Interpretation: Results suggest that the low reversion rate in ADNI may be driven by weighting the subjective Clinical Dementia Rating more heavily than the objective memory test, resulting in the continuation of the MCI diagnosis at 1-year of follow-up, despite normal memory and biomarkers. Had those subjects with normal memory performance been classified as cognitively normal at follow-up, the ADNI reversion rate would be more consistent with the literature.
3. Future directions: Future work will examine MCI-to-normal reversion using MCI criteria based on objective neuropsychological test performances. We will determine whether a comprehensive neuropsychological approach provides an improved method for differentiating stable MCI participants from "reverters" who are cognitively normal at follow-up.

References

[1] Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011;7:270–9.

[2] Manly JJ, Tang M-X, Schupf N, Stern Y, Vonsattel J-PG, Mayeux R. Frequency and course of mild cognitive impairment in a multiethnic community. Ann Neurol 2008;63:494–506.

[3] Pandya SY, Clem MA, Silva LM, Woon FL. Does mild cognitive impairment always lead to dementia? A review. J Neurol Sci 2016; 369:57–62.

[4] Roberts RO, Knopman DS, Mielke MM, Cha RH, Pankratz VS, Christianson TJH, et al. Higher risk of progression to dementia in mild cognitive impairment cases who revert to normal. Neurology 2014;82:317–25.

[5] Canevelli M, Grande G, Lacorte E, Quarchioni E, Cesari M, Mariani C, et al. Spontaneous reversion of mild cognitive impairment to normal cognition: A systematic review of literature and meta-analysis. J Am Med Dir Assoc 2016;17:943–8.

[6] Malek-Ahmadi M. Reversion from mild cognitive impairment to normal cognition. Alzheimer Dis Assoc Disord 2016;30:324–30.

[7] Petersen RC, Aisen PS, Beckett LA, Donohue MC, Gamst AC, Harvey DJ, et al. Alzheimer's Disease Neuroimaging Initiative (ADNI): Clinical characterization. Neurology 2010;74:201–9.

[8] Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment. Arch Neurol 1999;56:303.

[9] Winblad B, Palmer K, Kivipelto M, Jelic V, Fratiglioni L, Wahlund L-O, et al. Mild cognitive impairment—beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. J Intern Med 2004;256:240–6.

[10] Jak AJ, Bondi MW, Delano-Wood L, Wierenga C, Corey-Bloom J, Salmon DP, et al. Quantification of five neuropsychological approaches to defining mild cognitive impairment. Am J Geriatr Psychiatry 2009;17:368–75.

[11] Bondi MW, Edmonds EC, Jak AJ, Clark LR, Delano-Wood L, McDonald CR, et al. Neuropsychological criteria for mild cognitive impairment improves diagnostic precision, biomarker associations, and progression rates. J Alzheimers Dis 2014;42:275–89.

[12] Edmonds EC, Delano-Wood L, Clark LR, Jak AJ, Nation DA, McDonald CR, et al. Susceptibility of the conventional criteria for mild cognitive impairment to false-positive diagnostic errors. Alzheimers Dement 2015;11:415–24.

[13] Bangen KJ, Clark AL, Werhane M, Edmonds EC, Nation DA, Evangelista N, et al. Cortical amyloid burden differences across empirically-derived mild cognitive impairment subtypes and interaction with APOE ε4 genotype. J Alzheimers Dis 2016;52:849–61.

[14] Edmonds EC, Eppig J, Bondi MW, Leyden KM, Goodwin B, Delano-Wood L, et al. Heterogeneous cortical atrophy patterns in MCI not captured by conventional diagnostic criteria. Neurology 2016;87:2108–16.

[15] Thomas KR, Edmonds EC, Delano-Wood L, Bondi MW. Longitudinal trajectories of informant-reported daily functioning in empirically defined subtypes of mild cognitive impairment. J Int Neuropsychol Soc 2017;23:521–7.

[16] Eppig JS, Edmonds EC, Campbell L, Sanderson-Cimino M, Delano-Wood L, Bondi MW, et al. Statistically derived subtypes and associations with cerebrospinal fluid and genetic biomarkers in mild cognitive impairment: A latent profile analysis. J Int Neuropsychol Soc 2017;23:564–76.

[17] 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;34:939–44.

[18] Hansson O, Seibyl J, Stomrud E, Zetterberg H, Trojanowski JQ, Bittner T, et al. CSF biomarkers of Alzheimer's disease concord with amyloid-β PET and predict clinical progression: A study of fully automated immunoassays in BioFINDER and ADNI cohorts. Alzheimers Dement 2018;14:1470–81.

[19] Ryu SY, Lee SB, Kim TW, Lee TJ. Subjective memory complaints, depressive symptoms and instrumental activities of daily living in mild cognitive impairment. Int Psychogeriatr 2016;28:487–94.

[20] Yates JA, Clare L, Woods RT, CFAS M. Subjective memory complaints, mood and MCI: A follow-up study. Aging Ment Health 2017;21:313–21.

[21] Edmonds EC, Delano-Wood L, Galasko DR, Salmon DP, Bondi MW. Subjective cognitive complaints contribute to misdiagnosis of mild cognitive impairment. J Int Neuropsychol Soc 2014;20:836–47.

[22] Aerts L, Heffernan M, Kochan NA, Crawford JD, Draper B, Trollor JN, et al. Effects of MCI subtype and reversion on progression to dementia in a community sample. Neurology 2017;88:2225–32.

[23] Lineweaver TT, Bondi MW, Galasko D, Salmon DP. Effect of knowledge of APOE genotype on subjective and objective memory performance in healthy older adults. Am J Psychiatry 2014;171:201–8.

[24] Goldberg TE, Harvey PD, Wesnes KA, Snyder PJ, Schneider LS. Practice effects due to serial cognitive assessment: implications for preclinical Alzheimer's disease randomized controlled trials. Alzheimers Dement 2015;1:103–11.

[25] Loewenstein DA, Acevedo A, Small BJ, Agron J, Crocco E, Duara R. Stability of different subtypes of mild cognitive impairment among the elderly over a 2- to 3-year follow-up period. Dement Geriatr Cogn Disord 2009;27:418–23.

[26] Palmer BW, Boone KB, Lesser IM, Wohl MA. Base rates of "impaired" neuropsychological test performance among healthy older adults. Arch Clin Neuropsychol 1998;13:503–11.

[27] Binder LM, Iverson GL, Brooks BL. To err is human: "abnormal" neuropsychological scores and variability are common in healthy adults. Arch Clin Neuropsychol 2009;24:31–46.

[28] Crawford JR, Garthwaite PH, Gault CB. Estimating the percentage of the population with abnormally low scores (or abnormally large score differences) on standardized neuropsychological test batteries: A generic method with applications. Neuropsychology 2007;21:419–30.

[29] Edmonds EC, Ard MC, Edland SD, Galasko DR, Salmon DP, Bondi MW. Unmasking the benefits of donepezil via psychometrically precise identification of mild cognitive impairment: A secondary analysis of the ADCS vitamin E and donepezil in MCI study. Alzheimers Dement 2018;4:11–8.

[30] Shaw LM, Waligorska T, Fields L, Korecka M, Figurski M, Trojanowski JQ, et al. Derivation of cutoffs for the Elecsys® amyloid β (1–42) assay in Alzheimer's disease. Alzheimers Dement Diagn Assess Dis Monit 2018;10:698–705.

[31] Blennow K, Mattsson N, Schöll M, Hansson O, Zetterberg H. Amyloid biomarkers in Alzheimer's disease. Trends Pharmacol Sci 2015;36:297–309.

[32] Parnetti L, Tiraboschi P, Lanari A, Peducci M, Padiglioni C, D'Amore C, et al. Cerebrospinal fluid biomarkers in Parkinson's disease with dementia and dementia with lewy bodies. Biol Psychiatry 2008;64:850–5.

[33] Blennow K, Dubois B, Fagan AM, Lewczuk P, de Leon MJ, Hampel H. Clinical utility of cerebrospinal fluid biomarkers in the diagnosis of early Alzheimer's disease. Alzheimers Dement 2015;11:58–69.