Prodromal Alzheimer's Disease Demonstrates Increased Errors at a Simple and Automated Anti-Saccade Task

John G. Holden[1,2,#], Alexandre Cosnard[1,#], Brice Laurens[1], Julien Asselineau[3], Damien Biotti[4], Stéphanie Cubizolle[1], Sandrine Dupouy[1], Maïté Formaglio[4], Lejla Koric[5], Magali Seassau[6], Caroline Tilikete[4], Alain Vighetto[4], François Tison[1*]

[1]Institute for Neurodegenerative Diseases, Clinical Branch, University and University Hospital of Bordeaux, Bordeaux, France
[2]Department of Psychology, University of Cincinnati, OH, USA
[3]Bordeaux University Hospital, Department of Public Health, Clinical Epidemiology Unit, Bordeaux, France
[4]Neurocognition Department, Hôpital Neurologique et Neurochirurgical Pierre Wertheimer, Lyon, France.
[5]Department of Neurology and Neuropsychology University Hospital la Timone, Marseille, France
[6]e(ye)BRAIN, Ivry-sur-Seine, France
[#]These authors contributed equally to this work.

Running title: Saccades in aMCI
Accepted 23 July 2018
Handling Associate Editor: Gilles Chopard

***Correspondence to:** Prof. François Tison, Institut des Maladies Neurodégénératives, GH Pellegrin, CHU de Bordeaux, Place Amélie Raba Léon 33076 Bordeaux Cedex, France. E-mail: françois.tison@chu-bordeaux.fr

Abstract
Saccade alterations are potential early signs of Alzheimer's disease. However, uncertainty persists in how early and reliably automated saccade recording systems detect impairments. This multicenter pathophysiological case-control transversal study explored saccade execution in carefully diagnosed amnestic mild cognitive impairment patients fulfilling research criteria for prodromal Alzheimer's disease (n = 29), as compared to both aged-matched mild Alzheimer's disease patients (n = 23) and controls (n = 27). Auto-coded saccades from horizontal (gap) vertical (step) stimulus elicited pro-saccades, and anti-saccade (gap) tasks were compared across the 3 groups. Mild cognitive impairment patients committed significantly more anti-saccade errors compared to controls (46.9 versus 24.3%, $p < 0.001$). Conventional analyses of the auto-coded stimulus elicited saccades parameters did not distinguish the amnestic mild cognitive impairment from controls or the mild Alzheimer's disease group. However, an offline analysis of manually coded saccade latencies, using resampling statistics did reveal subtle differences among the groups. Analysis of the manually coded data revealed that the mild Alzheimer's disease group had a reliably larger self-corrected error-rate than in amnestic mild cognitive impairment and controls ($p = 0.003$). Analysis of the manually coded saccade latencies, using more sensitive lognormal bootstrap analysis revealed a continuum, from amnestic mild cognitive impairment to mild Alzheimer's disease, of an increased severity of impaired inhibition of stimulus elicited saccades and correct voluntary saccade initiation. Anti-saccade error rates and psychometric measures of executive and several other cognitive functions were moderately and negatively correlated. Overall, inhibitory impairments in stimulus elicited saccades, characteristic of Alzheimer's disease, may be detected early in presumed prodromal patients using a simple, automated anti-saccade task.
Keywords: Alzheimer's disease, amnestic mild cognitive impairment, executive function, saccadic eye movements

INTRODUCTION

Saccades are rapid eye movements that redirect one's sight-line toward different environmental fixation points [1]. Saccade alterations are increasingly implicated in the diagnosis and progression of Alzheimer's disease and may offer a potential early biomarker paralleling the course of disease severity using non-invasive video-tracking, thanks to its simplicity and availability [2]. Indeed, saccades are controlled by key frontal and parietal lobes, and their respective connections that may all be functionally or anatomically impacted by neurodegeneration related to Alzheimer's disease [3-7].

The basic execution parameters of stimulus elicited saccades (velocity, amplitude/gain) are historically inconsistent in their ability to discriminate Alzheimer's disease from the natural process of aging [8-10]. However, increased latencies were more consistently reported, and display better correlations with the underlying degenerative process [11, 12]. The inconsistencies in the literature may come from measurement procedures, statistical procedures, as well as heterogeneity in Alzheimer's disease patient's severity and control characteristics.

The most consistent abnormality found in the Alzheimer's disease literature is a relative increase in the rate of direction errors for the anti-saccade task compared to controls [9, 13, 14]. In the anti-saccade task, participants are asked to look in the opposite direction of a visual target, increasing the cognitive demand particularly by involving voluntary inhibition. Anti-saccade errors correlate well with global cognitive function measures, and alterations in executive functions and/or functional or structural alterations in frontal or fronto-parietal cortex, compared to controls [14]. These group differences are associated with functional or structural alterations in frontal eye fields or cortical fronto-parietal thinning [16].

Since most published reports involve a limited number of patients with full-blown Alzheimer's disease of various severities, diagnostic uncertainty remains about how early in the disease progression saccade abnormalities can be detected. Only three recent studies explored saccade execution in a limited number of mild cognitive impairment (MCI) patients [15-17]. All suggested, in different ways, an increased number of errors in the anti-saccade task, as found for Alzheimer's disease.

The present multicenter study was designed to explore the potential clinical diagnostic utility, accuracy, and sensitivity of automated video-tracking systems for early detection of impairments due to Alzheimer's disease associated neurodegeneration using a suite of simple eye-tracking tasks: stimulus elicited horizontal (gap condition) and vertical (step condition) pro-saccades, as well as horizontal (gap condition) anti-saccades in carefully diagnosed amnestic mild cognitive impairment (aMCI) patients fulfilling research criteria for prodromal Alzheimer's disease [5], compared to aged-matched mild Alzheimer's disease patients and controls. To explore if observed results are more likely pathophysiological or technical in nature, we conducted a two-step data coding and statistical analysis techniques. The first relied on standard, fully automated saccade coding procedures, and conventional linear statistical analyses to contrast the study groups using each participant's session-averaged saccade parameters as dependent measures. These procedures are easy-to-use in the clinical setting and representative of commonly reported data treatments in the Alzheimer's disease eye-movement literature [16, 17]. Because all algorithmic or automated saccade coding suffer from shortcomings in the detection and parsing of ocular-motor events [15, 18-22] particularly in special populations such as infants and the elderly [21], the second step was designed to be maximally sensitive to group differences: automatically coded saccades were manually verified and adjusted when appropriate to increase

reliability [22-23] and a computer-intensive parametric bootstrap analyses were then used to contrast the three groups' performances [24].

If the previously noted inconsistences in prosaccade outcomes are due largely to pathophysiological heterogeneity, then the results of the standard analysis of automatically coded prosaccades and those of the bootstrapped manually coded saccades should be similar and concur with the most consistent extant literature outcome: longer latencies for Alzheimer's disease patients relative to controls. Alternatively, if technical and statistical issues need to be considered to understand outcome discrepancies, then differences in the results of the two approaches should emerge. Indeed, Alzheimer's disease patients' anatomical inhibitory impairments should translate to faster prosaccade latencies than control groups. More anticipations, but an absence of mean latency differences [17] or slower prosaccade latencies [11] were observed, and apparently, faster prosaccade latencies for Alzheimer's disease patients, relative to controls has not been reported. We hypothesize that this may be due at least in part to technical and statistical issues.

MATERIALS AND METHODS

This was a multicenter (memory clinics of the university hospitals of Bordeaux, Lyon, and Marseille, France) pathophysiological case-control transversal study involving three groups of subjects: aMCI patients fulfilling research criteria for prodromal Alzheimer's disease [25], mild Alzheimer's disease, and controls, age-matched with aMCI. The study was coordinated and promoted by the university hospital of Bordeaux, funded by the French Health Ministry (PHRC#CHUBX 2011/22), approved by the Bordeaux ethical committee (#2011/105) and registered under ClinicalTrials.gov number NCT01630525.

Participants

The study was initially designed to enroll 90 participants, 30 per study group (aMCI, Alzheimer's disease, and controls) and per investigating center. Participants were recruited while seeking medical advice for a memory complaint and subsequently diagnosed as aMCI or were diagnosed and followed as mild Alzheimer's disease according to recent research criteria [25, 26]. The neuropsychological work-up as well as diagnosis using brain magnetic resonance imaging and cerebrospinal fluid biomarkers were routinely performed for diagnosis purposes in all three centers. Control participants were aMCI and Alzheimer's disease cognitively intact family members and relatives. The study recruitment period was two years.

All participants were at least 60 years old, had clinically assessed normal or corrected to normal binocular visual acuity (Parinaud's French vision scale measuring reading capacities and equivalent to the Jaeger scale), and no clinically assessed oculomotor deficits. Participants with severe depression were excluded. Severe depression was clinically ruled-out by the investigator on the basis of history, current diagnosis, current treatment, depressive symptoms based on the Geriatric Depression Scale (GDS, a score > 10 indicates a risk of depression), and clinical appreciation.

Inclusion criteria for aMCI patients were derived from prodromal Alzheimer's disease research criteria [25, 26] and include: memory complaints, progressive onset (>6 months), normal or slight restriction at the Lawton's instrumental activity of daily living [27], an "hippocampal-type" amnestic syndrome defined by poor free recall despite adequate (and controlled) encoding, decreased total recall because of insufficient effect of cuing or impaired recognition, numerous intrusions at the Free and Cued Selective Reminding test (FCSRT),

persistence of memory changes at a subsequent assessment (>3 months), a Mini-Mental State Examination (MMSE) >24, exclusion of other disorders that may cause MCI with adequate tests, a routinely performed 1.5 Tesla diagnosis brain magnetic resonance imaging showing absent or slight medio-temporal/hippocampal atrophy (score <2) [28] and if available (not mandatory) characteristic cerebrospinal fluid $A\beta_{42}$ tau ratio.

Inclusion criteria for mild Alzheimer's disease were based on NINDS-ADRDA diagnosis criteria, revised [25] and with a MMSE >20. Controls had to have a MMSE score >24 and no memory or any other significant cognitive complaints (by self-declaration and investigator interview) or any significant abnormalities at inclusion on neuropsychology testing suggestive of an unknown cognitive deficit.

Neuropsychological tests battery

The study adopted the battery that was used routinely for the exploration of a memory complaint in the French nationwide MEMENTO cohort [29] and for the diagnosis of Alzheimer's disease by the three centers and included: MMSE, FCSRT, a visual recognition memory test (DMS-48), phonemic verbal fluency (employing the letter P), Trail Making Test-A (speed of information processing, visual attention) and B (working memory and flexibility), Digit Symbol Test (psychomotor speed, visual scanning, attention, flexibility, working memory) from the Adult Intelligence Scale (Wechsler codes), the Clinical Dementia Rating Scale, language production (oral denomination 80 items), and the digit span (forward and backwards, immediate recall, auditory attention, working memory) also from the Adult Intelligence Scale (references available in [29]). The test battery was completed by an evaluation of Katz's activities of daily living [ADL; 30] and Lawton's instrumental ADL and of depressive symptoms by the GDS [31].

Saccade recording

The study was performed using an eye movement recording system (EyeBRAIN® tracker now distributed by SURICOG company, 21 inch screen with 1920 X 1080 pixel resolution) and an off-line automated analysis of eye movements (EyeBRAIN® software *meyeParadigm* and *meyeAnalysis* version 1.18.1*)* that allowed automated recording (binocular mode, eye position sampled at 300 Hz, spatial precision of 0.5°) and analysis of different parameters of pro-saccades tasks (horizontal gap and vertical step): mean latency, mean speed, and gain; and for the anti-saccade task: mean latency (correct saccades) and percentage of errors (wrong direction).

The participants were seated in a dark room wearing a helmet mounted camera and facing a screen located 60 cm in front of their eyes with their chin and their forehead placed against a chinrest support. Each recording session started with a brief training/calibration test in which the subjects looked at nine consecutive targets covering the entire visual field. The recording sessions were 20 min and used a fixed task sequence: horizontal (gap condition) then vertical (step condition) stimulus elicited pro-saccades, then the anti-saccade task (gap condition). The task event timing was modeled after [32], except the 1200 ms stimulus presentation range for the stimulus elicited and anti-saccade tasks was delayed by 700 ms and 1.3 s, respectively. Each task presented one block of 12 trials, each trial ending when the central fixation point returned. Performance on each task was recorded twice, using a separate fixed but randomly derived trial sequence each time, for a total of 24 trials that measured saccade latency (ms), speed (°/ms), and gain.

Horizontal stimulus elicited pro-saccades (gap condition)

A green central fixation square was presented with durations ranging from 2400 to 3600 ms, then disappeared for 200 ms (black screen), then a white target was presented ± 20° horizontally (right or left) in a random manner for 1000 ms. The participants were instructed to look as precisely and as fast as possible at the horizontal target and then back to the fixation point after disappearance of the target.

Vertical stimulus elicited pro-saccades (step condition)

A green central fixation point was presented with durations ranging from 2400 to 3600 ms, immediately followed by a white target presented for 1000 ms at ± 12 degrees vertically (up or down) in a random manner. The participants were instructed to look as precisely and as fast as possible at the vertical target and then back to the fixation point after the target disappeared.

Anti-saccades (gap condition)

Each trial in the anti-saccade task began with the presentation of a green center-screen fixation point for 3000 ms to 5500 ms, then 200 ms after its extinction (dark screen), a white target randomly appeared horizontally at ± 20° laterally (left or right) from center for 1000 ms. The participants were instructed to look in the direction opposite to the lateral target as fast as possible but told they could correct direction errors.

Data analysis and statistics

A two-step data coding and statistical analysis procedure was used to investigate a hypothesis that insensitivity in data coding and statistical practices are a potential source for inconsistent saccade study outcomes in the ocular-motor literature on Alzheimer's disease.

The first analysis was modeled after established practices in the ocular-motor literature. Saccade parameters were identified automatically using a computer algorithm. Exceeding the threshold for automated saccade detection required an amplitude and speed greater 2° and 30°/s respectively. Average eye-movement parameters were computed for each participant, and statistical contrasts were conducted on the participant's averaged parameters using t-tests and regression/ANOVA (or non-parametric equivalents), thus assuming a Gaussian statistical model.
The second analysis emphasized sensitivity to legitimate group differences. Increased sensitivity was achieved by manually checking the auto-coded data to correct coding errors [22]. A lognormal statistical model was adopted for the participant's raw latency distributions. Eye-movement latency distributions are positively skewed and better approximated by a lognormal, rather than a conventional Gaussian statistical model [34, 35]. The participant's individual latencies were combined into omnibus distributions as a function of group membership, and contrasts were completed using a lognormal bootstrap procedure [24, 36-38]. This approach preserved intrinsic within- and between-group variability that is otherwise lost when computing the subject-means needed for conventional t-tests.

Automated data coding procedures

An initial data set was generated that used the default automated EyeBrain® Inc. software settings to extract the trial-level observations used for horizontal gap and vertical step, and anti-saccade tasks. The software computed a measure of latency, average speed, and gain for each coded saccade. Gain is defined as the ratio between the actual saccade amplitude and the amplitude required to capture the target. The automated analysis classified saccades as correct,

erroneous, or missing. Saccade parameters were obtained only for correct saccades. To automatically calculate the error percentages, every situation in which the subject didn't look at the target (missing saccade) or in which the subject initiated the saccade before the target appeared was excluded; only direction errors for valid saccades were used. Self-corrected anti-saccades were not automatically recognized and coded.

Manual data coding procedures

Every participant's raw left and right eye-movement time-series was visually examined for measurement artifacts. The trace that expressed the fewest measurement artifacts was selected for analysis. Then each sequence of auto-coded saccades derived from the trace was verified manually. Saccades were deleted, added, or adjusted in cases where the automated analysis inserted false saccades, omitted true saccades, or inaccurately parsed saccade boundaries. Saccades that included a single glissade (momentary pause) were coded to include it. Notably, the manual analysis only adjusted saccade onsets and offsets, where necessary. The manual latencies were generated by recomputing each trace visually verified saccade onsets, as the interval between signal onset and the beginning of directed eye-movements. The manual analysis was limited to latencies because post-fixation saccade onsets are the simplest trace properties to judge reliably by means of a visual assessment. In addition, aside from errors, latencies are the most widely reported saccade parameter in the ocular-motor literature relating to neurodegeneration. Self-corrected anti-saccades were coded manually as a subset of correct anti-saccade trials for the manual analysis. The latencies of self-corrected saccades were typically longer, as their onsets were fixed at the point the saccade trajectory switched to the correct direction. By contrast, the automated analysis coded self-corrected saccades as errors. Express and anticipated saccade counts were derived from the manual analysis. Anticipated saccades are any saccade with a latency of less than 80 ms. Express saccades are any saccade with a latency equal or greater than 80ms but less than or equal to 130 ms. Percentages within each condition were computed as the portion all saccades with latencies of 130 ms or less that fell within the range of either the anticipated or express classification.

Pre-study power analyses

The sample size was calculated according to estimates found in a literature review. For sample size calculation, the most discriminative variable between Alzheimer's disease and controls chosen was the percentage of errors at the anti-saccade task [32]. The hypotheses were an expected percentage of errors in the anti-saccade task at 60% (SD = 30%) in the aMCI group and at 25% (SD = 38%) in the control group. Under these assumptions, with a type 1 error rate of 5% and a power of 80%, at least 20 subjects per group needed to be included (Nquery® software). To maintain a high power in the comparison for the other saccade parameters, sample size was enlarged to 30 subjects per group, for a total of 90 subjects.

Conventional statistical analyses

All the variables were described using means (SD) for quantitative or frequency (percentage) for qualitative measures. The primary objective was a sequential contrast of saccade execution parameters: step one contrasted the control and aMCI groups, step two contrasted the aMCI and Alzheimer's disease groups. Between groups comparisons of the saccade parameters used a Student test or a non-parametric Wilcoxon test if the normal distribution was not verified. For qualitative variables, groups were compared by using a Chi-Square (χ^2) test of a Fischer exact

test when the χ^2 test condition was not met. The correlation between the anti-saccades error percentages and mean neuropsychological test battery was estimated with the Pearson or Spearman (if normality was violated) correlation coefficient adjusted on age and clinic site. Confidence intervals (95%) were calculated empirically by using a re-sampling method. All the conventional statistical analyses were implemented with SAS® software (SAS Institute®, Cary, US. version 9.3).

Parametric bootstrap (resampling) analyses

Bootstrapping is a computer-based data resampling technique. For the manually coded latencies, parametric bootstrapping supported statistical contrasts, using a lognormal statistical model indicated by maximum likelihood fits. Lognormal bootstrapping began by first drawing a random sample of observations from an empirical latency distribution. This sample was the same size as the empirical distribution from which it was dawn and was drawn with replacement. Next, the lognormal mean (*MU*) and standard deviation (*Sigma*) were computed and retained in a bootstrap distribution. The resampling was repeated 300 times for all the statistical tests. Formally, the resulting bootstrap distribution is a standard-error distribution for the statistic of interest, given the sampled population [39]. *Z*-tests, computed from the bootstrapped lognormal parameter distributions of each patient group, were used to conduct statistical contrasts [40]. All the statistical bootstrap analyses were completed using MATLAB® (version R2016a) software.

RESULTS

Population at study

During the inclusion period 83 subjects out of the 90 planned were enrolled and available for analysis: 29 out of 30 in the aMCI group, 1 excluded patient was suffering from non-amnestic mild cognitive impairment; 23 out of 25 in the mild Alzheimer's disease group, one was excluded due to consent withdrawal and one because of dementia severity; 27 out of 28 for controls, one was excluded due to consent withdrawal. Fourteen out of 29 aMCI patients had a cerebrospinal fluid study before inclusion (48%), all demonstrated a characteristic cerebrospinal fluid $A\beta_{42}$ tau profile (IATI- Innotest Amyloid Tau Index ≤ 0.8, in favor of an AD diagnosis, one patient had 0.8).

Demographic and clinical characteristics of enrolled subjects are displayed in Table 1. It shows no clinically significant demographic differences between groups. Nineteen of 23 mild Alzheimer's disease and four out of 29 aMCI patients took anti-cholinesterasic drugs. The mild AD patients reported significantly more neuropsychiatric prescriptions than the aMCI group $\chi^2(1), = 7.7, p = 0.021$. The control group had less frequent neuropsychiatric conditions, benzodiazepine, or antidepressant drug use.

Neuropsychologic testing results are displayed in Table 2. The aMCI group had a mean MMSE score at 26.4 ± 1.8 significantly lower than controls (28.1 ± 1.6, $r^2 = 0.20$, $p < 0.001$), but higher than mild Alzheimer's disease (23.3 ± 2.6, $r^2 = 0.57$, $p < 0.001$), but with preserved instrumental activities of daily living. All cognitive dimensions: visual naming, episodic verbal memory, working memory and executive functions, verbal fluency, language production, visual naming and visuo-spatial functions, were significantly altered in aMCI compared to controls (with the exception of processing speed as assessed by the trail making test-A). aMCI patients had similar performances with those patients with mild Alzheimer's disease on memory tests. aMCI patients displayed significantly more depressive symptoms than controls.

Saccade execution results

Stimulus elicited pro-saccades

Horizontal (gap) saccades (Fig. 1A-C): *Conventional* analyses of the auto-coded horizontal saccades failed to reveal mean latency or velocity differences among the aMCI and controls and aMCI and mild Alzheimer's disease groups. However, there was a significantly lower average gain for aMCI compared to controls (mean ratio = 0.88 ± 0.06 versus 0.92 ± 0.07, $r^2 = 0.08$, $p = 0.03$). The percentage of correct saccades was 95.8% ± 7% for aMCI patients, significantly lower than in the control group (98.8% ± 3.5%, $r^2 = 0.10$, $p = 0.02$), but not significantly different from mild Alzheimer's disease (98.5% ± 2.8%, $p = 0.12$).

The mild Alzheimer's disease patients committed significantly more anticipated saccades than the controls or aMCI groups: controls versus mild Alzheimer's disease $\chi^2(1) = 7.77$, $r^2 = 0.20$, $p = 0.005$; aMCI versus mild Alzheimer's disease $\chi^2(1) = 7.46$, $r^2 = 0.21$, $p = 0.006$. The anticipated saccade rates were controls = 14.85%, aMCI = 14.81% and mild Alzheimer's disease = 32.58% of the 271 total early (≤80 ms) manually coded saccades. There were no reliable express saccade (>80 & ≤130 ms) differences among the groups (controls = 85.56%, aMCI = 86.11%, mild Alzheimer's disease = 69.41%).

The *bootstrap* analyses of the manually coded latency distributions for correct manually coded horizontal saccades contributed 634, 521, and 621 individual latencies to the bootstrapping analyses from the aMCI, mild Alzheimer's disease, and control groups, respectively. Contrasts revealed reliable differences between the aMCI group and the other two groups (Fig. 2A-D). The mild Alzheimer's disease group was slightly faster than both the controls ($Z = 2.13$, $p = 0.03$) and the aMCI groups ($Z = 3.49$, $p < 0.001$; mild Alzheimer's disease *MU* = 196.18 ms, controls *MU* = 206.35, aMCI *MU* = 216.97 ms). Furthermore, the average *Sigma* for the aMCI group was slightly larger than either the mild Alzheimer's disease or control groups, $Z = 2.39$, $p = 0.017$ and $Z = 2.27$, $p = 0.02$, respectively (aMCI *Sigma* = 87.01 ms, mild Alzheimer's disease Sigma = 67.50 ms, and controls *Sigma* = 73.20 ms). Thus, the aMCI horizontal latencies were both slightly slower and more variable than either the controls or mild Alzheimer's disease groups.

Vertical (step) saccades (Fig. 1D-F): *Conventional* analyses indicated the number of correct auto-coded vertical step saccades was 99.5% ± 7% for aMCI patients, similar to controls (99.5% ± 1.9%, $p = 0.98$) but significantly higher than in the mild Alzheimer's disease group (98.3% ± 2.3%, $r^2 = 0.09$, $p = 0.04$). Otherwise, conventional analysis of the automatically coded data did not reveal statistically reliable differences among either the aMCI and controls or the aMCI and mild Alzheimer's disease execution parameters. There were no group differences in anticipated or express saccades.

The *bootstrap* analyses of the manually coded latency distributions for correct manually coded vertical step saccades contributed 625, 502, and 585 observations to the bootstrap analysis for the aMCI, mild Alzheimer's disease, and control groups, respectively (Fig. 2E-H). The aMCI group was significantly slower than controls, $Z = 2.68$, p = 0.007, aMCI *MU* = 288.24 ms (*Sigma* = 70.31 ms), controls *MU* = 275.86 ms (*Sigma* = 61.14 ms). Likewise, aMCI was significantly slower than the AD group, $Z = 4.35$, $p < 0.001$, mild Alzheimer's disease *MU* = 269.04 ms, (*Sigma* = 59.88 ms). Overall, the aMCI group was slower than either the controls or mild Alzheimer's disease groups, which themselves were not significantly different from each other.

Anti-saccades (gap condition)

Auto-coded anti-saccades latencies and errors (Fig. 3A): *Conventional* analyses indicated the aMCI group produced a significantly larger proportion of errors compared to controls (46.9 ±

25.9% versus 24.3 ± 19.8%, $r^2 = 0.20$, $p < 0.001$), the proportion of errors was also larger in mild Alzheimer's disease but not statistically different from aMCI (56.8 ± 27%, $p = 0.18$). Thus, the percent of anti-saccade errors in aMCI fell between control and mild Alzheimer's disease values. The latency of correct saccades was significantly longer in the aMCI group (295 ms ±121.4 ms) compared to mild Alzheimer's disease (227.4 ms ±93.8 ms, $r^2 = .09$, $p = 0.03$) but not compared to controls (261.2 ms, ±96 ms, $p = 0.26$). Notably, self-corrected latencies coded as errors were excluded from analyses of the automatically coded data.

Manually coded anti-saccade errors (Fig. 3B): A *conventional* mixed 3X2 ANOVA on the manually coded data, crossing patient group with trial block on error counts mirrored the basic controls versus aMCI and mild Alzheimer's disease patient grouping that was revealed by the automatically coded errors. There was a reliable condition effect, $F(2,74) = 5.51$, $r^2 = 0.13$, $p = 0.006$. As percentages, the mean error rates were: controls = 3.54% aMCI = 17.20%, mild Alzheimer's disease = 17.63%, respectively. Post-hoc analyses were conducted determine the basis of the reliable differences indicated by the omnibus ANOVA. Two potential accounts of the patterns of differences were compared. The first implemented a hypothesis that the two patient groups committed more errors than controls, $F(1,74) = 5.65$, $r^2 = 0.071$, $p = 0.02$. The second potential pattern of differences was a linear trend of increasing error rates tracking the increasing cognitive impairment levels built into the study (i.e., controls < aMCI < Alzheimer's disease), $F(1,74) = 4.21$, $r^2 = 0.054$, $p = 0.04$. Thus, for simple anti-saccade errors, the patient-control distinction described the groups differences slightly better than a linearly increasing trend of impairment levels. Notably, these error percentages are lower than the auto-coded rates because they exclude self-corrects.

Anti-saccade self-corrections (Fig. 3C, D)*:* A *conventional* 3X2 mixed ANOVA on the self-corrected anti-saccade counts reliably distinguished the mild Alzheimer's disease patient group from both other groups. The main effect of condition was reliable, $F(2,73) = 5.92$, $r^2 = 0.14$, $p = 0.004$, controls $M = 24\%$, aMCI $M = 30.8\%$, and mild Alzheimer's disease $M = 43.3\%$, and a reliable main effect of block $F(1, 73) = 6.42$, $r^2 = 0.08$, $p = 0.01$, block 1 $M = 36.02\%$, block 2 $M = 29.38\%$. Again, two potential accounts of the differences were compared. The first implemented a hypothesis that the two patient groups committed more self-corrects than controls, $F(1,73) = 3.59$, $r^2 = 0.047$, $p = 0.06$. The second potential pattern of differences was a linear trend of increasing self-correct rates tracking the increasing cognitive impairment levels built into the study (i.e., controls < aMCI < Alzheimer's disease), $F(1,73) = 5.95$, $r^2 = 0.075$, $p = 0.02$. By contrast to the analysis of simple anti-saccade errors, the anti-saccade self-correction rates reliably tracked the linearly increasing trend of impairment levels built into the study's design.

Anti-saccade latencies (Fig. 2I-L): *Bootstrap* analyses of the manually coded latencies (including self-corrected latencies as valid observations) revealed the mild Alzheimer's disease group was significantly slower than controls, $Z = 2.36$, $p = 0.02$ (mild Alzheimer's disease = 385.64 ms, controls = 356.52 ms). The mild Alzheimer's disease group took longer than controls to organize and execute accurate anti-saccade eye-movements. The same test, contrasting the aMCI with the control group, and the mild Alzheimer's disease with aMCI group both failed to reveal reliable differences (aMCI = 375.11 ms).

Bootstrap contrasts on the lognormal standard deviation parameter (*Sigma*), indicated that the latencies of both the mild Alzheimer's disease and aMCI groups were more variable than the control group, $Z = 2.49$, $p = 0.01$, $Z = 3.00$, $p < 0.001$, (mild Alzheimer's disease *Sigma* = 144.81 ms, controls *Sigma* = 118.10 ms, aMCI *Sigma* = 149.41 ms). The average *Sigma* for the mild

Alzheimer's disease and aMCI groups were not reliably different. There were no significant differences in anticipated and express saccades frequencies across groups.

Finally, a *bootstrap Z*-test on the manually coded error latencies indicated the aMCI group's errors were significantly slower than mild Alzheimer's disease, $Z = 3.22$, $p < 0.001$, aMCI = 358 ms, mild Alzheimer's disease = 279 ms, controls = 420 ms). A bootstrap Z-test indicated a reliable difference between the mild Alzheimer's disease and controls ($Z = 2.64$, $p = 0.008$). Mild Alzheimer's disease and aMCI contributed 97 and 121 errors, respectively. However, controls committed only 21 errors, which increased statistical uncertainty for this group. Taken at face value, the error latencies suggest a rank ordering with mild Alzheimer's disease as fastest, followed by aMCI, and then the control group as slowest.

Correlations of saccades' execution with psychometric variables

Because the error rate for the anti-saccade task was the primary study variable and the only parameter that consistently distinguished aMCI from controls it was used for exploratory correlations with the psychometric test battery results, displayed in Table 3. Inverse correlations between the proportion of errors for the anti-saccade and psychometric values were weak to moderate in strength ($r_{(S)}^2 = 0.07$ to 0.30). At the level of the entire study population, most significant inverse correlations were found with: (1) MMSE, (2) executive functions: time at the Trail Making Test-B, global score at the DSST, verbal fluency and similitudes at the Wechsler Adult Intelligence Scale; and (3) with memory (digit span, FCSRT, DMS-48). Regarding the aMCI population a significant moderate inverse correlation was found with the MMSE, and verbal fluency, but the numbers were small. There was a weak positive correlation ($r_{(S)}^2 = 0.08$, $p = 0.01$) between the GDS score and anti-saccades errors that came from the control group only ($r_{(S)}^2 = 0.17$, $p = 0.05$), no significant correlation was found in the aMCI or mild Alzheimer's disease groups taken alone.

DISCUSSION

The present multicenter study explored how easy-to-use commercial automated video-tracking systems and software can detect saccades execution abnormalities in Alzheimer's disease by comparing carefully selected aMCI patients fulfilling diagnosis criteria for prodromal Alzheimer's disease, mild Alzheimer's disease, and age-matched controls. Only three recent studies explored saccade execution in MCI (totaling 54 aMCI and 36 non-specified MCI) [15-17], all using eye movement data analysis offline [16, 17] or electro-oculography recordings [15], which are more time-consuming in a clinical setting. To challenge the accuracy of results observed and to better compare them to the current literature, as well as to further explore their nature, we also performed an offline data analysis to maximize the sensitivity to group differences by manual verification and adjustment of auto-coded saccades [22, 23] and a computer-intensive parametric bootstrap analysis [37].

Our main result is that the anti-saccade task revealed inhibitory impairments in aMCI patients that can be captured by an automated system. In our study, aMCI group committed a significantly higher proportion of errors compared to controls, but comparable to the mild Alzheimer's disease group. This result corroborates outcomes reported by Alichniewicz et al. [15] and Peltsh et al. [17] which used different aMCI diagnosis criteria and offline analysis or electro-oculography. However, it is in contrast to the Heuer et al. [16] study, performed in uncharacterized mild cognitive impairment patients, in which only the Alzheimer's disease group showed an increased error rate. As found by Alichniewicz et al. [15] and Peltsh et al. [17] in their

aMCI patients, an increased latency of voluntary correct anti-saccades was also captured by the automated system, but in contrast to their studies, it was only significantly different compared to mild Alzheimer's disease but not to controls.

By contrast the automated coding and conventional statistics performed marginally in the context of the horizontal (gap) and vertical (step) saccade paradigms. They did not reliably distinguish aMCI from controls, with the sole exception of lower average gain (horizontal gap condition) compared to controls. A lower average gain was previously reported in mild Alzheimer's disease studies [8, 10]. Lower gain/accuracy may be due to the degenerative process affecting parieto-occipital lobes and subcortical structures [39]. The mismatch between the latency distribution's intrinsic lognormal behavior, and Gaussian statistical assumptions underpinning conventional analyses likely yielded low statistical power, explaining the lack of significance found in the mild Alzheimer's disease group. A Gaussian model assumes only location (mean) changes are associated with condition differences. However, the mean and variability of a lognormal are positively correlated. The heightened variability of the lognormal tends to undermine several conventional statistical practices. Outlier censoring criteria, derived from a Gaussian model, tend to eliminate the longer latencies expected by the lognormal. The increased variability yields more heterogeneity in mean subject latencies, and a loss of power. Thus, more participants are needed to reveal reliable differences with conventional analyses. Alternatively, the lognormal bootstrap was designed to accommodate the observed performance variability, allowing for more accurate and sensitive inferential tests.

In addition to auto-coded saccades, offline manually coded pro- and anti-saccades and conventional statistical analysis demonstrated two further characteristics. Firstly, off-line manually coded saccade demonstrated significantly increased anticipatory saccade rates in the Alzheimer's disease group alone and in the horizontal gap condition alone, that was twice that found in aMCI and controls (33% versus 15%, $p < 0.006$). This outcome is consistent with a lack of inhibition of saccades in mild Alzheimer's disease that is sensitized by the gap paradigm [39, 40]. Secondly, conventional post-hoc statistical analysis of manually coded error rates at anti-saccades revealed that the mild Alzheimer's disease group had a reliably larger self-corrected error-rate than aMCI and controls, but self-correction was not significantly impaired in the aMCI group compared to controls. Increased self-correction rates were demonstrated in previous Alzheimer's disease studies and are somewhat correlated with the MMSE and working memory scores [3, 6, 7, 11, 42].

If a more sensitive lognormal bootstrap analysis is applied on the manually coded pro-saccades data it reveals faster performances (smaller latencies) in the mild Alzheimer's disease group. In addition, with an increased rate of express and anticipatory saccades, this result is consistent with more impairment in tonic inhibitory function in mild Alzheimer's disease than in individuals with aMCI or controls. This result is in contrast to some reports of slower latencies found in the literature. The increased skew intrinsic to the lognormal latency distributions offers a potential explanation for the discrepancy. We speculate that, in addition to some faster latencies, patient-condition subjects may produce a portion of longer latencies, slowing conventionally computed condition means more than the lognormal *MU*. At the anti-saccade task, the bootstrap analyses of correct saccade latencies, which also included self-corrected latencies as valid observations, revealed that the mild Alzheimer's disease group took longer than controls to organize and execute accurate anti-saccade eye-movements. Regarding the errors' latency distributions, the mild Alzheimer's disease group was the fastest, followed by aMCI patients, and the control subjects the slowest.

Overall, automated coding and conventional statistics captures lack of inhibition in the aMCI and mild Alzheimer's disease group at the anti-saccade task, but only offline manual coding and more sensitive lognormal bootstrap analyses of pro-saccades could reveal abnormalities consistent with impairment of inhibition in the mild Alzheimer's disease group. This may explain some inconsistencies found in the literature on pro-saccades execution parameters in Alzheimer's disease.

At the level of the whole study population, we found significant but weak to moderate inverse correlations in between the auto-coded anti-saccade errors and the MMSE, some test exploring executive function (time Trail Making Test-B, global score at the DSST, verbal fluency and similitudes) or memory (digit span, FCSRT, DMS-48). If the smaller sample of the aMCI group is taken alone, only significant correlations were found with the MMSE and verbal fluency. Similar patterns of correlation were reported in several studies in mild Alzheimer's disease or aMCI [12]. This suggests an association between anti-saccade inhibitory impairment, global cognitive functions, and early executive impairment in aMCI. Perhaps stronger correlations with executive function and inhibition alterations would have emerged by using the Stroop test as reported in Alichniewicz et al. [15] and Peltsh et al. [17]. Unfortunately, this test was not commonly used in the psychometric test battery of the three investigating centers.

The present study relied on aMCI patients with a reduced clinical heterogeneity, as they all fulfilled the research criteria for prodromal Alzheimer's disease and almost half displayed characteristic cerebrospinal fluid. Assuming that our population is representative of prodromal Alzheimer's disease or at high risk of converting to Alzheimer's disease [44], results were consistent with a continuum of increased severity from preclinical to mild Alzheimer's disease of an impaired inhibition then impaired correct voluntary saccade initiation. Patients with aMCI displayed increased error rates for the anti-saccade task with relatively preserved self-correction while mild Alzheimer's disease patient presented more errors in the anti-saccade task with impaired self-correction, more anticipatory stimulus elicited saccades in the horizontal gap condition and slower correct and self-corrected anti-saccades.

Saccade execution alterations appears to parallel that of cognitive deficits, particularly that of global cognitive and executive functions [15] as well as functional changes preceding anatomical disease-related changes in key frontal structures such as the dorsolateral prefrontal cortex, the frontal eye field, and the supplementary eye field, as demonstrated and discussed by Peltsh et al. [17]. We however acknowledge that while carefully selected based on a persistent "hippocampal-type" amnestic syndrome, normal brain magnetic resonance imaging or the expression of slight medio-temporal/hippocampal atrophy, and a compatible cerebrospinal fluid $A\beta_{42}$/tau profile for whom it was assessed (50%), some aMCI clinical heterogeneity, may have persisted in our population, and that the concept of prodromal or preclinical Alzheimer's disease still debated [43]. A limitation of our study may also stem from the lack of power due to the relatively low number of aMCI patients. Depression may have also influenced eye movement performances as aMCI and mild Alzheimer's disease patients tended to have more depressive symptoms, but there were no statistically significant correlations between the GDS score and the rate of errors at anti-saccades, and none of the patients included had ongoing severe depression. A possible influence of psychotropic drugs on eye movement in aMCI and mild Alzheimer's disease patients may have occurred but could not be specifically investigated due to low numbers.

Conclusion

According to our results, although improvements in coding correct/corrected/incorrect and in express saccades detection are needed, the automated system could detect impairments early in presumed prodromal patients. The system used in this case is very similar to other available automated systems. Nevertheless, it remains to be determined if this conclusion may be applicable to all commercially available systems. Our study also strengthened the hypothesis that prodromal impairment of the inhibition of stimulus elicited saccades in Alzheimer's disease. Such changes are consistent with that of executive functions observed early in the disease process that may also contribute to conversion to dementia [14, 41]. Whether such saccades execution abnormalities can be used at the individual level as a diagnosis and prognosis biomarker (conversion to dementia) in aMCI remains to be further explored in large prospective cohorts.

ACKNOWLEDGMENTS

We wish to thank the patients, caregivers and control subjects and all the investigatory centers staff who made this work possible.

This work was supported by a French Gov. Ministry of Health grant (PHRC#CHUBX 2011/22)

Authors' disclosures available online (https://www.j-alz.com/manuscript-disclosures/18-0082r2).

REFERENCES

[1] Pierrot-Deseilligny C, Rivaud S, Gaymard B, Müri R, Vermersch AI (1995) Cortical control of saccades. *Ann Neurol* **37**, 557-567.

[2] Molitor RJ, Ko PC, Ally BA (2015) Eye movements in Alzheimer's disease. *J Alzheimers Dis* **44**, 1-12.

[3] Pierrot-Deseilligny C, Müri RM, Nyffeler T, Milea D (2005) The role of the human dorsolateral prefrontal cortex in ocular motor behavior. *Ann N Y Acad Sci* **1039**, 239-251.

[4] Pereira ML, Camargo MV, Aprahamian I, Forlenza OV (2014) Eye movement analysis and cognitive processing: detecting indicators of conversion to Alzheimer's disease. *Neuropsychiatr Dis Treat* **10**, 1273-1285.

[5] Dubois B, Feldman HH, Jacova C, Cummings JL, Dekosky ST, Barberger-Gateau P, Delacourte A, Frisoni G, Fox NC, Galasko D, Gauthier S, Hampel H, Jicha GA, Meguro K, O'Brien J, Pasquier F, Robert P, Rossor M, Salloway S, Sarazin M, de Souza LC, Stern Y, Visser PJ, Scheltens P (2010) Revising the definition of Alzheimer's disease: a new lexicon. *Lancet Neurol* **9**, 1118-1127.

[6] Rüb U, Del Tredici K, Schultz C, Büttner-Ennever JA, Braak H (2001) The premotor region essential for rapid vertical eye movements shows early involvement in Alzheimer's disease-related cytoskeletal pathology. *Vision Res **41***, 2149-2156.

[7] Hutton SB (2008) Cognitive control of saccadic eye movements. *Brain Cogn* **68**, 327-340.

[8] Shafiq-Antonacci R, Maruff P, Masters C, Currie J (2003) Spectrum of saccade system function in Alzheimer disease. *Arch Neurol* **60**, 1272-1278.

[9] Abel LA, Unverzagt F, Yee RD (2002) Effects of stimulus predictability and interstimulus gap on saccades in Alzheimer's disease. *Dement Geriatr Cogn Disord* **13**, 235-243.

[10] Crawford TJ, Devereaux A, Higham S, Kelly C (2015) The disengagement of visual attention in Alzheimer's disease: a longitudinal eye-tracking study. *Front Aging Neurosci* **7**, 118

[11] Boxer AL, Garbutt S, Seeley WW, Jafari A, Heuer HW, Mirsky J, Hellmuth J, Trojanowski JQ, Huang E, DeArmond S, Neuhaus J, Miller BL (2012) Saccade abnormalities in autopsy-confirmed frontotemporal lobar degeneration and Alzheimer's disease. *Arch Neurol* **69**, 509-517.

[12] Garbutt S, Matlin A, Hellmuth J, Schenk AK, Johnson JK, Rosen H, Dean D, Kramer J, Neuhaus J, Miller BL, Lisberger SG, Boxer AL (2008) Oculomotor function in frontotemporal lobar degeneration, related disorders and Alzheimer's disease. *Brain* **131 (Pt 5)**, 1268-1281.

[13] Boxer AL, Geschwind MD, Belfor N, Gorno-Tempini ML, Schauer GF, Miller BL, Weiner MW, Rosen HJ (2006) Patterns of brain atrophy that differentiate corticobasal degeneration syndrome from progressive supranuclear palsy. *Arch Neurol* **63**, 81-86.

[14] Kaufman LD, Pratt J, Levine B, Black SE (2010) Antisaccades: a probe into the dorsolateral prefrontal cortex in Alzheimer's disease. A critical review. *J Alzheimers Dis* **19**, 781-793.

[15] Alichniewicz KK, Brunner F, Klünemann HH, Greenlee MW (2013) Neural correlates of saccadic inhibition in healthy elderly and patients with amnestic mild cognitive impairment. *Front Psychol* **4**, 467

[16] Heuer HW, Mirsky JB, Kong EL, Dickerson BC, Miller BL, Kramer JH, Boxer AL (2013) Antisaccade task reflects cortical involvement in mild cognitive impairment. *Neurology* **81**, 1235-1243.

[17] Peltsch A, Hemraj A, Garcia A, Munoz DP (2014) Saccade deficits in amnestic mild cognitive impairment resemble mild Alzheimer's disease. *Eur J Neurosci* **39**, 2000-2013.

[18] Holmqvist K, Nyström, Andersson R, Dewhurst R, Jarodzka H, Van de Weijer J (2011) *Eye tracking: A comprehensive guide to methods and measures*. OUP, Oxford.

[19] Nyström M, Holmqvist K (2010) An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. *Behav Res Methods* **42**, 188-204.
[20] Shic F, Scassellati B, Chawarska K (2008) The incomplete fixation measure. *ETRA '08 Proceedings of the 2008 symposium on Eye tracking research & applications*, pp. 111-114.
[21] Wass SV, Smith TJ, Johnson M (2013) Parsing eye-tracking data of variable quality to provide accurate fixation duration estimates in infants and adults. *Behav Res Methods* **45**, 229-250.
[22] Komogortsev O V, Karpov A (2013) Automated classification and scoring of smooth pursuit eye movements in the presence of fixations and saccades. *Behav Res Methods* **45**, 203-215.
[23] Munn SM. Stefano L, Pelz JB (2008) Fixation-identification in dynamic scenes: Comparing an automated algorithm to manual coding. *Proceedings of the 5th symposium on Applied perception in graphics and visualization*, pp. 33-42.
[24] Liu Z, Holden JG, Serota RA (2016) Probability density of response times and neurophysiology of cognition. *Adv Complex Syst* **19**, 1650013.
[25] Dubois B, Feldman HH, Jacova C, Dekosky ST, Barberger-Gateau P, Cummings J, Delacourte A, Galasko D, Gauthier S, Jicha G, Meguro K, O'brien J, Pasquier F, Robert P, Rossor M, Salloway S, Stern Y, Visser PJ, Scheltens P (2007) Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *Lancet Neurol* **6**, 734-746.
[26] Dubois B, Albert ML (2004) Amnestic MCI or prodromal Alzheimer's disease? *Lancet Neurol* **3**, 246-248.
[27] Lawton MP, Brody EM (1969) Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist* **9**, 179-186.
[28] Scheltens P, Leys D, Barkhof F, Huglo D, Weinstein HC, Vermersch P, Kuiper M, Steinling M, Wolters EC, Valk J (1992) Atrophy of medial temporal lobes on MRI in "probable" Alzheimer's disease and normal ageing: diagnostic value and neuropsychological correlates. *J Neurol Neurosurg Psychiatry* **55**, 967-972.
[29] Dufouil C, Dubois B, Vellas B, Pasquier F, Blanc F, Hugon J, Hanon O, Dartigues JF, Harston S, Gabelle A, Ceccaldi M, Beauchet O, Krolak-Salmon P, David R, Rouaud O, Godefroy O, Belin C, Rouch I, Auguste N, Wallon D, Benetos A, Pariente J, Paccalin M, Moreaud O, Hommet C, Sellal F, Boutoleau-Bretonniére C, Jalenques I, Gentric A, Vandel P, Azouani C, Fillon L, Fischer C, Savarieau H, Operto G, Bertin H, Chupin M, Bouteloup V, Habert MO, Mangin JF, Chêne G; MEMENTO cohort Study Group (2017) Cognitive and imaging markers in non-demented subjects attending a memory clinic: study design and baseline findings of the MEMENTO cohort. *Alzheimers Res Ther* **9**, 67.
[30] Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW (1963) Studies of illness in the aged: The index of ADL: A standardized measure of biological and psychosocial function. *JAMA* **185**, 914-919.
[31] Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, Leirer VO (1982) Development and validation of a geriatric depression screening scale: a preliminary report. *J Psychiatr Res* **17**, 37-49.
[32] Mosimann UP, Müri RM, Burn DJ, Felblinger J, O'Brien JT, McKeith IG (2005) Saccadic eye movement changes in Parkinson's disease dementia and dementia with Lewy bodies. *Brain* **128 (Pt 6)**, 1267-1276.
[33] Credidio, HF, Teixeira EN, Reis SD, Moreira AA, Andrade Jr, JS (2012) Statistical patterns of visual search for hidden objects. *Sci Rep* **2**, 920.
[34] Stephen DG, Mirman D (2010) Interactions dominate the dynamics of visual cognition. *Cognition* **115**, 154-165.

[35] Kelty-Stephen DG, Mirman D (2013) Gaze fluctuations are not additively decomposable: Reply to Bogartz and Staub. *Cognition* **126**, 128-134.
[36] Zhou XH, Gao S, Hui SL (1997) Methods for comparing the means of two independent log-normal samples. *Biometrics* **53**, 1129-1135.
[37] Efron B, Tibshirani R (1991) Statistical data analysis in the computer age. *Science* **253**, 390-395.
[38] Mooney CZ, Duval RD (1993) *Bootstrapping: A nonparametric approach to statistical inference*. Sage Publications, Newbury Park, CA.
[39] Shakespeare TJ, Kaski D, Yong KX, Paterson RW, Slattery CF, Ryan NS, Schott JM, Crutch SJ (2015) Abnormalities of fixation, saccade and pursuit in posterior cortical atrophy. *Brain* **138 (Pt 7)**, 1976-1991
[40] Yang Q, Wang T, Su N, Xiao S, Kapoula Z (2013) Specific saccade deficits in patients with Alzheimer's disease at mild to moderate stage and in patients with amnestic mild cognitive impairment. *Age* **35**, 1287-1298.
[41] Crawford TJ, Higham S, Mayes J, Dale M, Shaunak S, Lekwuwa G (2013) The role of working memory and attentional disengagement on inhibitory control: effects of aging and Alzheimer's disease. *Age* **35**, 1637-1650.
[42] Crawford TJ, Higham S, Renvoize T, Patel J, Dale M, Suriya A, Tetley S (2005) Inhibitory control of saccadic eye movements and cognitive impairment in Alzheimer's disease. *Biol Psychiatry* **57**, 1052-1060.
[43] Epelbaum S, Genthon R, Cavedo E, Habert MO, Lamari F, Gagliardi G, Lista S, Teichmann M, Bakardjian H, Hampel H, Dubois B (2017) Preclinical Alzheimer's disease: A systematic review of the cohorts underlying the concept. *Alzheimers Dement* **13**, 454-467.

Table 1. Demographic and clinical characteristics of subjects by group, mean (SD)

Attribute	Control	aMCI	AD
N	27	29	23
Age (y)	69.5 (6.1)	71.3 (7.1)	70.6 (6.1)
Gender (m/f)	13/14	11/18	8/15
Disease duration	------	4.4 (4.2)	15.0 (15.8)
Parinaud score	4.4 (2.9)	4.4 (3.4)	3.4 (2.5)
Ophthalmologic history	1	1	0
Psychiatric history	3	10	8
Drugs			
Cholinesterase inhibitors	0	4	19
Benzodiazepines	2	5	6
Antidepressants	3	13	10

Parinaud scores measure visual acuity, and range from P28 to 1.8, where 2 equates to normal near-vision acuity. The AD patients reported significantly more neuropsychiatric prescriptions than the aMCI group $\chi^2(1), = 7.7$, $p = 0.021$; the empty cell precluded a complementary control versus aMCI contrast. The minimum p value attained contrasts of the remaining attributes was 0.2. aMCI, amnestic mild cognitive impairment; AD, Alzheimer's disease

Table 2. Neuropsychological tests results, mean (SD)

Test	Controls	aMCI	*p*	AD	*p**
MMSE	28.8 (1.6)	26.4 (1.8)	**< 0.001**	23.3 (2.6)	**< 0.001**
DMS 48 (immediate)	46.4 (1.8)	40.5 (7.0)	**0.002**	38.8 (5.6)	0.21
DMS 48 (delayed)	46.4 (2.0)	39.9 (7.8)	**< 0.001**	39.0 (5.5)	0.28
Free Recall (immediate)	28.3 (7.78)	11.4 (6.5)	**< 0.001**	10.4 (8.62)	0.27
FCSRT (immediate)	44.7 (5.2)	29.2 (8.9)	**< 0.001**	24.1 (12.1)	0.13
FCSRT (delayed)	15.3 (2.1)	10.0 (3.8)	**< 0.001**	8.2 (4.9)	0.26
Trail Making Test A (success.)	23.8 (0.6)	23.0 (4.5)	0.78	24.0 (0.2)	0.25
Trail Making Test A (time, s)	45.8 (17.2)	52.9 (19.4)	0.16	67.5 (23.4)	**0.02**
Trail Making Test B (success.)	23.2 (2.0)	22.7 (2.3)	0.11	20.4 (5.4)	0.40
Trail Making Test B (time, s)	94.5 (60.7)	133.0 (66.6)	**0.006**	191.5 (84.6)	**0.02**
Digit-span	14.6 (4.4)	14.1 (3.5)	0.93	11.9 (2.5)	**0.03**
Verbal fluency	31.9 (9.3)	23.0 (7.6)	**< 0.001**	19.3 (6.8)	0.07
Similarities	8.3 (1.8)	7.2 (2.1)	**0.04**	6.1 (2.5)	0.12
Oral Denomination 80 (total)	79.4(0.8)	78.2(4.1)	0.13	73.0 (10.2)	**0.02**
Oral Denomination 80 (time)	133.8 (32.8)	173.5 (84.4)	**0.03**	261.7(147.1	**0.006**
DSST (120 s)	55.5 (14.6)	43.6 (17.0)	0.09	35.4 (16.2)	0.24
ADL + IADL (60)	16.1 (8.8)	14.3 (2.9)	0.97	17.4 (4.8)	0.05
Clinical Dementia Rating Scale	0.0 (0.1)	0.5 (0.1)	**< 0.001**	0.9 (0.9)	**0.001**
Geriatric Depression Scale	3.9 (2.7)	7.6 (5.2)	**0.003**	5.4 (2.6)	0.14

p, statistical significance aMCI versus controls

p*, statistical significance aMCI versus mild Alzheimer

aMCI, amnestic mild cognitive impairment; AD, Alzheimer's disease; DMS 48, Delayed Matched to Sample; FCSRT, Free and Cued Selective Recall; DSST, Digit Symbol Substitution Test; ADL, Activities of Daily Living; IADL, Instrumental Activities of Daily Living. Their sum is used to assess the loss of autonomy diagnostic criteria.

Table 3. Significant correlations found between anti-saccades errors (%) and psychometric variables in all groups compared to aMCI taken alone.

Test	All	IC [95%]	p	aMCI	IC [95%]	*p*
MMSE	-0.55	[-0.7;-0.37]	**<0.0001**	-0.5	[-0.79;-0.08]	**0.001**
DMS 48 immediate	-0.35	[-0.54;-0.14]	**0.002**	-0.03	[-0.45;0.36]	0.880
DMS 48 delayed	-0.39	[-0.59;-0.17]	**0.0006**	-0.15	[-0.59;0.29]	0.46
FCSRT 16 immediate	-0.41	[-0.60;-0.20]	**0.0006**	-0.04	[-0.49;0.41]	0.84
FCSRT 16 delayed	-0.41	[-0.60;-0.18]	**0.0005**	-0.03	[-0.42;0.38]	0.91
TMT A (time)	0.37	[0.17;0.54]	**0.001**	0.37	[-0.03;0.76]	0.06
TMT B (time)	0.43	[0.20;0.61]	**0.0002**	0.32	[-0.1;0.68]	0.12
DSST	-0.45	[-0.66;-0.20]	**0.0004**	-0.40	[-0.74;0.09]	0.00
Digit-span direct	-0.27	[-0.51;-0.03]	**0.02**	-0.20	[-0.57;0.2]	0.33
Verbal fluency (categorial)	-0.42	[-0.59;-0.19]	**0.0002**	-0.42	[-0.69;-0.04]	**0.04**
Similarities	-0.31	[-0.47;-0.11]	**0.007**	-0.31	[-0.6;0.02]	0.14

aMCI, amnestic mild cognitive impairment; DMS 48, Delayed Matched to Sample; CI [95%], 95% confidence interval; FCSRT, Free and Cued Selective Recall; TMT, Trail Making Test; DSST, Digit Symbol Substitution Test

Fig. 1. Auto-coded side-gaze and step saccades. The left column of plots (A, B, C) depicts the outcomes of the auto-coded gap task. Bar graph A plots the gain results for the auto-coded gap task. The aMCI group returned significantly smaller amplitude gains than controls ($p = 0.036$). Graph B displays the proportion of valid saccades. The aMCI group returned a lower portion of valid saccades ($p = 0.022$). The average gap latencies appear on plot C, none of the group differences were statistically significant. The right column of plots (D, E, F) depicts the corresponding outcome measures for the step task, Gain, Proportion of valid saccades, and mean latencies. The AD group displayed a lower portion of valid saccades than the aMCI group ($p = 0.04$). No other statistically reliable group differences were detected. The whiskers in each plot represent one standard error of the mean. AD, Alzheimer's disease; C, Control, aMCI, prodromal Alzheimer's disease.

Fig. 2. Latency probability densities, latency models, and bootstrapped standard error distributions for the manual analysis of the side-gaze, step and anti-saccade tasks. Plots A, E, and I display Gaussian-kernel smoothed latency density functions for each group, across each task, respectively. The gap and step density functions are faster and more compact than the anti-saccade densities. Plots B, F, and J depict the probability density functions of the best-fit lognormal model for each group's latency distribution, across the three tasks. The relative ordering and spacing in the peak heights of the lognormal model modes tend to mirror the observed pattern of group differences in each task. The right column displays the bootstrapped standard error distributions of the lognormal *MU* and *Sigma* parameters that formed the basis of each Z-test described in the text (C, D, G, H, K, L). A lognormal is equivalent to a Gaussian on a logarithmic scale, so the *MU* and *Sigma* latency parameters are depicted in natural log units. The slower, more variable aMCI and AD latencies are largely consistent with previous anti-saccade reports. The more sensitive bootstrapping techniques also revealed a tendency for the AD group to express faster gap and step performances than C or aMCI, with variability comparable to C. AD, Alzheimer's disease; C, Control, aMCI, prodromal Alzheimer's disease.

Fig. 3. Antisaccades errors and self-corrections. The auto-coded anti saccade error percentages are presented by group in bar graph A. the significant differences between the controls (C) and aMCI group and the aMCI from mild Alzheimer's disease (AD), indicate an association of anti-saccade errors with disease progression. Graph B depicts the manually coded self-corrected anti-saccades. In this case only the AD group expressed significantly more anti-saccade self-corrects than the C and aMCI groups, taken together. Plot C displays the bootstrapped standard error distributions of the latencies for the manually coded error trials. The AD group produced reliably faster errors than the aMCI group. The control group only produced 21 errors, which explains the basis for the comparably broad standard error distribution.

Figure 1:

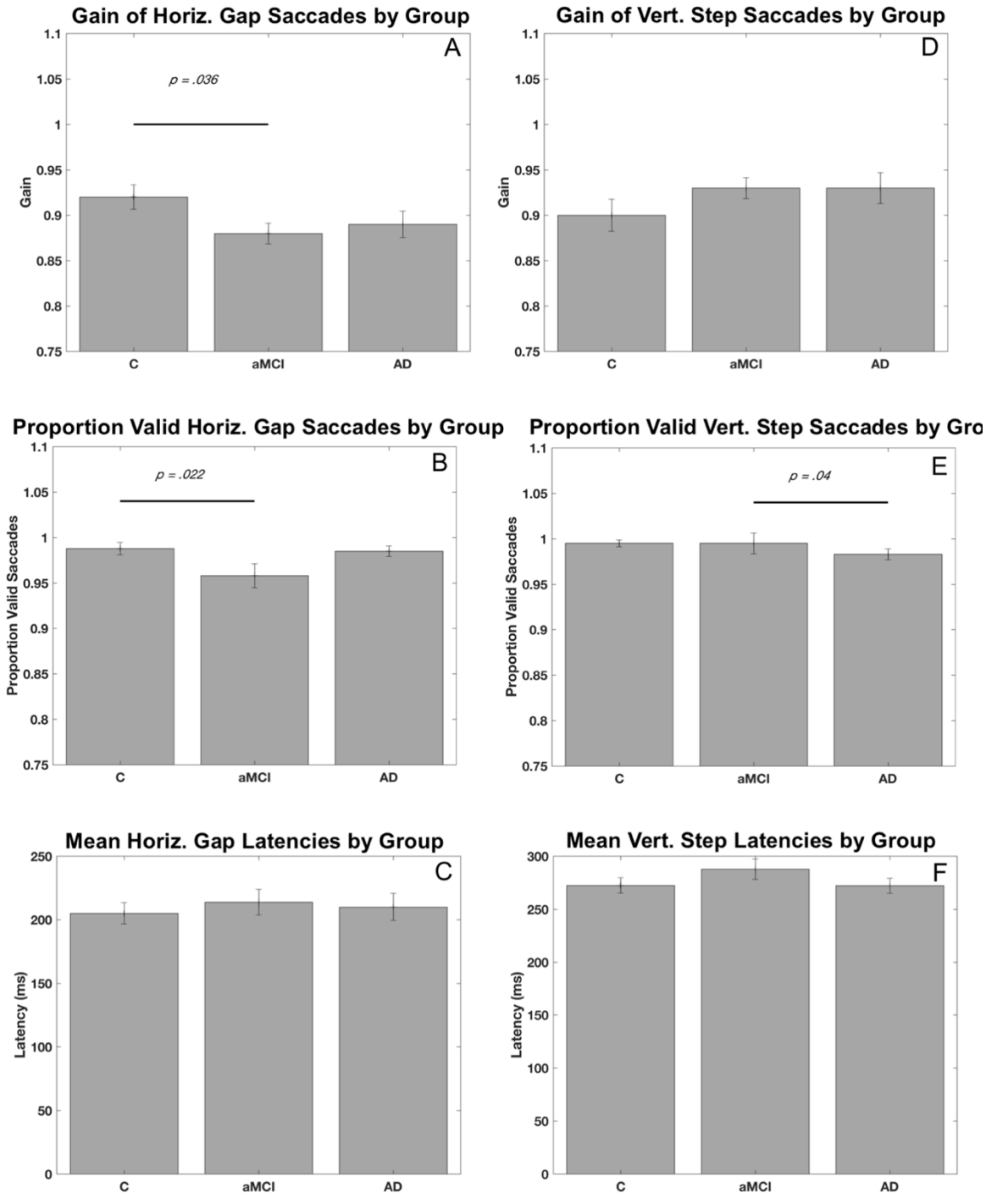

Gain of Horiz. Gap Saccades by Group
A
p = .036
Gain
1.1
1.05
1
0.95
0.9
0.85
0.8
0.75
C
aMCI
AD
Gain of Vert. Step Saccades by Group
D
Gain
Proportion Valid Horiz. Gap Saccades by Group
B
p = .022
Proportion Valid Saccades
Proportion Valid Vert. Step Saccades by Gro
E
p = .04
Proportion Valid Saccades
Mean Horiz. Gap Latencies by Group
C
Latency (ms)
250
200
150
100
50
0
Mean Vert. Step Latencies by Group
F
Latency (ms)
300

Figure 2.

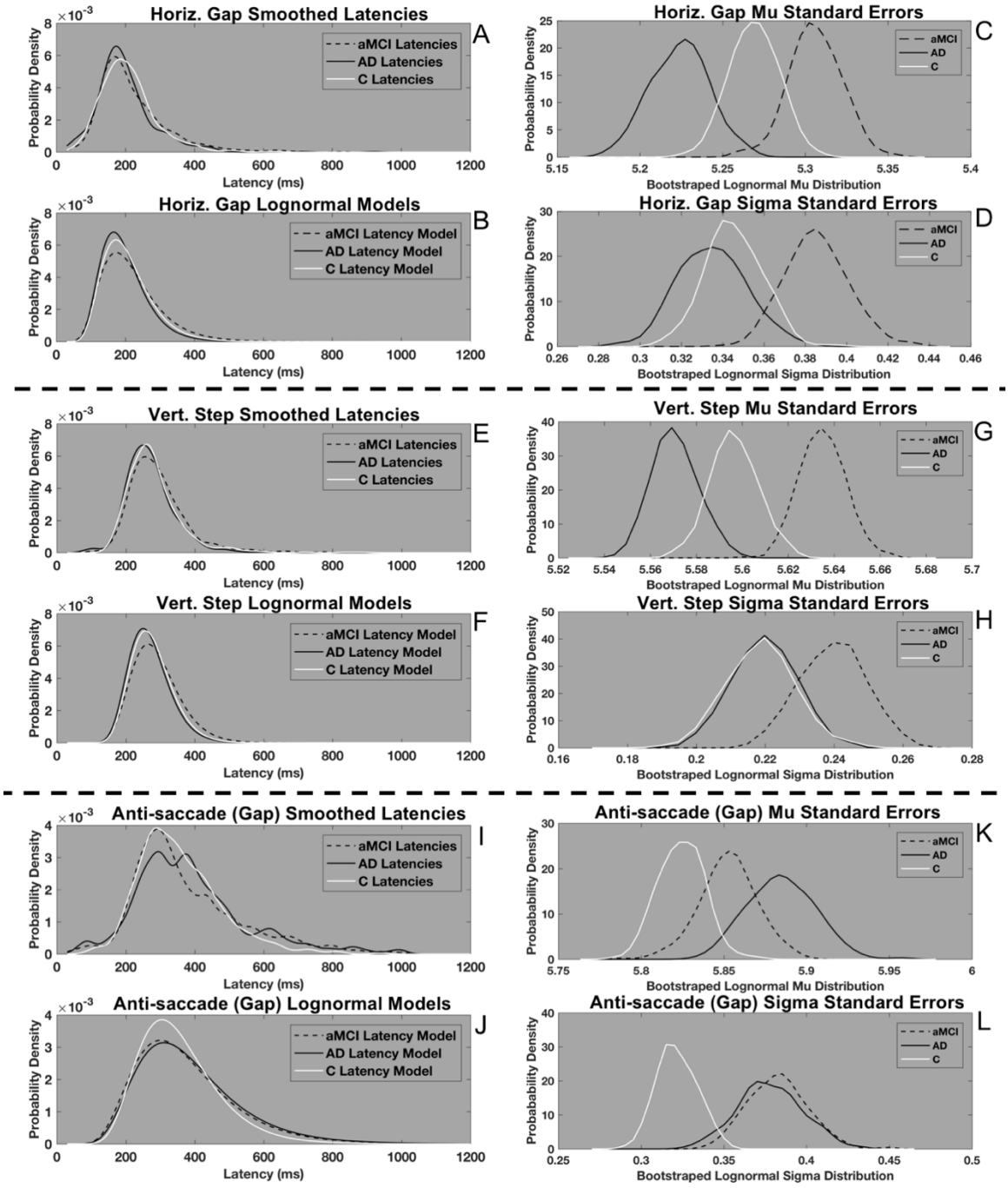
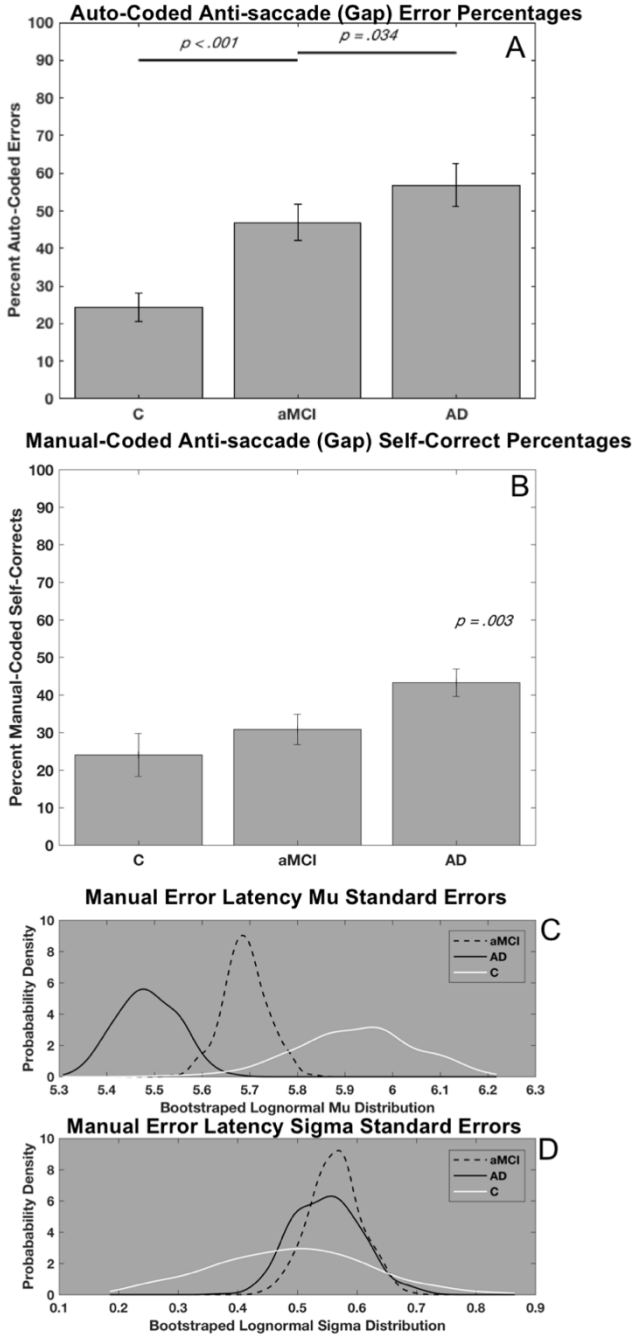

Horiz. Gap Smoothed Latencies
A
aMCI Latencies
AD Latencies
C Latencies
Probability Density
Latency (ms)
Horiz. Gap Mu Standard Errors
C
aMCI
AD
C
Probability Density
Bootstraped Lognormal Mu Distribution
Horiz. Gap Lognormal Models
B
aMCI Latency Model
AD Latency Model
C Latency Model
Probability Density
Latency (ms)
Horiz. Gap Sigma Standard Errors
D
aMCI
AD
C
Probability Density
Bootstraped Lognormal Sigma Distribution
Vert. Step Smoothed Latencies
E
aMCI Latencies
AD Latencies
C Latencies
Probability Density
Latency (ms)
Vert. Step Mu Standard Errors
G
aMCI
AD
C
Probability Density
Bootstraped Lognormal Mu Distribution
Vert. Step Lognormal Models
F
aMCI Latency Model
AD Latency Model
C Latency Model
Probability Density
Latency (ms)
Vert. Step Sigma Standard Errors
H
aMCI
AD
C
Probability Density
Bootstraped Lognormal Sigma Distribution
Anti-saccade (Gap) Smoothed Latencies
I
aMCI Latencies
AD Latencies
C Latencies
Probability Density
Latency (ms)
Anti-saccade (Gap) Mu Standard Errors
K
aMCI
AD
C
Probability Density
Bootstraped Lognormal Mu Distribution
Anti-saccade (Gap) Lognormal Models
J
aMCI Latency Model
AD Latency Model
C Latency Model
Probability Density
Latency (ms)
Anti-saccade (Gap) Sigma Standard Errors
L
aMCI
AD
C
Probability Density
Bootstraped Lognormal Sigma Distribution

Figure 3:


Auto-Coded Anti-saccade (Gap) Error Percentages
p < .001
p = .034
A
Percent Auto-Coded Errors
C
aMCI
AD
Manual-Coded Anti-saccade (Gap) Self-Correct Percentages
B
p = .003
Percent Manual-Coded Self-Corrects
Manual Error Latency Mu Standard Errors
C
aMCI
AD
C
Probability Density
Bootstraped Lognormal Mu Distribution
Manual Error Latency Sigma Standard Errors
D
Bootstraped Lognormal Sigma Distribution