Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

J Clin Exp Neuropsychol. Author manuscript; available in PMC 2023 January 16.

Published in final edited form as:

J Clin Exp Neuropsychol. 2021 October ; 43(8): 825–837. doi:10.1080/13803395.2021.2009447.

Sharper in the morning: Cognitive time of day effects revealed with high-frequency smartphone testing

Hannah Wilks[a], **Andrew J. Aschenbrenner**[a,b], **Brian A. Gordon**[b,c], **David A. Balota**[b], **Anne M. Fagan**[a], **Erik Musiek**[a], **Joyce Balls-Berry**[a], **Tammie L.S. Benzinger**[c], **Carlos Cruchaga**[d], **John C. Morris**[a], **Jason Hassenstab**[a,b]

[a]Department of Neurology, Washington University in St. Louis School of Medicine, St. Louis, MO, USA

[b]Department of Psychological & Brain Sciences, Washington University in St. Louis, St. Louis, MO, USA

[c]Department of Radiology, Washington University in St. Louis School of Medicine, St. Louis, MO, USA

[d]Department of Psychiatry, Washington University in St. Louis School of Medicine, St. Louis, MO, USA

Abstract

Decades of research has established a shift from an “eveningness” preference to a “morningness” preference with increasing age. Accordingly, older adults typically have better cognition in morning hours compared to evening hours. We present the first known attempt to capture circadian fluctuations in cognition in individuals at risk for Alzheimer disease (AD) using a remotely administered smartphone assessment that samples cognition rapidly and repeatedly over several days. Older adults (N = 169, aged 61–94 years; 93% cognitively normal) completed four brief smartphone-based testing sessions per day for 7 consecutive days at quasi-random time intervals, assessing associate memory, processing speed, and visual working memory. Scores completed during early hours were averaged for comparison with averaged scores completed during later hours. Mixed effects models evaluated time of day effects on cognition. Additional models included clinical status and cerebrospinal fluid (CSF) biomarkers for beta amyloid (Aβ42) and phosphorylated tau181 (pTau). Models with terms for age, gender, education, *APOE* ε4 status, and clinical status revealed significantly worse performance on associate memory in evening hours compared to morning hours. Contemporaneously reported mood and fatigue levels did not moderate relationships. Using CSF data to classify individuals with and without significant AD pathology, there were no group differences in performance in morning hours, but subtle impairment emerged in associate memory in evening hours in those with CSF-confirmed AD pathology. These findings indicate that memory is worse in evening hours in older adults, that

CONTACT Jason Hassenstab hassenstabj@wustl.edu Knight Alzheimer Disease Research Center, 4488 Forest Park Ave, Suite 301, St. Louis, MO 63108, USA.

Disclosure statement

Neither Dr. Morris nor his family owns stock or has equity interest (outside of mutual funds or other externally directed accounts) in any pharmaceutical or biotechnology company. All authors report no relevant conflicts of interest.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

this pattern is consistent across several days, and is independent of measures of mood and fatigue. Further, they provide preliminary evidence of a “cognitive sundowning” in the very earliest stages of AD. Time of day may be an important consideration for assessments in observational studies and clinical trials in AD populations.

Keywords

Aging; Alzheimer’s disease; smartphone-based cognitive testing; neuropsychological tests; circadian rhythms

Introduction

Circadian rhythms are physiological processes that regulate a number of critical functions including sleep–wake cycles, heart rate, and transcription of genes involved in metabolism, inflammation, and synaptic function (Musiek et al., 2018; Noya et al., 2019; Volicer et al., 2001). Age-related changes in circadian rhythm have been the focus of several comprehensive studies in normal aging and in Alzheimer’s disease (AD) populations (Ju et al., 2013; Musiek et al., 2015). Over the lifespan, circadian rhythms shift from an eveningness preference in young adults to a strong morningness preference in older adults, and this shift in preference is reflected in objective measurements of cognition captured in morning and evening hours (Anderson et al., 2014; Yoon et al., 1999). In younger adults, cognition is mostly invariant across peak and off-peak times, suggesting that time of day effects on cognition are minimal in younger adults. In contrast, older adults reliably demonstrate declines during off-peak evening hours on tests of inhibitory control, working memory, and recognition memory (Schmidt et al., 2007; Yoon et al., 1999). In AD studies, circadian disruption is exacerbated well beyond that observed in normal aging, and several studies in rodents and humans found that the primary circadian deficit seen in AD is fragmentation of activity rhythms, which is associated with an increased risk of cognitive decline and disease progression (Li et al., 2020). There is increasing evidence that circadian rhythm dysfunction occurs in very early symptomatic stages of AD (Naismith et al., 2013; Ortiz-Tudela et al., 2014) and even well before the onset of symptoms in a stage often referred to as “preclinical” AD (Musiek et al., 2018).

A related phenomenon that affects a substantial percentage of individuals with AD is known as sundowning. Sundowning symptoms can include increasing agitation, depressed mood, and cognitive difficulties as evening approaches (Little et al., 1995). Sundowning is rather loosely described in the literature, as it is difficult to study effectively and typically occurs in moderate-to-severe stages of dementias (Volicer et al., 2001). Sundowning has been hypothesized as a form of circadian dysfunction and has been tied to alterations in daily body temperature rhythms (Volicer et al., 2001). While sundowning is not considered a common phenomenon in earlier stages of AD, clinicians and caregivers of individuals with milder stages of AD commonly report mild behavioral disturbances and cognitive changes that can take the form of confusion, poor attention, and poor memory (Canevelli et al., 2016). Whether these behavioral disruptions and cognitive changes are worse in evening hours in early stages of AD is unclear. As of this writing, no studies have examined whether

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

individuals with early-stage and preclinical AD exhibit time of day effects over and above those observed in normal aging.

Measurement of time of day effects on cognition is reliant on repeated assessments, which is challenging using traditional assessment methodologies. Existing studies have used an in-person test administration repeated in morning and evening hours (e.g., Schmidt et al., 2007; Veneman et al., 2013). Such “one-shot” assessment designs suffer from poor between-subject reliability which ultimately results in low statistical power (Sliwinski, 2008). Administering multiple assessments with older adults in a memory clinic setting also presents significant logistical difficulties and is burdensome for participants and their study partners. Our approach was to use a novel smartphone application that employs a measurement burst protocol to remotely administer very brief cognitive assessments up to four times per day for 1 week in a well characterized sample of older adults enrolled in ongoing studies of normal aging and dementia. This approach is based on principles from ecological momentary assessment (EMA; for a review see, Shiffman et al., 2008), which involves repeated sampling of current behaviors and experiences in near real time. In addition to cognition, in-the-moment mood, fatigue, and other contextual information can be evaluated that provide important information that may moderate cognitive performance. EMA studies have rigorously demonstrated time of day effects on mood (Miller et al., 2015), interactions between sleep and anxiety (Cox et al., 2018), and impacts of insomnia on cognition and mood (Buysse et al., 2007).

We hypothesized that participants would perform worse on cognitive tests in evening hours compared to earlier hours, similar to existing studies of time of day and cognition. We further hypothesized that these effects would be more pronounced in participants with biomarker evidence of AD compared to those without abnormal biomarker levels. We also hypothesized that after removing symptomatic participants from analyses, time of day effects on cognition would remain present in cognitively normal participants with abnormal biomarker levels – suggesting a pattern of cognitive sundowning in preclinical AD.

Materials and methods

Participants

A total of 169 (94 male, 75 female) participants were recruited from ongoing studies of aging and dementia at the Charles F. and Joanne Knight Alzheimer Disease Research Center (Knight ADRC) at Washington University School of Medicine in Saint Louis. Due to concerns about independently operating a smartphone, we limited enrollment in the Ambulatory Research in Cognition (ARC) smartphone study to participants with a Clinical Dementia Rating™ (CDR™; Morris, 1993) score of 0 to 0.5, a range from normal cognition to very mild dementia. A CDR 0.5 score indicates that the symptomatic stage is roughly analogous to that encompassed by Mild Cognitive Impairment (Gross et al., 2017). In-person enrollment began in February 2020 and was halted in March 2020 due to SARS-CoV-2 (COVID-19) pandemic lockdowns. Because the Knight ADRC was unable to conduct in-person assessments during this stage of the pandemic, the vast majority of remaining participants were enrolled remotely beginning in April 2020. All participants

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

provided informed consent, and all procedures were approved by the Human Research Protections Office at Washington University in St. Louis.

Clinical assessments

Each participant underwent annual clinical assessments within approximately 1–2 months of smartphone-based cognitive testing (the mean interval between most recent clinical evaluation and smartphone testing was 34 ± 56 days). The presence and severity of dementia were determined by experienced clinicians using the CDR. The CDR is a semi-structured clinical interview that stages cognitive and functional deficits across six domains including memory, orientation, judgment & problems solving, community affairs, home & hobbies, and personal care. The global CDR stages are 0, indicating normal cognition, and 0.5, 1, 2, and 3, indicating very mild, mild, moderate, and severe dementia, respectively (Morris, 1993). Because of the ongoing COVID-19 pandemic, the CDR was administered by trained clinicians via phone or videoconferencing. Blood samples were drawn to determine apolipoprotein (*APOE*) genotyping, and participants were classified as *APOE* ε4 + (44, 34, 24) or *APOE* ε4− (33, 23, 22) (Cruchaga et al., 2013).

Smartphone assessments

The Ambulatory Research in Cognition (ARC) smartphone application administers extremely brief cognitive tests up to four times per day for 7 consecutive days. ARC is programmed to run on iOS and Android devices running recent major operating system (OS) versions (iOS 12.0+ and Android OS 8.0+). Participants were encouraged to use their own personal smartphones provided they met minimum technical requirements. Exclusion criteria for smartphones were software issues, limited phone storage, physical damage such as cracks to the screen or body of the phone, battery problems, or sluggish responsivity. Interested participants who did not have a smartphone or those with smartphones which did not meet our criteria were supplied a study smartphone (either iOS or Android) to keep for the duration of the study. All participants received detailed instructions on operation of the ARC application by a trained study coordinator. Participants unfamiliar with smartphones received additional guidance on the basics of setup and operation of the device. The study coordinator provided support for participants via phone, videoconferencing, e-mail, and text messaging throughout the study.

ARC assessments were administered at quasi-random intervals throughout the participant's awake hours, separated by at least 2 hours between each testing session. For example, a participant who selected 7am as their wake time and 10pm as their sleep time would receive four notifications to complete a test at nearly any time between 7am and 10pm, and each notification would be separated by at least 2 hours (see Figure 1 for a detailed example). During each testing session, mood was assessed with a simple visual analog scale (VAS) inquiring about their current mood from very poor mood to very good mood. Fatigue was also assessed with a simple VAS inquiring about their current level of energy from very tired to very alert. Three ARC cognitive measures were then administered in a random order during each assessment session (see Figure 2 for examples of the cognitive measures). Scores from each ARC measure are averaged across occasions to give an overall metric of cognition at different times of day (in this case, in early and late hours) and to minimize the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

short-term impacts of daily stressors and lapses in attention that can introduce variability in traditional “one-shot” paradigms (Sliwinski et al., 2016). This approach also provides some resiliency against practice effects, in that improvements in performance that typically occur during the first two to three retests (Collie et al., 2003) are averaged with multiple sessions after practice effects have plateaued.

Prices

Prices is an associate memory task with a learning and a recognition test. In the learning condition, participants are shown 10 item-price pairs for 3 seconds each and asked to try and remember the items and the price of each item. Items were common shopping items including food and household supplies. Prices were randomly assigned 3-digit prices that contain no repeated digits and no more than two sequential digits. In the recognition test, participants were asked to choose which of two prices was shown with the item previously. Based on extensive validation studies, the two price choices were separated by at least $3.00 to avoid ceiling and floor effects (Hassenstab et al., 2018). To protect against retest effects, 40 items are chosen without replacement, such that items are never repeated within the same day. In addition, we constrained the item and price pairings such that items are never shown with the same price over the 28 repeated sessions. Each administration of the Prices test requires approximately 60 seconds, and the primary outcome is the percent errors from the 10 recognition trials. Higher scores = worse performance.

Grids

Grids is a spatial working memory task in which three common objects (a key, a smartphone, and a pen) are randomly placed in a 5 × 5 grid. Participants are asked to remember the locations of the items. A brief distraction task is then completed followed by the presentation of a blank grid in which participants are asked to tap the locations where the items were shown. There were two trials of this task in each testing session. Stimuli are placed at random locations during each testing session to protect against retest effects. Scores reflect a Euclidean distance estimate such that a higher score indicates placement farther away from the correct placement (Sliwinski et al., 2016). Each administration of the Grids test requires approximately 30–40 seconds. Higher scores = worse performance.

Symbols

The Symbols test was based on designs used in Sliwinski et al. (2016) and is a processing speed measure in which three randomly assigned pairs of abstract shapes are shown and participants are asked to determine as quickly as possible which of two abstract shape pairs matched one of the target pairs. There were 12 trials of this task during each testing session. Item pairs are randomly assigned during each testing session to protect against retest effects. This test requires approximately 20–60 seconds to complete and the primary outcome is the response time (RT) on correct trials. Higher scores = worse performance.

Cerebrospinal fluid collection and processing

Most participants (n = 130) underwent lumbar puncture to collect cerebrospinal fluid (CSF) following overnight fasting. Participants at the Knight ADRC undergo lumbar puncture

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

approximately every 3 years, however, CSF collection was postponed in March 2020 due to the ongoing pandemic, eliminating the possibility of acquiring more recent samples. Therefore, we limited the use of CSF samples to those collected within 5 years of completion of ARC testing, resulting in 113 samples collected an average of 819.05 ± 398.2 days prior. Twenty to thirty millileters of CSF was collected in a 50 mL polypropylene tube via gravity drip using an atraumatic Sprotte 22-gauge spinal needle. CSF was kept on ice and centrifuged at low speed within 2 hours of collection. CSF was then transferred to another 50 mL tube. CSF was aliquoted at 500 μL into polypropylene tubes and stored at −80°C as previously described (Fagan et al., 2006). Prior to analysis, samples were brought to room temperature per manufacturer instructions. Samples were vortexed and transferred to polystyrene cuvettes for analysis. Concentrations of Aβ40, Aβ42, total tau (tTau), and tau phosphorylated at 181 (pTau) were measured by chemiluminescent enzyme immunoassay using a fully automated platform (LUMIPULSE G1200, Fujirebio, Malvern, PA) according to manufacturer's specifications. A single lot of reagents were used for all samples. A cutoff of > 0.0649 on the ratio of pTau:Aβ42 was used to indicate CSF biomarker positivity. This cutoff score is based on unpublished internal analyses of over 750 participants comparing CSF to amyloid positron emission tomography scans that found an AUC of 0.97 in predicting amyloid positivity.

Statistical analysis

To explore the potential effects of time of day on cognition, scores for the Prices, Grids, and Symbols measures were averaged for early and late testing times. Early testing sessions were defined as occurring prior to 1:00pm and late testing sessions were defined as occurring after 12:59pm. Early and late testing times were chosen based on a median split of the distribution of completed testing session times (early = 1,664 testing sessions; later = 1,908 testing sessions). Student's *t*-tests and Pearson correlations were used to describe relationships between early and late testing sessions, test performance and CSF biomarkers, and to describe group differences between CSF "Negative" and CSF "Positive" participants. Categorical variable comparisons were made with Chi-square tests, where appropriate. Linear mixed effects models were conducted using the lme4 package in R (Bates et al., 2015) with degrees of freedom calculated with the Satterthwaite approximation. Models included the following terms: fixed effects for age, gender, years of education, *APOE* ε4 status (carrier/non-carrier), CDR, time of day (early/late) and CSF biomarkers (positive/negative), and interval between CSF collection and ARC testing. Primary analyses focused on the main effect of time of day and the interaction between time of day and CSF biomarker status in predicting ARC test performance. All models included a random intercept term across all participants and a random slope for time of day.

Analyses proceeded in stages that began with determining if there was a main effect of time of day on ARC test performance, over and above effects of age, gender, and years of education. Additional models considered whether self-reported fatigue and mood moderated associations. Next, models considered whether disease-related variables were related to time of day effects and included terms for *APOE* ε4 status, CDR, CSF biomarker status, and interactions between CSF biomarkers and time of day. Finally, models were repeated in

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

CDR 0 (cognitively normal) participants to determine if there were interactions between AD biomarkers and time of day effects on cognition in preclinical AD.

Results

Sample characteristics are presented in Table 1. One-hundred sixty-nine participants were enrolled in the ARC smartphone study as of October 1, 2020. By design, the majority of participants were cognitively normal (90% CDR 0). Participants were highly educated and most self-reported their race as White. Participants were mostly male (66%), which differs from the total population enrolled in studies at the Knight ADRC, which is approximately 42% male. CSF Positive participants were no different than CSF Negative participants on age, education, or distribution of self-reported gender or self-reported race. CSF Positive participants had worse global CDR ($p = .001$) and were more likely to carry at least one copy of the *APOE* ε4 allele ($p < .001$). CSF Positive participants performed worse on the Symbols processing speed test ($t(41.36) = 3.28$, $p = .002$, $d = .83$) and the Grids spatial working memory test ($t(51.16) = 2.78$, $p = .008$, $d = .61$), but were no different than CSF Negative participants on the Prices associate memory test ($t(44.56) = 1.86$, $p = .07$, $d = .45$).

Adherence with ARC testing was defined as completion of all surveys and all three tests within one assessment session, divided by the total number of requested assessment sessions. Overall adherence was 75.7 ± 23.82%, meaning that, on average, participants completed 21/28 assessment sessions. These adherence rates exceed those reported in several fully remote assessment studies using smartphones (Pratap et al., 2020). As described above, early and late sessions were divided by a median split at 1:00pm. There were slightly more completed sessions in later times (Figure 3) and 3/169 participants did not complete any early sessions. Participants completed an average of 10.18 ± 5.1 (range 0–21) sessions in early hours and 11.46 ± 5.6 (range 1–26) sessions in late hours. Between subject reliability estimates were calculated using methods from Sliwinski et al. (2016). At a mean of 21 sessions completed, all measures demonstrated reliability coefficients of .80–.98, suggesting that ARC measures are highly reliable at this level of adherence.

Analyses of time of day effects (Table 2 & Figure 4) revealed that participants performed approximately 10% worse in evening hours ($M = 32.02$, $SD = 9.91$) compared to morning hours ($M = 29.55$, $SD = 9.76$) on the Prices (associate memory) test ($t(163.43) = 4.42$, $p < .001$, $d = .34$), but not on the Symbols (processing speed) test ($p = .05$) or Grids (spatial working memory) test ($p = .18$). When terms for age, gender, and years of education were added in mixed effect models, the time of day effect remained significant for the Prices test ($t(162.7) = 4.28$, $p < .001$). There was a significant main effect of gender on the Prices test, with females performing significantly better than males ($t(164.74) = 2.48$, $p = .014$), but the interaction between gender and time of day was not statistically significant, indicating that time of day did not affect female or male performance differentially.

Fatigue and mood levels collected contemporaneously (within minutes) of ARC cognitive scores were considered as potential modifiers of time of day effects on associate memory. Fatigue was weakly correlated with the Prices test in early ($r = −.16$, $p = .04$) and late ($r = −.19$, $p = .01$) periods, while mood demonstrated no significant correlations (Table 3). When

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

added to mixed effects models, neither fatigue (t(101.9) = .32, p = .75) or mood (t(102.9) = .53, p = .60) had significant associations with Prices performance and did not moderate the differences between morning and evening performance.

Considering the effects of CSF AD biomarkers on time of day and cognition, the pTau/AB42 ratio was chosen to reduce the number of comparisons and to best represent the two major proteinopathies of AD. A subset of participants completed CSF sampling, reducing the sample sizes of the comparisons to a total of 113 participants. Thirty-two participants met criteria for positivity on CSF pTau/AB42 and 81 participants were considered negative. The positive/negative classification was added to linear mixed effects models as a main effect and as an interaction term with time of day. There was no difference between AD biomarker positive and negative groups in early hours on any of the ARC measures (all p's > .20). However, in later hours there was worse performance on the Prices test in the AD biomarker positive group compared to the biomarker-negative group (t(108.6) = 2.03, p = .04) (Figure 4). When this same model was constrained to asymptomatic (CDR 0) participants, the interaction between time of day and CSF biomarker status was not significant (t(93.0) = 1.16, p = .27), suggesting that CSF indicators of AD pathology are not related to time of day effects on cognition in the preclinical stages of AD.

Due to possible differences in smartphone performance across operating systems (OS), sensitivity analyses were conducted that included a term for OS (iOS and Android). This term was not significant in any models and including this term did not alter any results or conclusions.

Discussion

In this study, we used a novel smartphone application to measure cognition rapidly and repeatedly in participants' natural environments over a 7-day period to determine if a measurement burst design was sensitive to time of day effects on cognition. We additionally sought to determine if diurnal patterns in cognition were exacerbated in those with biomarker confirmed AD. Our results, using this remote platform, support prior laboratory-based studies that found worse cognitive performance in evening hours compared to morning hours in older adults. Our results also suggest that AD pathology slightly exacerbates time of day effects in associate memory among those with abnormal levels of AD biomarkers.

Our first hypothesis, that a remote assessment with repeated testing administered over several days would demonstrate worse cognitive performance during later hours, was partially supported. Older adults performed worse during later hours on associate memory, but not spatial working memory or processing speed, than in early hours. Several studies found a similar pattern in older adults using conventional testing methods, but the domains of cognition assessed in these studies were somewhat different from our approach. For example, Yoon et al. (1999) found older adults had more difficulty inhibiting irrelevant information during the evening hours compared to the morning hours. Although the present study does not directly assess response inhibition, our results are similar in that older adults performed worse in the evening than in the morning for other closely related cognitive

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

processes. We are unaware of any studies that have found changes in processing speed related to time of day, and the lack of effects observed in the current study suggest that speed may be less impacted by diurnal rhythms. However, this was a voluntary study where participants chose their availability windows for completing assessments, so tests were typically not completed during periods closest to sleep onset when it is likely that effects on speed and other cognitive processes including spatial working memory may be observed.

We considered whether short-term changes in fatigue and mood would moderate the relationship between time of day and cognitive performance. The role of fatigue in cognition has been studied extensively in normal aging and in neurological disorders. When level of fatigue was added to our models, the results were unchanged, suggesting that subjective measures of fatigue, even when assessed repeatedly and in close temporal proximity to objective cognitive assessment, are not related to declines in performance in evening hours. There were weak correlations between subjective ratings of fatigue and associate memory in both early and late hours, but no correlations with processing speed or spatial working memory at any time of day. Previous studies using traditional cognitive assessment methods have also failed to demonstrate associations between subjective reports of fatigue or found very minor effects (Bailey et al., 2007; Niino et al., 2014).

Depression is related to fatigue and is similarly important to consider when assessing cognition in early AD (Javaherian et al., 2019). Depression is also associated with an increased eveningness preference (Drennan et al., 1991; Hasler et al., 2010), which may moderate any relationships between time of day and cognition. Similar to our fatigue question, mood was assessed with a simple VAS immediately prior to completing ARC measures. We found no relationship between mood and cognitive performance at any time of day and, when added to our models with CSF outcomes, mood did not alter the relationship between AD pathology and time of day effects on cognition. However, it is worth noting that Knight ADRC participants generally have very low levels of depressive symptoms, especially in the cognitively normal participants (Javaherian et al., 2019).

Our second hypothesis, that individuals with biomarker evidence of AD pathology would show even more deficits in later hours, revealed that AD pathology slightly exacerbated time of day effects on the Prices associate memory test, with an effect size suggesting approximately 10% worse performance in later hours compared to early hours. This trend was present when all participants (both CDR 0 and CDR .5) with available CSF were included in the models, suggesting that the period leading up to and surrounding the transition from cognitive normality to symptomatic AD may represent a critical period for cognitive sundowning. Exploring this result with more detailed analyses revealed some interesting and somewhat unexpected findings. While there was no statistically significant difference from morning to evening performance within individuals with a clinical diagnosis of very mild dementia (as indicated by a CDR of .5), there was a strong trend for individuals with CSF biomarker evidence of AD pathology to perform worse in evening hours. Most individuals with a CDR of .5 present with the clinical phenotype of AD in our sample (Morris et al., 2001), but there are cases where the etiology is uncertain, or when the clinical classification of an AD phenotype is incorrect (Beach et al., 2012). Thus, the effect observed appears to be more driven by AD pathology specifically than by clinical symptoms of

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

dementia. This pattern will require additional participants, longitudinal ARC assessments, and contemporaneous biomarker data to confirm.

Finally, we considered whether a cognitive sundowning effect would be observed among individuals who are cognitively normal but have evidence of significant AD biomarker abnormalities (preclinical AD). Although the direction of the effect remained unchanged, restricting the sample to CDR 0 individuals with biomarker evidence of AD pathology removed any trends, suggesting cognitive sundowning may manifest only after the transition to symptomatic AD.

A consideration raised by these findings is how time of day may impact cognitive assessment methods and study design in both observational and interventional studies of AD. It is standard practice in neuropsychology to conduct cognitive testing during daytime hours as both a matter of logistical practicality and to encourage optimal performance since older adults exhibit strong preferences for morning hours (Yoon et al., 1999). Detecting the earliest cognitive changes in AD can be elusive, and using traditional methodology including one-shot assessments administered at peak performance times may miss important disease-related signals that repeated measures can capture (Baker et al., 2020; Weston et al., 2018). Our results suggest that the magnitude of difference seen in evening hours with repeated testing is indicative of disease state, and perhaps this outcome, as a difference score tested over time, may be a useful indicator of disease progression or even as evidence of a treatment effect.

An overall interpretation of these data brings to light some significant limitations. Our participants, composed of highly motivated older adults who are comprehensively phenotyped and engaged in imaging and biomarker studies, are not representative of the general population. Our sample is primarily White and highly educated, which biases toward more technology access and familiarity (Pew Research, 2021). A relatively small number of cognitively impaired individuals were remotely enrolled in ARC, which poses risks to a lack of generalizability of the sample. It is possible that the cognitively impaired individuals in the Knight ADRC sample who were able to enroll in a remote, smartphone-based cognitive assessment performed better on the cognitive measures than the broader population of cognitively impaired individuals. Cerebrospinal fluid biomarkers were collected within 5 years of enrollment in ARC assessments, and it is possible some participants progressed from CSF biomarker-negative to positive since their last sample collection, which may result in misclassification of individuals as biomarker-negative. Seasonality is an important issue in the circadian rhythm literature. Differences in exposure to sunlight may be another factor that impacts the effect of time of day on cognition. Because the majority of data for this study was collected between April and September 2020, we chose not to examine the role of sunlight exposure due to limited variability in daylight hours. Finally, participants self-selected hours they wished to complete ARC assessments. Presumably, some participants may have chosen to take the tests only at their cognitive peak. Most participants completed their tests between 9am and 5pm, so there is limited data on cognition in the later evening hours when sundowning behaviors are likely to be more pronounced. Another possibility is that participants may have avoided taking tests during moments when they felt their cognitive performance was suboptimal, which could logically include later times during the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

day. This problem is common in all studies, but especially important for remote studies where there is no examiner or study proctor facilitating the assessments.

In summary, this is the first known study to examine time-of-day related fluctuations in cognitive performance in older adults at risk for AD using brief and repeatable smartphone-based cognitive tests. We found decrements in performance in associate memory in evening hours across all participants, confirming prior studies that used in-clinic measures and also extending these results by establishing that this is a consistent pattern over several consecutive days. There was subtle evidence that time of day effects on memory are exacerbated in very mildly symptomatic individuals with abnormal CSF AD biomarker levels, suggesting that a cognitive sundowning may occur in the earliest symptomatic stages of AD. The ARC smartphone study is ongoing and will include longitudinal assessments in a larger and more diverse sample with additional biomarker data. Future studies will examine if changes in time of day effects over longer time periods represent an early indicator of the transition from cognitive normality to symptomatic AD.

Acknowledgments

We would like to thank Marisol Tahan, BA, for her contributions as the primary study coordinator for the ARC smartphone study.

Funding

This work was supported by the National Institute on Aging under Grants P01AG03991, P01 AG026276, and P30AG066444.

References

Anderson JAE, Campbell KL, Amer T, Grady CL, & Hasher L (2014). Timing is everything: Age differences in the cognitive control network are modulated by time of day. Psychology and Aging, 29(3), 648–657. 10.1037/a0037243 [PubMed: 24999661]

Bailey A, Channon S, & Beaumont JG (2007). The relationship between subjective fatigue and cognitive fatigue in advanced multiple sclerosis. Multiple Sclerosis Journal, 13(1), 73–80. 10.1177/1352458506071162 [PubMed: 17294614]

Baker JE, Bruns L Jr, Jr, Hassenstab J, Masters CL, Maruff P, & Lim YY (2020). Use of an experimental language acquisition paradigm for standardized neuropsychological assessment of learning: A pilot study in young and older adults. Journal of Clinical and Experimental Neuropsychology, 42(1), 55–65. 10.1080/13803395.2019.1665626 [PubMed: 31514573]

Bates D, Maechler M, Bolker B, & Walker S (2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software, 67(1), 1–48. 10.18637/jss.v067.i01

Beach TG, Monsell SE, Phillips LE, & Kukull W (2012). Accuracy of the clinical diagnosis of Alzheimer disease at national institute on aging Alzheimer disease centers, 2005–2010. Journal of Neuropathology and Experimental Neurology, 71(4), 266–273. 10.1097/NEN.0b013e31824b211b [PubMed: 22437338]

Buysse DJ, Thompson W, Scott J, Franzen PL, Germain A, Hall M, Moul DE, Nofzinger EA, & Kupfer DJ (2007). Daytime symptoms in primary insomnia: A prospective analysis using ecological momentary assessment. Sleep Medicine, 8(3), 198–208. 10.1016/j.sleep.2006.10.006 [PubMed: 17368098]

Canevelli M, Valletta M, Trebbastoni A, Sarli G, D'Antonio F, Tariciotti L, de Lena C, & Bruno G (2016). Sundowning in dementia: Clinical relevance, pathophysiological determinants, and therapeutic approaches. Frontiers in Medicine, 3, 1–7. 10.3389/fmed.2016.00073 [PubMed: 26870733]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Collie A, Maruff P, Darby DG, & McStephen M (2003). The effects of practice on the cognitive test performance of neurologically normal individuals assessed at brief test–retest intervals. Journal of the International Neuropsychological Society, 9(3), 419–428. 10.1017/S1355617703930074 [PubMed: 12666766]

Cox RC, Sterba SK, Cole DA, Upender RP, & Olatunji BO (2018). Time of day effects on the relationship between daily sleep and anxiety: An ecological momentary assessment approach. Behaviour Research and Therapy, 111, 44–51. 10.1016/j.brat.2018.09.008 [PubMed: 30300779]

Cruchaga C, Kauwe JSK, Harari O, Jin SC, Cai Y, Karch CM, Benitez BA, Jeng AT, Skorupa T, Carrell D, Bertelsen S, Bailey M, McKean D, Shulman JM, De Jager PL, Chibnik L, Bennett DA, Arnold SE, Harold D, Sims R, … Goate AM (2013). GWAS of cerebrospinal fluid tau levels identifies risk variants for Alzheimer's disease. Neuron, 78(2), 256–268. 10.1016/j.neuron.2013.02.026 [PubMed: 23562540]

Drennan MD, Klauber MR, Kripke DF, & Goyette LM (1991). The effects of depression and age on the Horne-Ostberg morningness-eveningness score. Journal of Affective Disorders, 23(2), 93–98. 10.1016/0165-0327(91)90096-B [PubMed: 1753041]

Fagan AM, Mintun MA, Mach RH, Lee S-Y, Dence CS, Shah AR, LaRossa GN, Spinner ML, Klunk WE, Mathis CA, DeKosky ST, Morris JC, & Holtzman DM (2006). Inverse relation between in vivo amyloid imaging load and cerebrospinal fluid Aβ 42in humans. Annals of Neurology, 59(3), 512–519. 10.1002/ana.20730 [PubMed: 16372280]

Gross AL, Hassenstab JJ, Johnson SC, Clark LR, Resnick SM, Kitner-Triolo M, Masters CL, Maruff P, Morris JC, Soldan A, Pettigrew C, & Albert MS (2017). A classification algorithm for predicting progression from normal cognition to mild cognitive impairment across five cohorts: The preclinical AD consortium. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring, 8, 147–155. 10.1016/j.dadm.2017.05.003

Hasler BP, Buysse DJ, Kupfer DJ, & Germain A (2010). Phase relationships between core body temperature, melatonin, and sleep are associated with depression severity: Further evidence for circadian misalignment in non-seasonal depression. Psychiatry Research, 178(1), 205–207. 10.1016/j.psychres.2010.04.027 [PubMed: 20471106]

Hassenstab J, Aschenbrenner AJ, Balota DA, McDade E, Lim Y, Fagan AM, Benzinger TLS, Cruchaga C, Goate A, Morris JC, & Bateman RJ (2018). Comparing smartphone-administered cognitive assessments with conventional tests and biomarkers in sporadic and dominantly inherited Alzheimer disease. Alzheimer's & Dementia, 14(7), P224–P225. 10.1016/j.jalz.2018.06.2350

Javaherian K, Newman BM, Weng H, Hassenstab J, Xiong C, Coble D, Fagan AM, Benzinger T, & Morris JC (2019). Examining the complicated relationship between depressive symptoms and cognitive impairment in preclinical Alzheimer disease. Alzheimer Disease and Associated Disorders, 33(1), 15–20. 10.1097/WAD.0000000000000284 [PubMed: 30489279]

Ju Y-ES, McLeland JS, Toedebusch CD, Xiong C, Fagan AM, Duntley SP, Morris JC, & Holtzman DM (2013). Sleep quality and preclinical Alzheimer disease. JAMA Neurology, 70(5), 587. 10.1001/jamaneurol.2013.2334 [PubMed: 23479184]

Li P, Gao L, Gaba A, Yu L, Cui L, Fan W, Lim ASP, Bennett DA, Buchman AS, & Hu K (2020). Circadian disturbances in Alzheimer's disease progression: A prospective observational cohort study of community-based older adults. The Lancet Healthy Longevity, 1(3), e96–e105. 10.1016/S2666-7568(20)30015-5 [PubMed: 34179863]

Little JT, Satlin A, Sunderland T, & Volicer L (1995). Sundown syndrome in severely demented patients with probable Alzheimer's disease. Journal of Geriatric Psychiatry and Neurology, 8(2), 103–106. 10.1177/089198879500800205 [PubMed: 7794472]

Miller MA, Rothenberger SD, Hasler BP, Donofry SD, Wong PM, Manuck SB, Kamarck TW, & Roecklein KA (2015). Chronotype predicts positive affect rhythms measured by ecological momentary assessment. Chronobiology International, 32 (3), 376–384. 10.3109/07420528.2014.983602 [PubMed: 25410882]

Morris JC, Storandt M, Miller JP, McKeel DW, Price JL, Rubin EH, & Berg L (2001). Mild cognitive impairment represents early-stage Alzheimer disease. Archives of Neurology, 58(3), 3. 10.1001/archneur.58.3.397

Morris JC (1993). The Clinical Dementia Rating (CDR): Current version and scoring rules. Neurology, 43(11), 2412–2414. 10.1212/WNL.43.11.2412-a

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Musiek ES, Bhimasani M, Zangrilli MA, Morris JC, Holtzman DM, & Ju Y-ES (2018). Circadian rest-activity pattern changes in aging and preclinical Alzheimer disease. JAMA Neurology, 75(5), 582. 10.1001/jamaneurol.2017.4719 [PubMed: 29379963]

Musiek ES, Xiong DD, & Holtzman DM (2015). Sleep, circadian rhythms, and the pathogenesis of Alzheimer disease. Experimental & Molecular Medicine, 47(3), e148–e148. 10.1038/emm.2014.121 [PubMed: 25766617]

Naismith SL, Hickie IB, Terpening Z, Rajaratnam SW, Hodges JR, Bolitho S, Rogers NL, & Lewis SJG (2013). Circadian misalignment and sleep disruption in mild cognitive impairment. Journal of Alzheimer's Disease, 38(4), 857–866. 10.3233/JAD-131217

Niino M, Mifune N, Kohriyama T, Mori M, Ohashi T, Kawachi I, Shimizu Y, Fukaura H, Nakashima I, Kusunoki S, Miyamoto K, Yoshida K, Kanda T, Nomura K, Yamamura T, Yoshii F, Kira J, Nakane S, Yokoyama K, & Kikuchi S (2014). Apathy/depression, but not subjective fatigue, is related with cognitive dysfunction in patients with multiple sclerosis. BMC Neurology, 14(1), 3. 10.1186/1471-2377-14-3 [PubMed: 24393373]

Noya SB, Colameo D, Brüning F, Spinnler A, Mircsof D, Opitz L, Mann M, Tyagarajan SK, Robles MS, & Brown SA (2019). The forebrain synaptic transcriptome is organized by clocks but its proteome is driven by sleep. Science, 366(6462), eaav2642. 10.1126/science.aav2642 [PubMed: 31601739]

NW, 1615 L. St, Suite 800Washington, & Inquiries, D. 20036USA202–419-4300 | M.−857–8562 | F.−419–4372 | M. (2021). Demographics of mobile device ownership and adoption in the United States. Pew Research Center: Internet, Science & Tech. Retrieved April 29, 2021, from https://www.pewresearch.org/internet/factsheet/mobile/

Ortiz-Tudela E, Martinez-Nicolas A, Díaz-Mardomingo C, García-Herranz S, Pereda-Pérez I, Valencia A, Peraita H, Venero C, Madrid JA, & Rol MA (2014). The characterization of biological rhythms in mild cognitive impairment. BioMed Research International, 2014, 1–7. 10.1155/2014/524971

Pratap A, Neto EC, Snyder P, Stepnowsky C, Elhadad N, Grant D, Mohebbi MH, Mooney S, Suver C, Wilbanks J, Mangravite L, Heagerty PJ, Areán P, & Omberg L (2020). Indicators of retention in remote digital health studies: A cross-study evaluation of 100,000 participants. Npj Digital Medicine, 3(1), 21. 10.1038/s41746-020-0224-8 [PubMed: 32128451]

Schmidt C, Collette F, Cajochen C, & Peigneux P (2007). A time to think: Circadian rhythms in human cognition. Cognitive Neuropsychology, 24(7), 755–789. 10.1080/02643290701754158 [PubMed: 18066734]

Shiffman S, Stone AA, & Hufford MR (2008). Ecological momentary assessment. Annual Review of Clinical Psychology, 4(1), 1–32. 10.1146/annurev.clinpsy.3.022806.091415

Sliwinski MJ, Mogle JA, Hyun J, Munoz E, Smyth JM, & Lipton RB (2016). Reliability and validity of ambulatory cognitive assessments. Assessment, 25(1), 1–17. 10.1177/1073191116643164

Sliwinski MJ (2008). Measurement-Burst Designs for Social Health Research. Social and Personality Psychology Compass, 2(1), 245–261. 10.1111/j.1751-9004.2007.00043.x

Veneman CE, Gordon-Salant S, Matthews LJ, & Dubno JR (2013). Age and measurement time-of-day effects on speech recognition in noise. Ear and Hearing, 34 (3), 288–299. 10.1097/AUD.0b013e31826d0b81 [PubMed: 23187606]

Volicer L, Harper DG, Manning BC, Goldstein R, & Satlin A (2001). Sundowning and circadian rhythms in Alzheimer's disease. American Journal of Psychiatry, 158 (5), 704–711. 10.1176/appi.ajp.158.5.704 [PubMed: 11329390]

Weston PSJ, Nicholas JM, Henley SMD, Liang Y, Macpherson K, Donnachie E, Schott JM, Rossor MN, Crutch SJ, Butler CR, Zeman AZ, & Fox NC (2018). Accelerated long-term forgetting in pre-symptomatic autosomal dominant Alzheimer's disease: A cross-sectional study. The Lancet Neurology, 17(2), 123–132. 10.1016/S1474-4422(17)30434-9 [PubMed: 29413314]

Yoon C, May CP, & Hasher L (1999). Aging, circadian arousal patterns, and cognition. In Schwarz N, Park DC, Knäuper B, & Sudman S (Eds.), Cognition, aging, and self-reports (pp. 117–117). Taylor & Francis. 10.4324/9780203345115_chapter_6

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

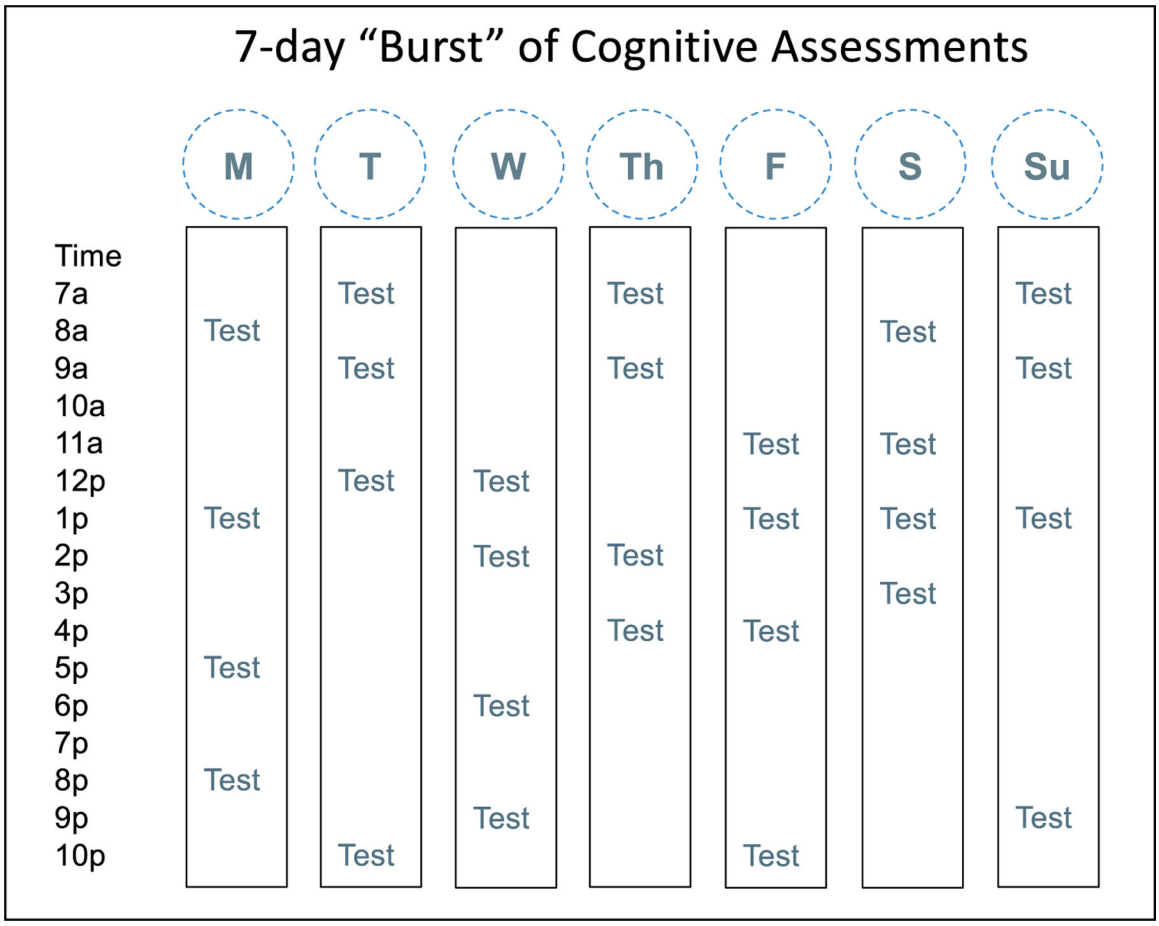

Figure 1.

Testing frequency and time of day. Example of a hypothetical participant available to complete tests between 7am and 10pm. Tests are quasi-randomly distributed throughout the available window (a minimum window of 8 hours is required). Testing sessions are separated by at least 2 hours and after a notification is sent, participants have up to 2 hours to complete the testing session.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

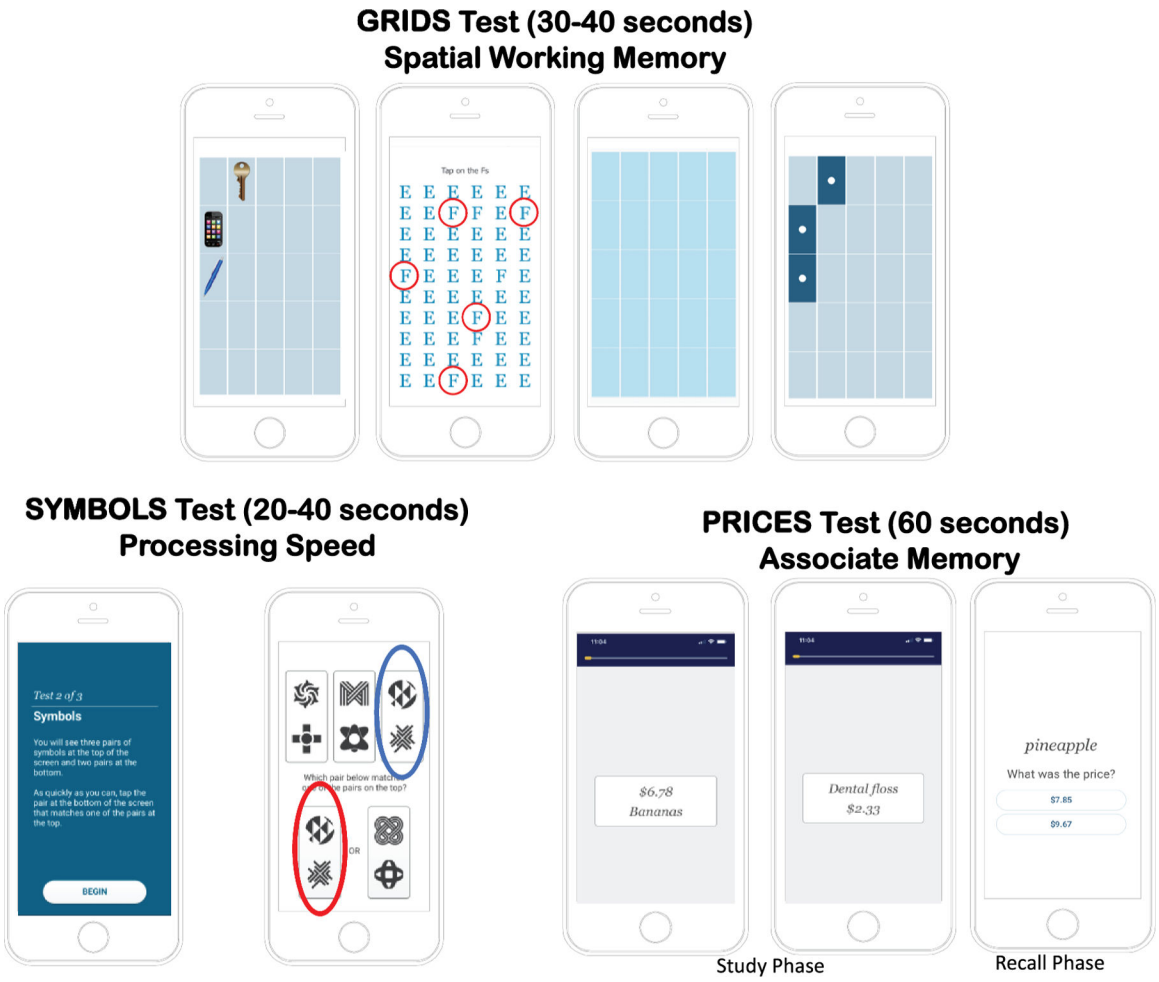

Figure 2.

Examples of ARC cognitive measures. Example of Grids, Symbols, and Prices tests. For all measures, higher scores = worse performance. The Grids test assesses spatial working memory, which requires encoding the location of three items, a brief distraction task, and a recall of the location of the three items. The Symbols test assesses the processing speed task, which requires matching pairs of symbols to a target under timed conditions. The Prices test is an associate memory task, which tests memory for common shopping items paired with randomly selected prices. The test includes a study phase and recall phase.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

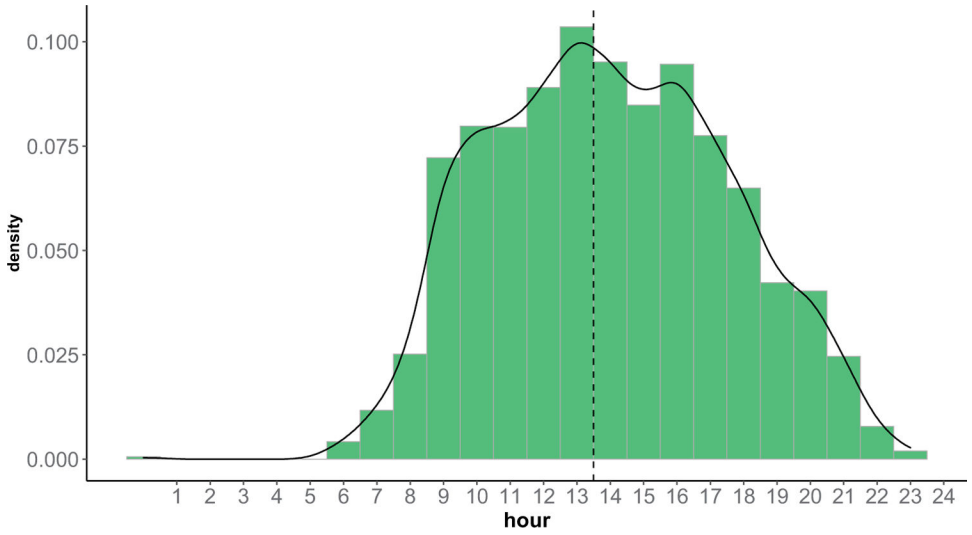

Figure 3.

ARC assessment completion times (total number of assessments N = 3,572). The dashed line represents 13:00 (1pm), the cutoff for early and late assessment times.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

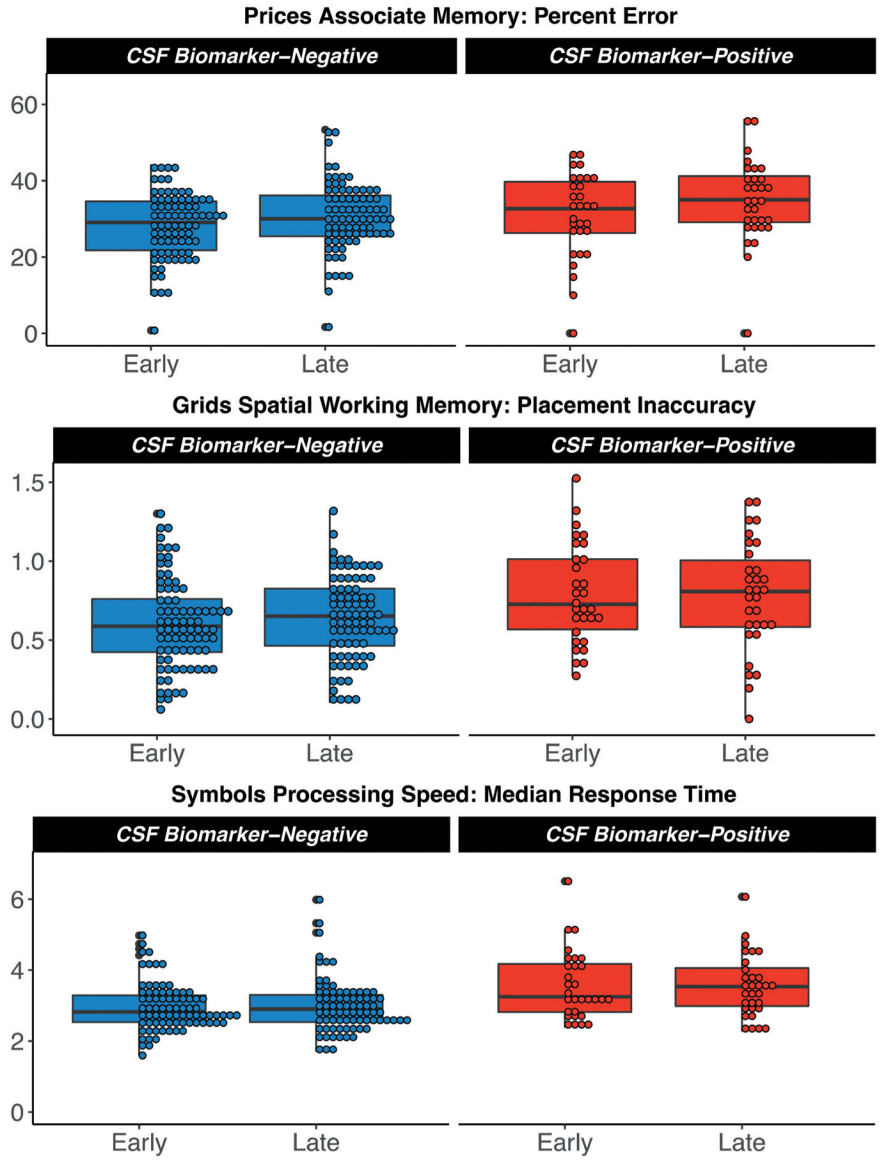

Figure 4.

CSF ptau181/AB42 ratio and time of day. Results of ARC measures in individuals with CSF data separated by CSF biomarker status and time of day. Compared to morning hours, worse performance was observed on the Prices associate memory test in the CSF biomarker positive group. There were no significant differences on the Symbols processing speed test or the Grids spatial working memory test.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 1.

Demographic, biomarker, and smartphone-based information by CSF Biomarker Status. Statistical differences calculated for continuous and categorical variables using t-tests and Chi-square tests, respectively. A p-value compares CSF negative and CSF positive groups. CSF positivity was defined as a $ptau_{181}/A\beta42$ ratio of > .0649. Effect sizes are shown as Cohen's *d*, where appropriate.

Variables	All Participants (N = 169)	CSF Negative (N = 81)	CSF Positive (N = 32)	p-value CSF Neg vs CSF Pos	Cohen's *d*
Demographics					
Age (years, SD)	75.9 (5.6)	74.6 (5.1)	76.1 (6.3)	.23	.28
Gender (n female, %)	75 (44.3)	30 (37.0)	11 (35.5)	.78	
Education (years, SD)	16.4 (2.4)	16.7 (2.3)	16.5 (2.5)	.61	.11
Race (n White, %)	138 (81.7)	73 (90.1)	30 (94.0)	.14	
Race (n Black, %)	29 (17.2)	8 (9.9)	1 (3.1)		
Race (n Asian, %)	1 (.6)	0	1 (3.1)		
*APOE ε*4 carrier (n, %)	62 (37.6)	19 (23.8)	21 (65.6)	<.001	
CDR Global (mean, SD)	.05 (.2)	.02 (.1)	.18 (.2)	.001	.97
CDR >0 (n, %)	17 (10.1%)	4 (4.9%)	11 (34.4%)	<.001	
Cerebrospinal Fluid Biomarkers					
CSF Aβ42 pg/mL (mean, SD)	–	1078.3 (413.1)	543.3 (191.7)	<.001	1.47
CSF $ptau_{181}$ pg/mL (mean, SD)	–	37.3 (13.9)	69.1 (28.7)	<.001	1.65
CSF ptau/Aβ42 ratio (mean, SD)	–	.04 (.01)	.14 (.09)	<.001	2.22
ARC Testing (Higher Scores = Worse Performance)					
Grids Euclidean Distance (mean, SD)	.7 (.3)	.6 (.3)	.8 (.3)	.008	.61
Symbols Median Response Time (mean, SD)	3.3 (.9)	2.9 (.7)	3.6 (1.1)	.002	.83
Prices % Error Score (mean, SD)	31.0 (9.0)	29.5 (7.9)	33.5 (10.8)	.07	.45

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 2.

Smartphone-based time of day effects.

Statistical differences calculated using Student's paired t-tests. A p-value compares mean test performance during early times (< 1:00pm) and later times (> 12:59pm). Higher scores = worse performance.

	Prices (Associate Memory)				Grids (Spatial Working Memory)				Symbols (Processing Speed)			
	Early (mean, SD)	Late (mean, SD)	p	Cohen's *d*	Early (mean, SD)	Late (mean, SD)	p	Cohen's *d*	Early (mean, SD)	Late (mean, SD)	p	Cohen's *d*
All (N = 169)	29.55 (9.76)	32.02 (9.91)	<.001	.34	.70 (.32)	.72 (.30)	.17	.11	3.25 (.95)	3.29 (1.01)	.05	.15
CSF Negative (N = 81)	28.22 (8.45)	30.50 (8.8)	.004	.33	.61 (.28)	.64 (.26)	.06	.27	2.96 (.66)	2.98 (.74)	.53	.07
CSF Positive (N = 32)	30.90 (11.18)	35.76 (12.31)	.001	.66	.79 (.32)	.78 (.35)	.94	.14	3.57 (.96)	3.68 (1.15)	.13	.29

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 3.

Correlations between ARC tests during early and late periods and predictor variables.

	Prices Early	Prices Late	Symbols Early	Symbols Late	Grids Early	Grids Late	Age	Education	CDR	Fatigue	Mood	CSF AB42	CSF Ptau
Prices Early	–												
Prices Late	.74***	–											
Symbols Early	.33***	.31***	–										
Symbols Late	.33***	.31***	.93***	–									
Grids Early	.22**	.26***	.35***	.37***	–								
Grids Late	.25**	.28***	.33***	.34***	.77***	–							
Age	.13	.21**	.3***	.29***	−.01	−.01	–						
Education	−.13	.00	−.07	−.06	−.06	−.11	−.14	–					
CDR	.23**	.30***	.18*	.17*	.36***	.33***	−.04	−.02	–				
Fatigue	−.16*	−.19*	−.07	−.08	.03	−.03	−.1	−.04	.01	–			
Mood	.07	.08	−.08	−.06	−.16*	−.08	−.16*	−.09	−.06	−.2*	–		
CSF AB42	−.13	−.15	−.31***	−.29***	−.17	−.11	−.07	−.01	−.33***	.13	−.02	–	
CSF Ptau	.02	.06	.22*	.2*	−.03	.11	.08	−.17	.15	.12	−.09	−.1	–
Ptau/AB42 Ratio	.08	.11	.27**	.24**	.04	.14	.05	−.16	.24**	.07	−.03	−.44***	.91***

* p < .05

** p < .01

*** p < .001.

Abbreviations: CDR = Clinical Dementia Rating; CSF = Cerebrospinal Fluid; AB42 = beta amyloid 1–42; Ptau = phosphorylated tau181.