UNIVERSITY OF LEEDS

This is a repository copy of *Reaction Time Measures Predict Incident Dementia in Community-Living Older Adults: The Sydney Memory and Ageing Study*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/94357/

Version: Accepted Version

Article:
Kochan, NA, Bunce, D, Pont, S et al. (3 more authors) (2016) Reaction Time Measures Predict Incident Dementia in Community-Living Older Adults: The Sydney Memory and Ageing Study. American Journal of Geriatric Psychiatry, 24 (3). pp. 221-231. ISSN 1064-7481

https://doi.org/10.1016/j.jagp.2015.12.005

© 2016, Elsevier. Licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International
http://creativecommons.org/licenses/by-nc-nd/4.0/

Reuse
Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

White Rose university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Reaction time measures predict incident dementia in community-living older adults: The Sydney Memory and Ageing Study.

Nicole A Kochan[1 2], PhD; David Bunce[5], PhD; Sarah Pont, BPsych(Hons)-BSc[1]; John D Crawford, PhD[1]; Henry Brodaty[1 3 4], MD, DSc; Perminder S Sachdev, MD, PhD[1 2]

Affiliations

1. CHeBA, Centre for Healthy Brain Ageing, School of Psychiatry, UNSW Medicine, The University of New South Wales, Sydney, NSW 2052, Australia
2. Neuropsychiatric Institute, Prince of Wales Hospital, Barker Street, Randwick, NSW 2031, Australia
3. Academic Department for Old Age Psychiatry, Prince of Wales Hospital, Avoca Street, Randwick, NSW 2031, Australia
4. Dementia Collaborative Research Centre, School of Psychiatry, UNSW Medicine, The University of New South Wales, Sydney, NSW 2052, Australia
5. School of Psychology, Faculty of Medicine and Health, University of Leeds, Leeds, United Kingdom

Corresponding author: Dr Nicole Kochan, CHeBA, Centre for Healthy Brain Ageing, School of Psychiatry, UNSW Medicine, The University of New South Wales, Sydney, NSW 2052. Email: n.kochan@unsw.edu.au; T: +612 93823735

Conflicts of Interest and Source of Funding: For all authors no COI declared. This study was supported by a National Health & Medical Research Council of Australia Program Grant Program Grant (ID 350833) and Capacity Building Grant (ID 568940). Further support is provided by the Dementia Collaborative Research Centre – Assessment and Better Care part of the Australian Government's Dementia Initiative. Financial support for Dr Kochan is provided by NHMRC Early Career Fellowship (ID 1053804).

dementia; variability; reaction time; neuropsychological tests; risk prediction; early diagnosis

Abstract

Objectives: To examine the utility of intra-individual variability of reaction times (IIV_{RT}) and mean reaction time (RT) as behavioural markers of incident all-cause dementia. ***Design:*** Longitudinal cohort study followed biennially for four years. ***Setting:*** The community-based Sydney Memory and Ageing Study. ***Participants:*** 861 initially non-demented participants aged 70-90. ***Measurements:*** 1) Incident all-cause dementia determined by consensus; 2) RT measures from simple and complex tasks; 3) Mini-mental State Examination and neuropsychological tests; 4) Geriatric Depression Scale, Goldberg Anxiety Scale; 5) cardiovascular risk score; 6) apolipoprotein ε4 status; 7) Bayer ADL Scale. The associations of baseline IIV_{RT} and mean RT with time to dementia were evaluated with hazard ratios (HR) using Cox proportional-hazards models with and without controlling for dementia risk factors. ***Results:*** 48 cases developed dementia. Greater Complex IIV_{RT} predicted a 40% (HR 1.43) and mean RT a 50-60% (Simple RT: HR 1.53; Complex RT: HR 1.59) per standard deviation increased risk of developing dementia, remaining significant after controlling for age, education, sex, general cognitive function, mood, cerebrovascular disease and genetic susceptibility. Prediction of incident dementia using demographical information and RT measures combined was comparable to several traditional neuropsychological measures (AUC 0.75) although lower than a full neuropsychological battery (AUC 0.90). Prediction of functional decline by RT measures combined was equal to the neuropsychological battery (multiple Rs of .233 and .238, respectively). ***Conclusions:*** Brief RT measures, can provide information on risk of imminent dementia and functional decline within four years in older adults at a population level, with mean RT the stronger predictor.

Cognitive slowing, as indicated by computer-administered reaction time (RT) measures, has been a major focus of research into cognitive ageing and is regarded as a marker of Mild Cognitive Impairment (MCI) and dementia (1-4). Less is known in this context of the predictive utility of speeded RT measures relative to the intra-individual RT variability (IIV_{RT}) obtained from the same task. IIV_{RT} represents transient within-person trial-to-trial fluctuations in RT, and is thought to arise from momentary fluctuations in attentional or executive control (5, 6). It has been purported to be an indicator of the functional integrity of brain networks (7-9). Increased IIV_{RT} is associated with frontal lobe lesions (10), alterations to white matter tracts, particularly those located frontally (11), and reduced anterior dopamine (D2 receptor) binding (12). IIV_{RT} is increased in normal ageing (7), Mild Cognitive Impairment (MCI) (13-15), Parkinson's disease (9), and mild dementia (8, 16). Taken together, behavioural and neuroimaging studies suggest that increased IIV_{RT} is sensitive to disturbances in the integrity of several neural systems. Therefore, as well as speeded RT measures such as mean RT, IIV_{RT} may be a sensitive *early* cognitive marker of all-cause dementia.

Longitudinal studies are scarce but the few available highlight the potential of mean RT and IIV_{RT} as behavioural markers of cognitive decline to mild impairment states such as MCI (17, 18) with evidence supporting IIV_{RT} as the stronger predictor. Only one small clinic-based study has evaluated RT measures in relation to incident all-cause dementia, showing that higher variability (but not median RT) at baseline differentiated patients with amnestic MCI who converted to dementia in 2.5 years from those who remained MCI (19). However, global cognitive function was not controlled for and converters were more severely impaired than nonconverters.

Hence it remains unclear whether measures of variability have better prognostic value than brief cognitive screening instruments such as the Mini Mental State Examination (MMSE)(20), complement traditional neuropsychological instruments, or have generalizability beyond specialised memory disorders clinic settings. Moreover, since IIV_{RT} can increase with RT slowing (21), the relative value of the two measures in predicting dementia is worthy of examination. Studies in healthy older persons and MCI patients have shown independence of IIV_{RT} from mean RT from the same task (e.g. 11, 22) although this is not a consistent finding (15).

Hence, our objective was to examine the unique potential of RT measures from simple and complex RT tasks to identify those older adults who were at increased risk of future dementia over four years in a large community-living cohort. First, we separately examined the predictive utility of mean RT and IIV_{RT} with and without adjustment for dementia risk factors derived from basic clinical measures of mood and cognition, specialised medical indices of vascular health and apolipoprotein ε4 status. Second, we compared the relative prognostic performance of RT measures to a broad range of traditional neuropsychological measures. Third, we examined whether mean RT and IIV_{RT} are independent predictors.

Methods

Participants

Participants were drawn from the Sydney Memory and Ageing Study (MAS), a longitudinal cohort of community-living older adults recruited through the electoral roll, aged 70-90 years and not demented at study entry (23). MAS exclusion criteria

were history of dementia, or suspected dementia based on baseline assessment and consensus diagnosis from an expert panel (see below), or MMSE score less than 24 adjusted for age, education, and non-English-speaking background, psychotic symptoms, schizophrenia, bipolar disorder, multiple sclerosis, motor neuron disease, developmental disability, progressive malignancy or insufficient English to complete a psychometric assessment. Of the MAS baseline sample of 1037 participants, the study sample consisted of 861 native English-speakers who completed baseline RT tasks and neuropsychological assessments (98.6% of the native English speaking group). The study was approved by the institutional Research and Ethics Committee. Participants gave written informed consent.

Assessment

Comprehensive assessments incorporating medical history, physical examination, cognitive measures and informant interviews were administered by trained psychology graduates at baseline, 2-year and 4-year follow-up. The neuropsychological battery included 10 standardized tests measuring attention/processing speed, memory, language, visuospatial and executive function (23, 24) (Supplementary Material 1). Informant ratings of instrumental activities of daily living were made using the Bayer Activities of Daily Living (ADL) Scale at each wave (25).

Reaction time tasks and measures

Simple and complex RT tasks were administered using a touch screen computer with millisecond accuracy and stylus pen (Figure 1). For the Simple RT task, following 4 practice trials, 36 test trials were administered over two sessions. A

yellow square was presented (interstimulus interval (ISI) of 1, 2 or 4 seconds, in random order) and participants were instructed to press it as quickly as possible. For the Complex RT task, 40 test trials were presented over two sessions where two coloured squares appeared vertically (Red-Red, Yellow-Yellow, Red-Yellow or Yellow-Red; 10 of each type pseudorandomly presented; ISI 3 seconds). If the squares were of the same colour, participants had to press the upper square; if the squares were of a different colour, they pressed the lower square. To enhance accuracy, practice trials were administered until four correct responses were made and brief instructions were repeated after each error during the test.

In line with established procedures in RT research (8), prior to computing RT metrics, unusually fast (Simple: <250 ms, Complex: <400ms) and long trials (>3 SD above age group mean [≤75, 76-80; 81-85 and ≥86 years] were removed. The number of trials removed was small (SRT: lower trim n=11, upper trim n=282 (0.95%), CRT: lower trim n=70, upper trim n=388 (1.34%). After error trials were removed from the Complex task, total number of trials replaced was 3.77% (n= 1298). Missing trials were replaced by imputing values using a regression procedure to guard against aggregate values being inflated by extreme scores. Intraindividual mean RTs were computed for each participant, for each task. The intraindividual standard deviation of RTs (ISD) was used as the measure of $IIV_{RT.}$ Computation of ISD followed established methods whereby a regression procedure was used to partial out effects of time-on-task and age (and their interaction) and the residuals obtained were standardised (7). RT and IIV_{RT} scores were averaged across Session 1 and Session 2 to obtain the most reliable estimates.

Dementia risk factors

Two categories of dementia risk factors were examined: 1) baseline clinical measures - MMSE score (20) for global cognitive function, Goldberg Anxiety Scale (26), Geriatric Depression Scale (15 item: 27) and 2) baseline medical and genetic measures- a Framingham-type cardiovascular disease (CVD) risk score derived from current smoking status, diabetic status, systolic blood pressure, total cholesterol level, high-density lipoprotein (HDL) level, current anti-hypertensive medications; and ε4 status derived from genomic DNA extracted using standard methods and *APOE* ε2/3/4 genotyping using Taqman assays (described in 28) with ε4 carriers possessing at least one ε4 allele.

Diagnosis of dementia

At 2-year and 4-year follow-ups, consensus diagnoses were made by at least three experienced clinicians from an expert panel of neuropsychiatrists, psychogeriatricians and neuropsychologists using all available clinical, neuropsychological, laboratory and imaging data, and collateral information from informants. Diagnosis of all-cause dementia was made in accordance with DSM-IV criteria (29) and required deficits in at least two cognitive domains including memory, and impairment in instrumental activities of daily living.

Statistical analyses

Baseline characteristics of those with and without incident dementia were compared using Student's *t* tests and chi-square (χ^2) tests. Cross-tabulations were performed to explore the relationship between baseline RT performance (high versus low

according to median spilt) and follow-up cognitive status (dementia or no dementia). Cox proportional-hazards models were used to examine the influence of baseline mean RT and IIV_{RT} on time to all-cause dementia over 4-year follow-up. Time to dementia was calculated at the midway point between the follow-up assessment when dementia was diagnosed, and the previous assessment. Mean RT and IIV_{RT} measures were transformed to Z-scores and each measure entered into separate models for Simple and Complex tasks, and their effects were estimated both with and without adjustment for dementia risk factors and demographic variables. In the first model, no covariates were included. Subsequent models included demographic (age, sex, years of education) plus 1) clinical (baseline MMSE score, depression and anxiety scores) or 2) medical (APOE ε4 carrier status, and CVD risk score) or 3) both clinical and medical variables. Receiver operating characteristic (ROC) analyses were used to calculate area under the curve (AUC) for RT measures and their combination and each neuropsychological measure and their combination for the prediction of incident dementia over 4 years. The relationship between mean RT and IIV from the same task was examined with Pearson correlations, and hierarchical Cox models were used to estimate incremental prediction of IIV_{RT} over mean RT, and mean RT over IIV_{RT}, from the same task (adjusting for age, sex, and years of education). The statistical significance of these additional variables was obtained using the χ^2 test for the change in log likelihood ratio. The assumption of proportional hazards was checked using Schoenfeld residuals of the covariates to calculate goodness of fit (significance values ranged from .100 to .971) and by examining log-log survival curves associated with different values of the covariate (curves were approximately parallel). Based on these checks, the proportional hazards

assumption did not appear to be violated for any measure. SPSS Version 22.0 was used.

Results

Sample characteristics

Of the 861 participants, 48 cases of incident dementia (5.6%) were identified and 600 participants (69.7%) remained non-demented over follow-up (median 3.9 years), 86 participants (10%) died and 127 (14.8%) dropped out before the last follow-up. Of all incident dementia cases, one reverted to MCI at four-year follow-up although it is notable that two years later, further cognitive decline and functional impairment indicated that this case had progressed to dementia. Dementia diagnoses were subtyped according to established criteria; Alzheimer's disease (probable *n*= 25, possible *n*= 7); vascular dementia (*n*=8); Dementia with Lewy bodies (*n*=2); Parkinson's disease dementia (*n*=1) and dementia with multiple aetiologies or where no specific subtype could be determined (*n*=5).

Those with incident dementia were significantly older, had lower baseline MMSE, higher frequency of ε4 allele, slower mean RT on Simple and Complex tasks, and higher IIV_{RT} for the Complex task though error rate was low and did not differ from those who remained nondemented (Table 1). Among participants with a slower or more variable performance (i.e., equal or higher score than the median of the sample at baseline), dementia incidence was increased by two-fold for Simple mean RT, by almost three-fold for Complex mean RT and close to 2-fold for Complex IIV_{RT} (Table 2).

[Insert Table 1 – Baseline sample characteristics]

[Insert Table 2 –Cognitive Status by median split]

Predictive capability of RT measures for incident dementia – time to dementia

In unadjusted Cox models (Model 1 in Table 3), longer mean RTs for Simple and Complex tasks, and higher IIV_{RT} for the Complex task, when examined individually, were associated with a significantly shorter time to incident dementia while Simple IIV_{RT} was not. A one standard deviation increase in mean RT raised the hazard by approximately 50-60%. A one standard deviation increase in Complex IIV_{RT} raised the hazard by approximately 40%. Kaplan-Meier survival curves (Figure 2) show that longer mean RT (+1 SD above mean) for the Simple (A) and Complex (B) tasks and higher IIV_{RT} score (+1 SD above mean) on the Complex task (C) were associated with shorter time to dementia.

[Insert Table 3 –Cox Models]

[Insert Figure 2 – Survival curves]

In multivariable Cox regression models formed by inclusion of control variables together with each individual RT measure singly, Simple and Complex mean RT measures remained significant in the final model after adjusting for all clinical and medical dementia risk covariates (Model 4 in Table 3). Complex IIV_{RT} effects remained significant but slightly weaker with clinical and medical covariates included in separate models (Models 2 and 3), but failed to reach significance when all covariates were included (Model 4). Age, MMSE score and presence of APOE ε4 were also identified as significant risk factors for a shorter time to incident dementia in final models (results not shown).

In an exploratory analysis, RT measures were examined separately for 'same colour' and 'different colour' trials based on a large literature demonstrating RT disparities for same-different judgements (e.g. 30). Analyses of the two trial types from the Complex task showed 'different colour' trial measures had larger effects than respective mean RT and IIV_{RT} measures for 'same colour' trials. IIV_{RT} for 'different colour' trials remained a significant predictor of all-cause dementia after adjusting for all dementia risk factors (HR=1.36 (1.06-1.74), Wald = 6.04(1), p=.01).

Independence of RT measures in predicting time to incident dementia

Complex mean RT and IIV were strongly correlated (Pearson's r (861) =0.71 p<0.001). Yet, the addition of Complex mean RT to Complex IIV_{RT} in a Cox regression model did significantly improve prediction of dementia (χ^2(1)= 7.1, p=.008) although the addition of Complex IIV_{RT} to the model containing Complex mean RT did not (χ^2(1)=0.00, p=.99) . Exploratory hierarchical Cox regression models examined relative predictive strengths of Simple versus Complex RT measures for all-cause dementia. Addition of Complex mean RT and Complex IIV_{RT} to Simple mean RT failed to improve the model (χ^2(2)=2.37, p=.31), and similarly, addition of Simple mean RT to both Complex measures also did not reach significance ($\chi 2(1)$=2.87, p=.09) suggesting that neither task provided significant additional prediction over the other.

Predictive capability of RT and neuropsychological measures for incident dementia – ROC for incident dementia

ROC analyses (Table 4) revealed that the neuropsychological measures combined and delayed recall memory tests had the highest AUCs. However, AUCs for the four reaction time measures in combination compared favourably to a number of traditional neuropsychological measures particularly when demographic information (age, sex, years of education) was included, ranking equal fourth best predictor. Mean RT measures were slightly stronger predictors of incident dementia than IIV_{RT} measures.

[Insert Table 4 –ROC analyses]

Utility of RT and neuropsychological measures for predicting functional decline

Considering the relatively low number of incident dementia cases, we examined prediction of change in functional ability; defined as the difference between baseline and 4-year follow-up scores on the Bayer ADL, using ordinary least squares regression with Bonferroni adjustment for multiple comparisons (0.05/16, $p<0.003$). Prediction of functional decline based on four RT measures combined had virtually the same effect strength (Multiple R = .233, $F(4, 31.092)= 8.04$, $p<.001$) as the 10 neuropsychological tests combined (Multiple R = .238, $F(10, 30.398) = 3.27$, $p<.001$). Considering individual measures, three RT and five neuropsychological measures significantly predicted 4-year functional decline. The best RT measures were stronger predictors than the best neuropsychological measures (Multiple R: Simple mean RT =.22; Complex mean RT =.17; Category fluency = .14; Trail Making B =.14).

Discussion

In this prospective study, both types of RT measure - mean RT and IIV_{RT} - independently predicted time to all-cause dementia in a large community-based

cohort of older adults free of dementia at baseline. Slower Simple and Complex mean RT increased risk of developing dementia over 4 years by 50-60% and higher variability on the Complex task increased risk by 40% per SD increase. Mean RT had the strongest association, with Simple and Complex mean RT showing comparable effects after controlling for demographic, clinical and medical dementia-risk variables. Complex IIV_{RT} independently predicted dementia over demographic, clinical and medical covariates when considered in separate models, but not in the full model inclusive of all dementia risk factors. Furthermore, RT measures were comparable to several neuropsychological tests for predicting 4-year incident dementia including those measuring processing speed, visual attention, spatial problem-solving, mental flexibility and language, although not surprisingly classical memory measures and a full neuropsychological battery were superior. Mean RT measures showed superiority over IIV_{RT} across all models, and when compared in the same model Complex mean RT added incremental prediction to Complex IIV_{RT}, but not the reverse.

Our findings support and extend the previous literature, which suggests that slower processing speed and higher IIV_{RT} may be useful behavioural markers in individuals destined to subsequently cognitively decline. The focus of most previous work has been on mild impairment states including MCI where eventual outcome is not known (17, 18). Only a single memory clinic study (19) followed amnestic MCI patients and observed higher variability but not slower RT in a small group of MCI converters (n=13) who developed dementia over 2.5 years compared to non-converters (n=26). However, no other cognitive measures or dementia risk factors were controlled for in the analysis. We extend the current literature by demonstrating that RT measures

are also sensitive predictors of future dementia in a large unselected old-age community cohort, even after controlling for global cognitive function (MMSE) and other major dementia risk factors.

Notably, mean RT was a stronger predictor of imminent dementia than $IIV_{RT,}$ and no dissociation of IIV_{RT} from mean RT was observed when examined in the same model. The few previous longitudinal studies suggest that IIV_{RT} is more sensitive to cognitive decline than mean RT obtained from the same task (17-19). However, cross-sectional work suggests that severity of cognitive impairment may be a factor. IIV_{RT} may be more sensitive to subtle cognitive disturbances related to early neurobiological dysfunction in mildly impaired individuals while mean RT may be more discriminatory in more severely impaired individuals or those with dementia (2, 14), although this is not an entirely consistent finding (21). Our findings based on survival analyses favour mean RT as predictive of a shorter time to dementia diagnosis. Given a few years proximity to dementia in some of our participants, the neuropathological cascade accompanying dementia may be more advanced, thereby perhaps reducing the discriminatory power of IIV_{RT} relative to mean RT. Task complexity may be an important factor since moderately demanding tasks appear to be more predictive of subsequent decline (18, 31). Consistent with this, Complex IIV_{RT} was a predictor of dementia while Simple IIV_{RT} was not. Our Complex task incorporating same-different judgements was more sensitive to incident dementia, particularly the ‘different colour’ trials perhaps because of greater demands on higher level attention and executive control processes.

There is debate about whether IIV_{RT} is independent of mean RT or is a result of cognitive slowing. The two measures were highly correlated in our study, consistent with others (8, 17). However, we did not observe an effect of IIV_{RT} independent of mean RT. The few studies examining IIV_{RT} while controlling for mean RT in the analyses are inconsistent; with variability predicting mild cognitive disorder independent of mean RT level in some (22, 32) but not others (15, 21). This inconsistency may stem from the method of computing variability, specifically the extent to which it is independent of mean performance. For example, Cherbuin *et al* (17) used two different computation methods and observed that the IIV_{RT} measure that did not adjust for RT and was highly correlated with RT, was the stronger predictor of transition to mild cognitive disorders compared to the IIV_{RT} measure that corrected for RT. Hence, whether increased IIV_{RT} is an independent marker of impending cognitive decline to mild neurocognitive disorders and to dementia (separable from slowed RT) has yet to be fully established.

Another aim of this study was to compare the predictive validity of RT measures with a broad range of psychometrically validated neuropsychological measures. To date, only the aforementioned Cherbuin *et al* study has examined RT performance along with a small number of neuropsychological tests and found that IIV_{RT} was the best cognitive predictor overall of transition to a variety of mild cognitive disorders in community-living 60-64 year-olds (17). We examined cognitive predictors with and without adjunctive demographical information (age, sex, years of education) as these data are traditionally used in conjunction with cognitive performance. The neuropsychological battery in full (10 tests) (plus demographics) had the highest predictive accuracy which is not surprising given that neuropsychological

performance was a major component of the diagnostic formulation of dementia. The best individual tests were delayed memory recall measures (Logical Memory, Rey Auditory Verbal Learning Test) and category fluency (animals), which have high predictive utility for dementia (33). The combined RT measures (mean RT and IIV_{RT} from Simple and Complex tasks) with demographical information had good predictive accuracy, ranked equal fourth best cognitive performer in predicting 4-year incident dementia and comparable or better to other non-memory tests. Moreover, the combined RT measures predicted functional decline over four years comparably to a full neuropsychological battery. Hence, RT measures are sensitive to incipient decline in everyday tasks which has implications for capacity for independent living as well as for progression to dementia. Therefore, RT measures have value-added appeal since four measures can be derived from two brief tasks which only take approximately four minutes to perform, can be easily measured and do not require high levels of training to administer and score as do neuropsychological tests, were entirely independent of the diagnostic formulation of dementia and performed very well in comparison to the gold standard neuropsychological armoury of psychometric tests. Moreover, our findings suggest that reducing to a single RT task does not reduce predictive accuracy substantially, with Simple mean RT performing comparably to Complex mean RT in predicting dementia, and the strongest predictor of functional decline.

This study highlights the potential of RT tasks to detect early cognitive changes associated with a variety of dementia types within a representative community-based cohort. While two-thirds of our dementia cases were due to Alzheimer's disease, we expect that RT measures may also be appropriate for early detection of Vascular

dementia, Parkinson's disease dementia and Dementia with Lewy bodies since they capture cognitive slowing and attention/executive difficulties, are sensitive to white matter degradation and increased hyperintensity load (11, 34), and dopamine binding (12). Furthermore, RT measures have potential scalability for pragmatic, time- and cost-effective screening of at-risk individuals from a broad section of the population including those from culturally and linguistically diverse backgrounds in a variety of settings; primary care, research or clinical trials because they are brief, do not require linguistic content or high levels of expertise for administration.

Study Strengths and Limitations

This study has several strengths, including large sample size, a well characterised cohort and access to consensus diagnoses of dementia made by an expert panel using standardised criteria. Notably, this is the first study to examine the prognostic utility of RT measures on time to dementia while controlling for an array of clinical and medical dementia risk factors. The sample was sourced from a community-based older population, hence the findings have potential value for developing community screening measures particularly when more detailed neuropsychological assessment is not feasible. The results of this study are subject to some limitations. First, low numbers prevented examination of predictors for different dementia types. Second, responses via mouse or key press rather than a stylus would have minimised the motor component, more clearly reflecting the central processing element. Replication is needed with different RT paradigms. For example, our exploratory analysis of 'same' and 'different' trials suggests that task conditions with greater cognitive demands may be more sensitive. Also a larger number of trials may be required given that some have suggested that IIV_{RT} measures have lower

reliability than mean RT (35) although other work suggests that relatively few RT trials produce statistically reliable predictions of potential neuropathology (36).

Figure Legends

Figure 1. Schematic of stimuli and trial configurations for the Simple and Complex RT tasks.

Figure 2. Survival curves for prediction of time to dementia based on unadjusted models (shown at the mean, 1 SD above and 1 SD below the mean) as a function of variation in (A) Simple mean reaction time, (B) Complex mean reaction time, and (C) Complex IIV_{RT}.

Supplemental Digital Content 1 (text): List of measures from neuropsychological test battery

Acknowledgements: The authors thank all participants and their supporters in the Sydney Memory and Ageing Study (MAS), and the MAS research team.

References

1. Gordon B,Carson K: The basis for choice reaction time slowing in Alzheimer's disease. Brain Cogn 1990; 13:148-166

2. Gorus E, De Raedt R, Lambert M, et al: Reaction times and performance variability in normal aging, mild cognitive impairment, and Alzheimer's disease. J Geriatr Psychiatry Neurol 2008; 21:204-218

3. Mendez MF, Cherrier MM,Perryman KM: Differences between Alzheimer's disease and vascular dementia on information processing measures. Brain Cogn 1997; 34:301-310

4. Salthouse TA: Aging and measures of processing speed. Biol Psychol 2000; 54:35-54

5. Bunce D, MacDonald SW,Hultsch DF: Inconsistency in serial choice decision and motor reaction times dissociate in younger and older adults. Brain Cogn 2004; 56:320-327

6. West R, Murphy KJ, Armilio ML, et al: Lapses of intention and performance variability reveal age-related increases in fluctuations of executive control. Brain Cogn 2002; 49:402-419

7. Hultsch DF, Strauss E, Hunter MA, et al: Intraindividual variability, cognition and aging, in The handbook of aging and cognition. Edited by Craik FIMaS, T.A. New York, Psychology Press, 2008, pp 491-556

8. Hultsch DF, MacDonald SW, Hunter MA, et al: Intraindividual variability in cognitive performance in older adults: comparison of adults with mild dementia, adults with arthritis, and healthy adults. Neuropsychology 2000; 14:588-598

9. de Frias CM, Dixon RA, Fisher N, et al: Intraindividual variability in neurocognitive speed: a comparison of Parkinson's disease and normal older adults. Neuropsychologia 2007; 45:2499-2507

10. Stuss DT, Murphy KJ, Binns MA, et al: Staying on the job: the frontal lobes control individual performance variability. Brain 2003; 126:2363-2380

11. Bunce D, Anstey KJ, Christensen H, et al: White matter hyperintensities and within-person variability in community-dwelling adults aged 60-64 years. Neuropsychologia 2007; 45:2009-2015

12. MacDonald SW, Cervenka S, Farde L, et al: Extrastriatal dopamine D2 receptor binding modulates intraindividual variability in episodic recognition and executive functioning. Neuropsychologia 2009; 47:2299-2304

13. Duchek JM, Balota DA, Tse CS, et al: The utility of intraindividual variability in selective attention tasks as an early marker for Alzheimer's disease. Neuropsychology 2009; 23:746-758

14. Strauss E, Bielak AAM, Bunce D, et al: Within-person variability in response speed as an indicator of cognitive impairment in older adults. Neuropsychol 2007; 14:608-630

15. Christensen H, Dear KB, Anstey KJ, et al: Within-occasion intraindividual variability and preclinical diagnostic status: is intraindividual variability an indicator of mild cognitive impairment? Neuropsychology 2005; 19:309-317

16. Murtha S, Cismaru R, Waechter R, et al: Increased variability accompanies frontal lobe damage in dementia. J Int Neuropsychol Soc 2002; 8:360-372

17. Cherbuin N, Sachdev P,Anstey KJ: Neuropsychological predictors of transition from healthy cognitive aging to mild cognitive impairment: The PATH through life study. Am J Geriatr Psychiatry 2010; 18:723-733

18. Bielak AAM, Hultsch DF, Strauss E, et al: Intraindividual variability in reaction time predicts cognitive outcomes 5 years later. Neuropsychology 2010; 24:731-741

19. Tales A, Leonards U, Bompas A, et al: Intra-individual reaction time variability in amnestic mild cognitive impairment: a precursor to dementia? J Alzheimers Dis 2012; 32:457-466

20. Folstein MF, Folstein SE,McHugh PR: "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975; 12:189-198

21. Phillips M, Rogers P, Haworth J, et al: Intra-individual reaction time variability in mild cognitive impairment and Alzheimer's disease: gender, processing load and speed factors. PLoS ONE 2013; 8:e65712

22. Dixon RA, Garrett DD, Lentz TL, et al: Neurocognitive markers of cognitive impairment: exploring the roles of speed and inconsistency. Neuropsychology 2007; 21:381-399

23. Sachdev PS, Brodaty H, Reppermund S, et al: The Sydney Memory and Ageing Study (MAS). Methodology and baseline medical and neuropsychiatric characteristics of an elderly epidemiological non-demented cohort of Australians aged 70-90 years. Int Psychogeriatr 2010; 22:1248-1264

24. Kochan NA, Slavin MJ, Brodaty H, et al: Effect of different impairment criteria on prevalence of objective mild cognitive impairment in a community sample. Am J Geriatr Psychiatry 2010; 18:711-722

25. Hindmarch I, Lehfeld H, de Jongh P, et al: The Bayer Activities of Daily Living Scale (B-ADL). Dement Geriatr Cogn Disord 1998; 9:20-26

26. Goldberg D, Bridges K, Duncan-Jones P, et al: Detecting anxiety and depression in general medical settings. Br Med J 1988; 297:897-899

27. Sheikh J,Yesavage J: Geriatric Depression Scale (GDS) Recent evidence and development of a shorter version, in Clinical Gerontology : A Guide to Assessment and Intervention. Edited by TL B. New York, The Haworth Press, 1986, pp 165-173

28. Song F, Poljak A, Crawford J, et al: Plasma apolipoprotein levels are associated with cognitive status and decline in a community cohort of older individuals. PLoS ONE [Electronic Resource] 2012; 7:e34078

29. American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders. 4th Edition (DSM-IV), Washington, D.C., 1994

30. Farell B: Same-different judgments: a review of current controversies in perceptual comparisons. Psychol Bull 1985; 98:419-456

31. de Frias CM, Dixon RA,Camicioli R: Neurocognitive speed and inconsistency in Parkinson's disease with and without incipient dementia: an 18-month prospective cohort study. J Int Neuropsychol Soc 2012; 18:764-772

32. Lovden M, Li S-C, Shing YL, et al: Within-person trial-to-trial variability precedes and predicts cognitive decline in old and very old age: longitudinal data from the Berlin Aging Study. Neuropsychologia 2007; 45:2827-2838

33. Gallagher D, Mhaolain AN, Coen R, et al: Detecting prodromal Alzheimer's disease in mild cognitive impairment: utility of the CAMCOG and other neuropsychological predictors. Int J Geriatr Psychiatry 2010; 25:1280-1287

34. Fjell AM, Westlye LT, Amlien IK, et al: Reduced white matter integrity is related to cognitive instability. The Journal of Neuroscience 2011; 31:18060-18072

35. Schmiedek F, Lovden M,Lindenberger U: On the relation of mean reaction time and intraindividual reaction time variability. Psychol Aging 2009; 24:841-857

36. Bunce D, Bielak AA, Cherbuin N, et al: Utility of intraindividual reaction time variability to predict white matter hyperintensities: a potential assessment tool for clinical contexts? J Int Neuropsychol Soc 2013; 19:971-976

Table 1. Baseline sample characteristics based on cognitive outcome after 4 years.

Variable	No dementia (n=813) Mean (SD)	Incident dementia (n=48) Mean (SD)	Test statistic	*p* value
Age, years	78.55 (4.75)	80.28 (4.75)	t(859) =-2.46	.014
Sex, number of Ps				
Male	355	25	χ^2=1.30	.254
Female	458	23		
Education, years	11.62 (3.48)	11.92 (4.05)	t(85971)=-.57	.572
MMSE score	28.59 (1.34)	28.13 (1.12)	t(55.16)=2.77	.008
NART IQ	107.54 (9.97)	107.66 (11.29)	t(845)=-.10	.938
GDS score	2.21 (2.00)	2.23 (1.89)	t(854)=-.08	.934
GAS score	0.98 (1.68)	0.81 (1.47)	t(853)=.67	.502
CVD risk score	17.12 (3.41)	17.70 (3.44)	t(828)=-1.11	.270
APOE ε4 allele %[a]	22.01%	39.58%	χ^2=7.87	.005
Simple Mean RT	615.02 (180.27)	734.01 (241.57)	t(50.14)=-3.36	.002
Simple IIV_{RT}	6.19 (4.86)	6.98 (5.31)	t(859)=-1.08	.279
Complex Mean RT	943.88 (190.03)	1064.19 (239.17)	t(50.57)=-3.42	.001
Complex IIV_{RT}	6.77 (2.89)	7.92 (3.22)	t(859)=-2.66	.008
Errors[b]	2.42 (2.71)	2.60 (2.72)	t(859)=-.45	.656

Abbreviations: SD, standard deviation; MMSE, Mini-Mental State Examination; NART, National Adult Reading Test; GDS, Geriatric Depression Scale; GAS, Goldberg Anxiety Scale; CVD, Cardiovascular disease; APOE, apolipoprotein E; RT, reaction time measured in milliseconds; IIV, intraindividual variability.

Means and standard deviations in brackets are presented unless otherwise specified.

[a]Percentage of the 768 participants with APOE data.

[b] Errors on the Complex task

Table 2. Cognitive status at follow-up according to median split of sample's baseline performance on each RT measure

	Median split group	No dementia n (%)	Incident dementia n (%)	χ^2 (p) [a]
Simple Mean RT	Below median	313 (95.4%)	15 (4.6%)	7.22 (.007)
	Above median	296 (88.2%)	33 (10.0%)	
Simple IIV_{RT}	Below median	308 (93.3%)	22 (6.7%)	.40 (.53)
	Above median	301 (92.0%)	26 (8.0%)	
Complex Mean RT	Below median	316 (96.3%)	12 (3.7%)	12.87 (<0.001)
	Above median	293 (89.1%)	36 (10.9%)	
Complex IIV_{RT}	Below median	311 (94.8%)	17 (5.2%)	4.36 (0.037)
	Above median	298 (90.6%)	31 (9.4%)	

RT, reaction time; IIV_{RT}, intraindividual variability.

N=657 i.e. Participants with cognitive status data at 4-year follow-up. Percentages summed across rows.

[a] p-value (2-tailed) is from χ^2 tests with 1 degree of freedom

Median values: Simple Mean RT = 573.28 ms; Complex Mean RT = 908.36 ms; Simple IIV_{RT} = 4.97; Complex IIV_{RT} = 6.11

Table 3. Cox proportional-hazards models of time to all-cause dementia.

	Model 1 Unadjusted			Model 2 Demographics + Clinical			Model 3 Demographics + Medical			Model 4 Demographics + Clinical + Medical		
RT measure	Wald	HR (95% CI)	*p*	Wald	HR (95% CI)	*p*	Wald	HR (95% CI)	*p*	Wald	HR (95% CI)	*p*
Simple Mean RT	22.98	1.53 (1.28-1.81)	<.001	15.07	1.46 (1.20-1.76)	<.001	18.82	1.50 (1.25-1.80)	<.001	16.00	1.48 (1.22-1.80)	<.001
Simple IIV_{RT}	1.43	1.16 (0.91-1.46)	0.23	0.98	1.13 (0.89-1.42)	0.35	2.04	1.20 (0.94-1.53)	0.19	1.16	1.14 (0.90-1.45)	0.31
Complex Mean RT	21.68	1.59 (1.31-1.93)	<.001	13.69	1.52 (1.22-1.89)	<.001	14.30	1.53 (1.23-1.91)	<.001	12.79	1.53 (1.21-1.93)	<.001
Complex IIV_{RT}	8.57	1.43 (1.13-1.82)	.003	5.05	1.36 (1.04-1.77)	.025	3.98	1.33 (1.01-1.75)	.046	3.63	1.32 (0.99-1.75)	.057

Models were run separately for the four RT measures as follows:

Model 1: unadjusted (n=844)

Model 2: adjustments for age, sex, education, MMSE, GDS, GAS (n=833)

Model 3: adjustments for age, sex, education, CVD, APOE ε4 status (n=776)

Model 4: adjustments for age, sex, education, MMSE, GDS, GAS, CVD, APOE ε4 status (n=771)

df =1 for RT measures in all models in Table 3.

Data from 17 participants were excluded from the Cox regression analyses because they were censored before the earliest event.

Abbreviations: HR, hazard ratio; CI, confidence interval; RT, reaction time; IIV_{RT}, intraindividual variability; MMSE, Mini-Mental State Examination; GDS, Geriatric Depression Scale; GAS, Goldberg Anxiety Scale; APOE, apolipoprotein E; CVD, cardiovascular disease risk score.

Table 4. Receiver operating characteristics of reaction time and traditional neuropsychological measures for the prediction of incident dementia

	Measure alone			Measure plus demographic variables[a]		
	AUC	95% CI	p[b]	AUC	95% CI	p[b]
RT measures						
Complex Mean RT	0.68	0.60-0.76	<0.001	0.71	0.63-0.79	<0.001
Simple Mean RT	0.67	0.58-0.76	<0.001	0.72	0.64-0.80	<0.001
Complex IIV_{RT}	0.63	0.55-0.71	0.003	0.67	0.59-0.75	<0.001
Simple IIV_{RT}	0.56	0.48-0.64	0.19	0.67	0.59-0.74	<0.001
Combined RT measures	0.71	0.63-0.79	<0.001	0.74	0.67-0.81	<0.001
Neuropsychological measures						
Logical Memory delayed	0.79	0.72-0.87	<0.001	0.81	0.73-0.88	<0.001
RAVLT delayed	0.78	0.71-0.84	<0.001	0.80	0.74-0.86	<0.001
Category fluency (Animals)	0.75	0.68-0.82	<0.001	0.77	0.70-0.84	<0.001
Coding	0.72	0.64-0.79	<0.001	0.74	0.66-0.81	<0.001
Block Design	0.69	0.61-0.76	<0.001	0.72	0.64-0.80	<0.001
Benton Visual Retention	0.69	0.61-0.77	<0.001	0.72	0.64-0.79	<0.001
Trail Making Test B	0.68	0.60-0.76	<0.001	0.70	0.62-0.77	<0.001
Boston Naming Test	0.66	0.58-0.75	<0.001	0.71	0.64-0.78	<0.001
Trail Making Test A	0.64	0.55-0.72	0.002	0.68	0.61-0.76	<0.001
Letter fluency	0.56	0.48-0.64	0.17	0.66	0.58-0.73	<0.001
Combined neuropsych	0.88	0.83-0.94	<0.001	0.89	0.83-0.94	<0.001

[a] age, years of education, sex

[b] p-value is based on large sample z-approximation

Abbreviations: AUC, area under the curve; CI, confidence interval; RT, reaction time; IIV_{RT}, intraindividual variability; RAVLT, Rey Auditory Verbal Learning Test

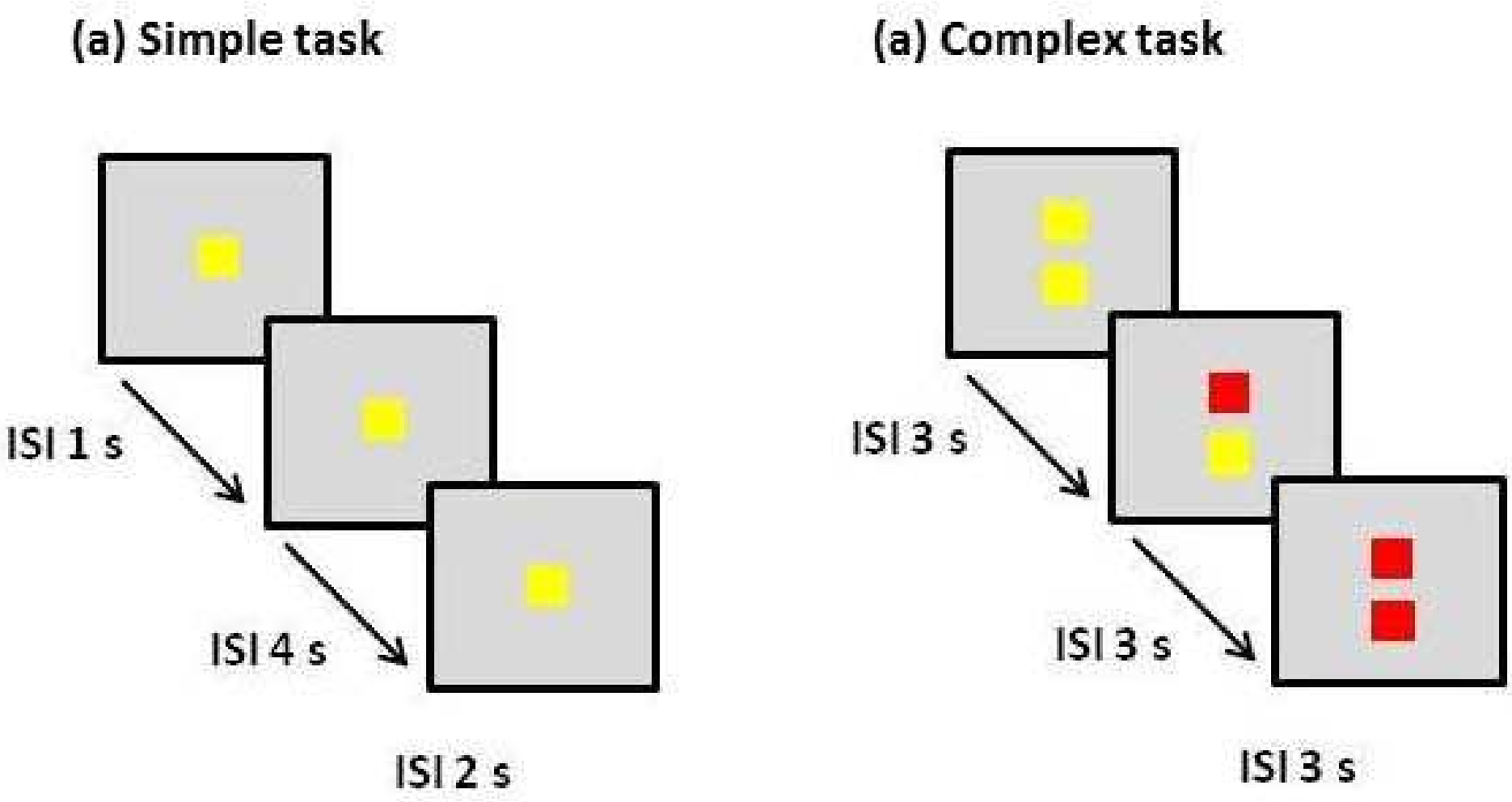
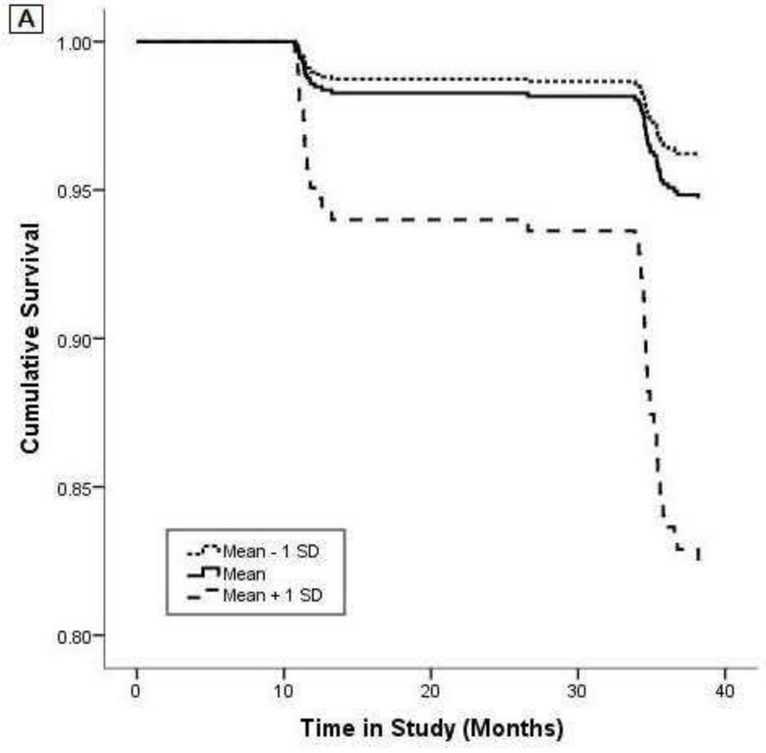
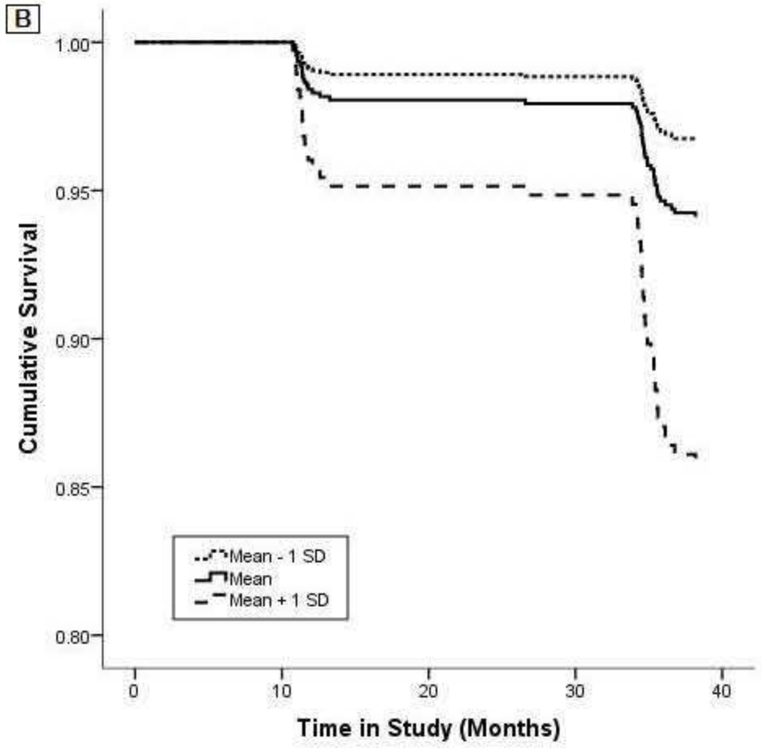
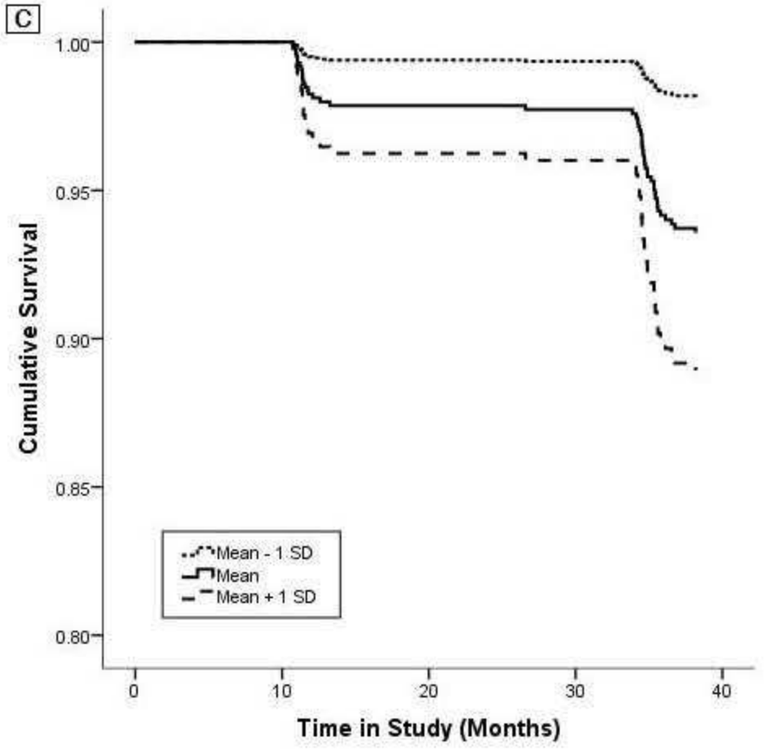


Fig. 1 Stimuli presentation schematic


A
1.00
0.95
0.90
0.85
0.80
Cumulative Survival
Mean - 1 SD
Mean
Mean + 1 SD
0
10
20
30
40
Time in Study (Months)


B
1.00
0.95
0.90
0.85
0.80
Cumulative Survival
Mean - 1 SD
Mean
Mean + 1 SD
0
10
20
30
40
Time in Study (Months)


C
1.00
0.95
0.90
0.85
0.80
Cumulative Survival
Mean - 1 SD
Mean
Mean + 1 SD
0
10
20
30
40
Time in Study (Months)