

Mobile Digital Neurocognitive Assessment for Mild Cognitive Impairment and Alzheimer's Disease: Validation in an Elderly Chinese Population

Kayo Sakamoto¹, Qianhua Zhao² and Simon L Collinson^{1,3,*}

¹Savonix Inc 490 - 43rd Street #50, Oakland, California 94609, United States

²Institute of Neurology, Huashan Hospital, Fudan University, Shanghai, China

³Department of Psychological Medicine, Yong Loo Lin School of Medicine National University of Singapore, Singapore

Abstract

Background: Mobile digital neurocognitive assessment has the potential to facilitate mass screening for MCI and dementia in vulnerable populations by delivering accessible, accurate, standardized measurement of cognition at a vastly reduced administration cost.

Method: 11 subtests of Savonix, a mobile/tablet-based neurocognitive assessment system, were administered to elderly people in hospital and community sites in China in order to assess their applicability in an elderly Chinese population: 29 Mild Cognitive Impairment (MCI) patients, 21 mild Alzheimer's disease (AD) patients, and 265 neuro-normal elderly people. The diagnostic validity in detecting MCI and AD cases from the neuro-normal cases was assessed with Receiver Operating Characteristics (ROC) curve analyses. The convergent validity was assessed by correlating to traditional paper-and-pen screenings such as Mini-Mental State Examination (MMSE), the Alzheimer Disease Assessment Scale Cognitive subscale (ADAS-Cog), and Clinical Dementia Rating (CDR).

Results: Diagnostic accuracies for all 11 tests were established for AD vs. neuro-normal cases (sensitivity 100%, specificity 95%), MCI vs. neuro-normal cases (sensitivity 85.7%, specificity 93.3%), and AD vs. MCI cases (sensitivity 81%, specificity 85.7%). Fewer numbers of tests also showed comparable accuracies to overall 11 tests. However, the three different total scores demonstrated only weak to moderate correlation to MMSE, ADAS-Cog, and CDRSB for the MCI and AD subjects.

Conclusion: Savonix demonstrated sensitive diagnostic performance in detecting both MCI and AD for Chinese population even when a relatively small number of tests are employed making it suitable for mass screening.

Background

In recent years, the movement toward primary prevention strategies for mild cognitive impairment (MCI) and dementia, especially Alzheimer's Disease (AD), has highlighted the importance of early screening and risk identification at scale. Diagnostic criteria for both MCI and AD emphasise the value of biomarkers including β -amyloid protein (A β) and tau protein as early indicators of neuropathology that are associated with AD [1-3], but the cost and accessibility of these methods are significant barriers to widespread adoption.

Neurocognitive screening is another method for routine identification of cognitive decline. Several neurocognitive tests have been shown to be sensitive to both underlying A β and Tau burden [4] as well as the diagnosis of MCI and AD itself [1-3]. Neurocognitive tests are also non-invasive and relatively easy to administer. However, whilst 'in clinic' testing of cognition offers several advantages, primary care physicians, the most likely first channel for potential patients tend to lack technical support, infrastructure, and experience to effectively use such methods [5,6]. Furthermore, the relatively high cost of delivery of neuropsychological services in clinics and hospitals remains prohibitive in many countries especially in regions with the highest rates of dementia and/or looming dementia crises, such as Asia [7].

The advent of COVID-19 has created new barriers to 'in clinic' screening for MCI and dementia as significant numbers of elderly people at higher risk of cognitive decline appear to be avoiding attendance at clinics and hospitals due to infection concerns [8]. The problem is worse for researchers engaged in important work based in tertiary medical centres such as University Hospitals which rely

on voluntary participation. This situation is unlikely to change in the near future, leading to under diagnosis and later diagnosis with all of the attendant costs to patients, their families, healthcare systems and society [9].

Computerised neuropsychological assessment has been in existence for over 20 years and has promised to provide a solution to this problem. This approach addresses the problems of administration and interpretation of neurocognitive tests as they can provide standardized and accurate delivery and reporting of cognitive function [10-12] but to date they have failed to provide the complete solution that is needed for screening at scale. Often such tests are highly expensive, require a physician or assistant to administer and are delivered via desktop or hardware solutions that must be separately purchased and/or may not possess the necessary clinical validity and patient experience to ensure widespread uptake.

A new approach to digital neurocognitive assessment involves the delivery of app-based tests via mobile phone. More than 3 Billion people own or use a smartphone worldwide and is forecast to further

Corresponding Author: Prof. Simon L Collinson, Department of Psychological Medicine, Yong Loo Lin School of Medicine National University of Singapore, 21 Lower Kent Ridge Rd, Singapore 119077; E-mail: scollinson@savonix.com

Citation: Sakamoto K, Zhao Q, Collinson SL (2021) Mobile Digital Neurocognitive Assessment for Mild Cognitive Impairment and Alzheimer's Disease: Validation in an Elderly Chinese Population. Int J Digit Hlthc 1: 105. doi: <https://doi.org/10.15344/ijdh/2021/104>

Copyright: © 2021 Sakamoto et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

grow by several hundred million in the next few years. Contrary to popular perception, older people are increasingly adopting smart phones, nearly three-quarters (74%) of Americans aged 50-64 are smartphone owners, as are 42% of those 65 and older [13]. In China, there is widespread acceptance and adoption of smartphones in the for health applications [14] with over 85 percent of elderly mobile users having downloaded more than 20 apps, whilst half have more than 30 apps installed on their phones. Among middle-aged and elderly internet users (over 40 years old), 65.7 percent spend at least a quarter of their free time on mobile phones every day, and nearly 30 percent spend more than half of their free time on their phones [15].

This study is a pilot to investigate the diagnostic validity of Savonix, a mobile/tablet-based neurocognitive assessment system. We investigated the 11 tests of the Savonix system in older adults from Shanghai, China in order to determine their sensitivity, specificity and to MCI and AD and to examine their convergent validity in relation to standard paper and pencil screening tests - the MMSE, ADAS-Cog and CDR, SB.

Methods

Participants

Forty-nine cognitively impaired older patients (mean age =72 years) were recruited from Hua Shan hospital, affiliated with Fudan University, Shanghai, China. Twenty-eight patients were diagnosed with MCI, and 21 with mild AD. A slightly younger and larger cohort of 265 neuro-normal older adults (mean age = 62) were recruited from two community sites, Jian Ai Charity and Jinmei Care, both located in Shanghai, China for normative purposes. No user indicated color blindness (see Table 1).

Clinical diagnosis

Clinical cases were diagnosed according to National Institute on Aging (NIA)-Alzheimer's Association guidelines [16,17] and Alzheimer's Dementia diagnostic criteria followed The Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) [18], whereas Mild Cognitive Impairment (MCI) diagnostic criteria follows the Petersen (2004) criteria [19]. AD or MCI Diagnosis were typically initiated following an informant or self-reported cognitive

change. History taking was conducted by an experienced clinician. The Chinese version of Mini-Mental State Examination [20] was also performed to evaluate general cognitive level and for diagnosis. Laboratory and imaging (MRI) examinations were used to rule out other neurological or a non-neurological medical comorbidity, including, but not limited to, vascular diseases, trauma, dementia with Lewy bodies. CSF and/or PET are not routinely conducted for AD/MCI patients in the hospital.

Cognitive Assessments

Bedside mental status examinations

Besides the MMSE, clinical patients were administered The Alzheimer Disease Assessment Scale Cognitive subscale Chinese version [21] and Clinical Dementia Rating (CDR) Chinese version [22] by an experienced clinician. All paper-and-pen tests, including MMSE, were taken by patients at least 3 months prior to the mobile cognitive assessment.

Savonix mobile cognitive assessment

All the participants were administered the Savonix mobile cognitive assessment in November and December 2016. For the purposes of initial validation, Savonix was administered an experienced clinician administrator. The assessment included following 11 cognitive subtests:

Immediate/Delayed Verbal Learning: This task measures immediate and delayed verbal memory. In 1916, Édouard Claparède published the Test de mémoire des mots (Test of Memory for Words). This test would become the basis for André Rey's Auditory Verbal Learning Test [23], and subsequent tests of verbal learning, including the Savonix Verbal Learning Task. The task presents patients with a list of words, which are to be committed to memory. Patients are tested on their ability to remember the words immediately after the original list is removed, and then again after at least a 15-minute delay.

Verbal Interference Part 1/Part 2: This task is based on the Stroop effect [24], which has been widely used in clinical and research settings for the past 50 years as an accurate measure of cognitive control and processing speed. Participants were presented with color words such

Characteristics	Normal (N=265)	MCI (N=28)	AD (N=21)	Comparison test P value*	Univariate logistic regression P value
Age, mean (SD)	62.3 (8.00)	72.14 (8.20)	72.57 (7.00)	< 0.0001	< 0.0001 for Normal vs. MCI < 0.0001 for Normal vs. AD < 0.0001 for Normal vs. MCI/AD 0.844 for MCI vs. AD
Gender, number of female (%)	191 (72.1)	15 (53.6)	11 (52.4)	0.891	NA
Education Years, mean (SD)	10.43 (2.39)	11.57 (3.94)	13.38 (2.46)	< 0.001	< 0.0001 for Normal vs. MCI < 0.05 for Normal vs. AD < 0.0001 for Normal vs. MCI/AD < 0.1 for MCI vs. AD
MMSE total score, mean (SD)		27.64 (1.39)	22.29 (2.12)	< 0.0001	< 0.05
ADAS-Cog total score, mean (SD)		16.36 (4.96)	23.14 (8.47)	< 0.01	< 0.001
CDRSB, mean (SD)		1.61 (1.07)	2.95 (1.77)	< 0.01	< 0.01

Table 1: Population characteristics according to clinical diagnosis and cognitive assessment.

*: *p* value of one-way ANOVA for Age and Education Years. *p* value of Chi-squared test for Gender. *p* value of t-test for MMSE, ADAS-Cog, and CDRSB.

as red, yellow, green, or blue, which were printed in a different color of ink (e.g., red printed in blue ink). Participants were able to easily read the words but found it more difficult to name the color of the ink. Stroop observed that the learned response of reading the words "interfered" with the unfamiliar task of naming the ink color. Naming the ink color requires inhibiting the well-practiced, interfering tendency to read words. This task measures focus by assessing the subject's ability to inhibit automatic and irrelevant responses.

Trail Making Part 1/Part 2: This task is based on the classic test developed by Reitan (1958) [25]. Part 1 of this task is a measure of visual scanning and information processing speed. The user is presented with a pattern of 13 numbers (1-13) on the screen and is required to touch numbers in ascending sequence (i.e., 1, 2, 3...). As each number is touched in correct order, a line is drawn automatically to connect it to the preceding number or letter in the sequence. Part 2 of this task is a measure of cognitive flexibility and attention switching. The user is presented with a pattern of 13 numbers (1-13) and 12 letters (A-L) on the screen and is required to touch numbers and letters alternatively in ascending sequence (i.e. 1, A, 2, B, 3, C...). As each number or letter is touched in correct order, a line is drawn automatically to connect it to the preceding number or letter in the sequence.

Digit Span Forward/ Reverse: The Digit Span task is similar to that subtest of the Wechsler Adult Intelligence Scale [26]. The Forward Digit Span task consists of a number of trials where a series of digits are presented at a constant rate on the device screen. Immediately after each trial, the user is required to enter the digits on a keypad in the order in which they were flashed. In the Reverse Digit Span task, the user is required to enter the digits in reverse order. Sequence length varies between three and 10, with two trials for each length and with trials presented in ascending sequence order. The task ends when the participant fails two trials of any sequence length or when all trials are completed. This task is used as a measure of attention and working memory.

Maze Part 1/Part 2: The Savonix Maze task is similar to the procedure reported by Milner in 1965 [27], which is sometimes referred to as the Austin Maze and is sensitive to hippocampal, frontal lobe, and cerebellar lesions and/degeneration. In the original maze, bolts were attached to a wooden board; no markings were on the board; subjects were required to discover the path of the "maze" through trial and error. When subjects touched the bolts with a metal-tipped stylus, a sound would alert the subject if they were on the right path. Though the Savonix test is similar to the Austin Maze, the maze is displayed on a device screen and feedback is provided visually rather than auditorily.

Complex Figure: The Savonix Complex Figure task is based on the classical Rey-Osterrieth Complex Figure Test (ROCF) [29], a measure of visual spatial and constructional abilities as well as spatial memory, and has been used clinically to evaluate attention, planning, and executive function. The Savonix version is adapted for a touchscreen device. Similar to the ROCF, the outcome variable of the Savonix Complex Figure is accuracy of the reproduced figure against the original figure, as measured by errors versus correct features. Following at least a 15-minute delay period, a free recall is administered in which subjects attempt to redraw the original design from memory. Both delayed cued and free recall memory paradigms have been identified as being important measures used in evaluating Alzheimer's disease.

The Savonix application was preinstalled on a 9-inch tablet. User information such Gender, and Education was entered by an administrator on the tablet, and the tablet was handed to the patients before starting the test. Savonix tests were self-administrated after user information was collected.

Statistical analysis

Statistical analyses were conducted using R- Studio version 1. 3. 1073 with R version 4. 0. 2, pROC package version 1. 16. 2, [29], psych package version 2. 0. 9 [30] and dplyr package version 1. 0. 2, [31].

Demographic variables were compared in the AD, MCI, and neuro-normal groups by one-way Analyses of Variance (ANOVA) for age and education years, and a Person χ^2 analysis for gender. MMSE total scores, ADAS-Cog total scores, and CDR Sum of Boxes were compared between AD and MCI subjects by t-tests. Univariate logistic regression analyses were applied for variables which showed statistical significance in the comparison tests (t- tests or χ^2 analysis).

For each of the Savonix subtests, only one outcome variable was selected to be submitted for subsequent analysis: Ratio of incorrect responses for Immediate/Delayed Verbal Learning, Digit Span Forward/ Reverse, Maze 1 / 2, and Complex Figure; Total reaction time for Verbal Interference Part 1, Trail Making Part 1 / Part 2; and Ratio of incorrect responses divided by Total reaction times for Verbal Interference Part 2.

Nine cognitive domain scores were also composed from the 11 subtests: Instant Verbal Memory, Delayed Verbal Memory, Information Processing Speed, Flexible Thinking, Working Memory, Spatial Memory, Attention, Focus, and Executive Function. Inverse minmax norms of the raw data which were used for following analyses were compared in AD, MCI, and neuro-normal participants by one-way Kruskal-Wallis rank sum tests as several scores were non-normally distributed.

Convergent validity between traditional paper-and-pen screening tests and Savonix tests was measured using the Spearman's correlation. MMSE's total scores, ADAS-Cog's total scores, CDR Sum of Boxes (CDRSB), total Savonix 11 test scores, total Savonix 9 domain scores, and total Savonix 6 memory-related domain scores (Instant Verbal Memory, Delayed Verbal Memory, Information Processing Speed, Flexible Thinking, Working Memory, Spatial Memory) for the MCI and AD subjects were used to calculate those correlation coefficients.

Logistic regression analyses were conducted to evaluate discriminatory ability of Savonix tests relative to the traditional paper-and-pen screening tests (only for MCI vs. AD classification). Based on univariate logistic regression analyses of age and education years, models for Normal vs. MCI discrimination and Normal vs. AD and for Normal vs. MCI/AD were adjusted by age and education years, and models for MCI vs. AD were adjusted by education years. For the MCI vs. AD models, all the sample (21 AD subjects and 28 MCI subjects) were used to estimate one model. For the other discrimination models, down-sampling for the Normal subjects was performed to balance the data: 1000 randomly down-sampled 30 Normal subjects (out of 265) and 28 MCI subjects to estimate 1000 Normal vs. MCI models; 1000 of 20 Normal subjects and 21 AD subjects for the Normal vs. AD models, and 1000 of 50 Normal subjects and 49 MCI-or-AD subjects for the Normal vs. MCI/AD models.

Tests	AUC (95% CI)	Sensitivity	Specificity	PPV	NPV	B		SD	Wald Chisq	P value
Savonix 11 subtests total score	0.866	0.810	0.857	0.810	0.857	Test score	-1.27358	0.39176	10.5685	p<0.01
	(0.760-0.972)					Education	0.071	0.112	0.396	p = 0.529 n.s.
Savonix Selected 3 subtest total score	0.855	0.714	0.929	0.882	0.813	Test score	-2.39999	0.7702	9.784	p<0.01
	(0.744-0.938)					Education	0.109	0.112	0.961	p = 0.327 n.s.
MMSE*	0.991	0.952	1	1	0.966	Test score	-2.567	1.1193	5.2601	p<0.05
	(0.971-1)					Education	0.242	0.353	0.469	p = 0.469 n.s.
ADAS-Cog	0.841	0.750	0.821	0.750	0.821	Test score	0.206	0.064	10.492	p<0.01
	(0.722-0.9601)					Education	0.107	0.120	0.798	p = 0.372 n.s.
CDRSB	0.832	0.900	0.643	0.643	0.900	Test score	0.935	0.354	6.993	p<0.01
	(0.721-0.944)					Education	0.213	0.354	2.864	p = 0.091 n.s.

Table 2: Discriminant indices of Savonix test, MMSE, ADAS-Cog, and CDR Sum of Boxes (CDRSB) to detect AD from MCI. Summary of logistic regression for logistic regressions with the predictor tests adjusted education years.

*: MMSE was used for diagnosis. AUC: Area Under the Receiver Operating Characteristic Curve. CI: Confidence Interval. PPV: Positive Predictive Values. NPV: Negative Predictive Values. B: regression coefficients. SD: Standard Deviations for the regression coefficients. Wald Chisq: Wald Chi square statistics. P value: P values for Wald Chi square statistics. Selected 3 subtests: Trail Making 1; Trail Making 2; and Complex Figure.

Savonix Test composites	AUC (95% CI)	Sensitivity	Specificity	PPV	NPV	B		SD	Wald Chisq	P value
11 Subtests total score										
Max	0.933	0.857	0.933	0.923	0.875	Test score	-2.08	0.691	9.066	p < 0.01
	(0.872-0.994)					Education	0.135	0.151	0.792	p = 0.374 n.s.
						Age	0.123	0.063	3.858	p < 0.05
Min	0.758	0.821	0.633	0.676	0.792	Test score	-0.894	0.430	4.312	p<0.05
	(0.633-0.883)					Education	0.041	0.098	0.177	p = 0.674 n.s.
						Age	0.050	0.048	1.046	p = 0.306 n.s.
Selected 3 subtests total score										
Max	0.917	0.929	0.8	0.813	0.923	Test score	-4.175	1.437	8.453	p < 0.01
	(0.847-0.987)					Education	0.17	0.145	1.36	p = 0.244 n.s.
						Age	0.128	0.062	4.219	p < 0.05
Min	0.745	0.679	0.733	0.704	0.710	Test score	-1.659	0.809	4.197	p < 0.05
	(0.618-0.873)					Education	-0.014	0.098	0.021	p = 0.884 n.s.
						Age	0.073	0.048	2.312	p = 0.128 n.s.

Table 3: Discriminant indices of Savonix test scores, to detect MCI from Normal. Summary of logistic regression for logistic regressions with the predictor tests adjusted education years and ages.

Max/Mix were determined by the AUC values of logistic regression models of 1000 randomly down-sampled 30 Normal subjects (out of 265) and 28 MCI subjects. AUC: Area Under the Receiver Operating Characteristic Curve. CI: Confidence Interval. PPV: Positive Predictive Values. NPV: Negative Predictive Values. B: regression coefficients. SD: Standard Deviations for the regression coefficients. Wald Chisq: Wald Chi square statistics. P value: P values for Wald Chi square statistics. Selected 3 subtests: Trail Making 2; Maze 2; and Complex Figure.

Savonix Test composites	AUC (95% CI)	Sensitivity	Specificity	PPV	NPV	B		SD	Wald Chisq	P value
11 Subtests total score										
Max	0.988	1	0.95	0.955	1	Test score	-1.107	0.47	5.563	p<0.05
	(0.967-1)					Education	0.972	0.539	3.25	p = 0.071 n.s.
						Age	0.287	0.144	3.944	p<0.05
Min	0.921	0.857	0.95	0.947	0.864	Test score	-1.371	0.47	8.51	p<0.01
	(0.831-1)					Education	0.143	0.172	0.689	p = 0.406 n.s.
						Age	0.136	0.066	4.203	p < 0.05
Selected 2 subtests total score										
Max	0.995	1	0.95	0.955	1	Test score	-7.538	2.893	6.787	p < 0.01
	(0.984-1)					Education	0.25	0.343	0.532	p = 0.466 n.s.
						Age	0.326	0.192	2.881	p = 0.090 n.s.
Min	0.902	0.81	0.8	0.81	0.8	Test score	-4.035	1.355	8.874	p < 0.01
	(0.814-0.99)					Education	0.211	0.184	1.311	p = 0.252 n.s.
						Age	0.068	0.076	0.796	p = 0.372 n.s.

Table 4: Discriminant indices of Savonix test scores, to detect AD from Normal. Summary of logistic regression for logistic regressions with the predictor tests adjusted education years and ages.

Max/Mix were determined by the AUC values of logistic regression models of 1000 of 20 Normal subjects and 21 AD subjects. AUC: Area Under the Receiver Operating Characteristic Curve. CI: Confidence Interval. PPV: Positive Predictive Values. NPV: Negative Predictive Values. B: regression coefficients. SD: Standard Deviations for the regression coefficients. Wald Chisq: Wald Chi square statistics. P value: P values for Wald Chi square statistics. Selected 2 subtests: Trail Making 2; Maze 2; and Complex Figure.

The discrimination accuracies of the logistic regression models were evaluated based on the Receiver Operating Characteristic (ROC) analyses. The discriminant indices were: sensitivities, specificities, positive predictive values (PPVs), negative predictive values (NPVs), and Area under the ROC curves (AUCs). For Normal vs. MCI,

Normal vs. AD, and Normal vs. MCI/AD models, only the models with statistically significant regression coefficients for the test scores were considered out of 1000 sampled models, and of those considered models, the models with minimum and maximum AUCs were selected for discriminant indices [See Table 2-5, Figure 1].

Savonix Test composite	AUC (95% CI)	Sensitivity	Specificity	PPV	NPV	B		SD	Wald Chisq	p
11 Subtests total score										
Max	0.949 (0.907-0.992)	0.878	0.98	0.977	0.891	Test score	-2.561	0.653	15.375	p < 0.0001
						Education	0.029	0.134	0.047	p = 0.828 n.s.
						Age	0.082	0.047	3.02	p = 0.082 n.s.
Min	0.825 (0.742-0.908)	0.653	0.96	0.941	0.738	Test score	-0.554	0.207	7.167	p < 0.01
						Education	0.111	0.086	1.656	p = 0.198 n.s.
						Age	0.134	0.041	10.726	p < 0.01

Table 5: Discriminant indices of Savonix test scores, to detect AD or MCI from Normal. Summary of logistic regression for logistic regressions with the predictor tests adjusted education years and ages.

Max/Mix were determined by the AUC values of logistic regression models of 1000 of 50 Normal subjects and 49 MCI-or-AD subjects. AUC: Area Under the Receiver Operating Characteristic Curve. CI: Confidence Interval. PPV: Positive Predictive Values. NPV: Negative Predictive Values. B: regression coefficients. SD: Standard Deviations for the regression coefficients. Wald Chisq: Wald Chi square statistics. P value: P values for Wald Chi square statistics.

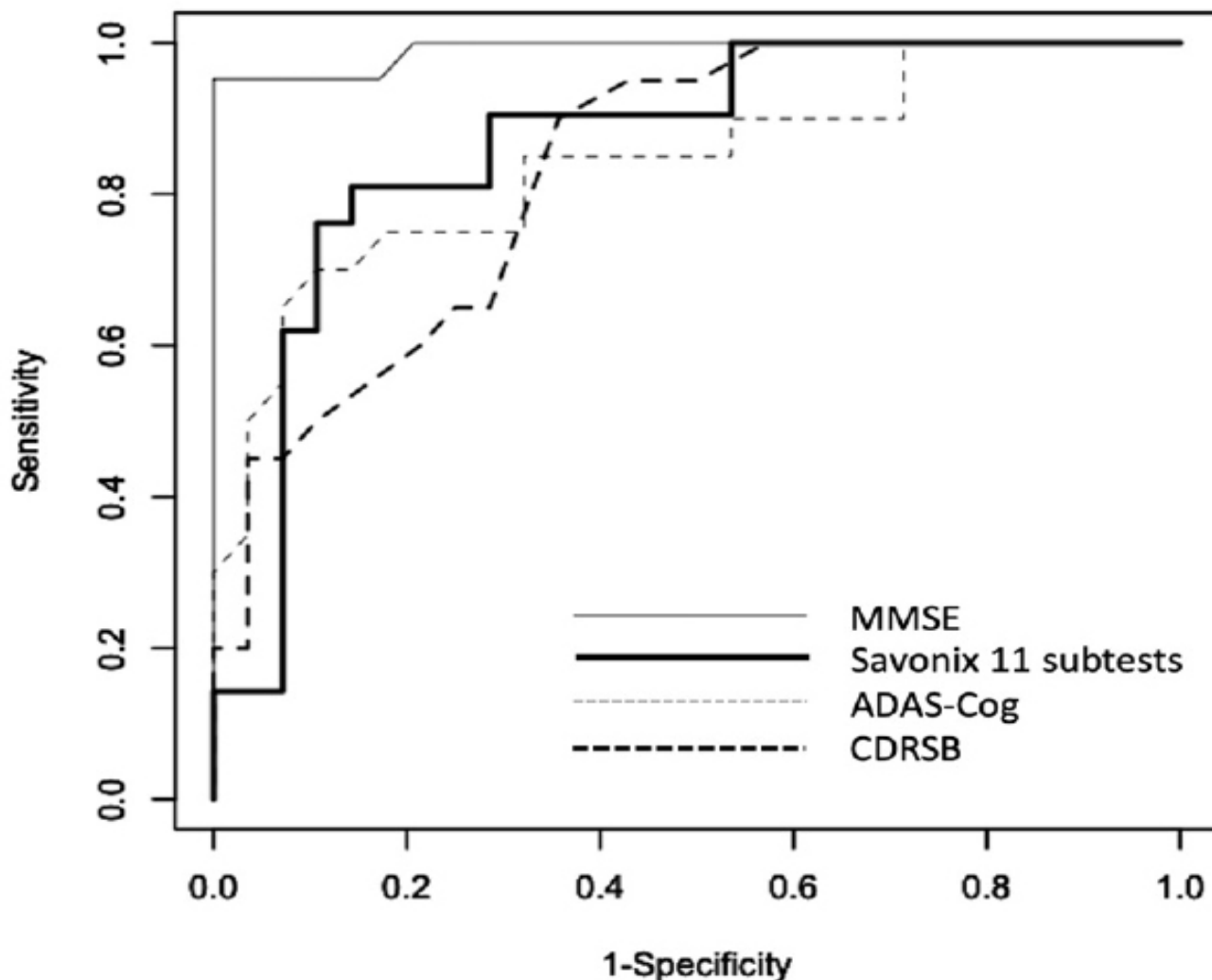


Figure 1: ROC curve analysis for Savonix 11 subtests, MMSE, ADAS-Cog, and CDRSB using scores adjusted with education years.

Results

Participant characteristics

Table 1 lists the descriptive statistics of participants' characteristics, such as age, education gender, years; the traditional paper-and-pen screening tests; as well as results of comparison tests and univariate logistic regression analyses.

Between-group one-way ANOVA revealed significant differences in age [$F(1, 312) = 57.03, p < 0.0001$] and education [$F(1, 312) = 13.45, p < 0.001$]. Bonferroni post-hoc showed significant differences in age between Normal vs. MCI ($p < 0.0001$) and Normal vs. AD ($p < 0.0001$). It also showed significant differences in education years between the Normal vs. AD classification ($p < 0.0001$) and the MCI vs. AD classification ($p < 0.05$). This means that the neuro-normal subjects are younger than those with MCI or AD classification, and AD subjects had more education than neuro-normal subjects and MCI subjects.

Univariate logistic analyses to discriminate the three groups (Normal, MCI, and AD) with age and education years suggested that the cognitive assessment scores need to be adjusted by age and education for models of Normal vs. MCI; Normal vs. AD; and Normal vs. MCI or AD discrimination. It also suggested that these scores need to be adjusted by education for MCI vs. AD discrimination.

Between-group t-tests showed significant differences in the MMSE, the ADAS-Cog, and the CDRSB scores between MCI and AD subjects [$t(47) = 10.05, p < 0.0001$ for MMSE; $t(47) = 3.27, p < 0.01$ for ADAS-Cog; and $t(47) = 3.08, p < 0.01$ for CDRSB]. This means that MCI subjects showed better cognitive abilities than AD subjects.

Savonix test scores

Table 6 shows the variables used to score the 11 subtests; descriptive statistics of the minmax normalized scores; and the results of Kruskal-Wallis tests to compare between neuro-normal, MCI, and AD subjects.

The Kruskal-Wallis tests showed significant differences between Normal, MCI, and AD subject groups in most of the scores except for Verbal Interference Part 1.

Mann-Whitney tests with Bonferroni correction for post-hoc showed pair wise differences: between Normal vs. MCI ($p < 0.01$), Normal vs. AD ($p < 0.0001$), and MCI vs. AD ($p < 0.05$) in Immediate Verbal Learning; between Normal vs. MCI ($p < 0.001$) and Normal vs. AD ($p < 0.0001$) in Delayed Verbal Learning; between Normal vs. AD ($p < 0.001$), and MCI vs. AD ($p < 0.01$) in Verbal Interference Part 2; between Normal vs. MCI ($p < 0.001$), Normal vs. AD ($p < 0.0001$), and MCI vs. AD ($p < 0.001$) in Trail Making Part 1; between Normal vs. MCI ($p < 0.0001$), Normal vs. AD ($p < 0.0001$), and MCI vs. AD ($p < 0.01$) in Trail Making Part 2; between Normal vs. MCI ($p < 0.01$), and Normal vs. AD ($p < 0.01$) in Digit Span Part 1; between Normal vs. MCI ($p < 0.01$), Normal vs. AD ($p < 0.0001$), and MCI vs. AD ($p < 0.05$) in Digit Span Part 2; between Normal vs. AD ($p < 0.01$) in Maze 1; between Normal vs. MCI ($p < 0.001$), and Normal vs. AD ($p < 0.0001$) in Maze 2; between Normal vs. MCI ($p < 0.0001$), Normal vs. AD ($p < 0.0001$), and MCI vs. AD ($p < 0.01$) in Complex Figure; between Normal vs. MCI ($p < 0.0001$), Normal vs. AD ($p < 0.0001$), and MCI vs. AD ($p < 0.001$) in Total 11 test score; between Normal vs. MCI ($p < 0.0001$), Normal vs. AD ($p < 0.0001$), and MCI vs. AD ($p < 0.001$) in Total 6 domain score; and between Normal vs. MCI ($p < 0.0001$), Normal vs. AD ($p < 0.0001$), and MCI vs. AD ($p < 0.001$) in Total 9 domain score.

This result of post-hoc suggests that the discriminant abilities of the Savonix 11 subtests may differ in the sensitive areas such as sensitive only between Normal vs. MCI/AD, or sensitive across all types of discrimination.

Convergent validity

Correlation coefficients between three types of Savonix assessment total scores and the three traditional screening tests are shown in Table 7, which are all significant. The Savonix cognitive assessment has weak to moderate correlation to MMSE, ADAS-Cog, and CDRSB for the MCI and AD subjects.

Subtest	Feature for scoring	Normal (N=265)		MCI (N=28)		AD (N=21)		P value
		M	SD	M	SD	M	SD	
Immediate Verbal Learning	Incorrect response rate	0.85	0.11	0.76	0.15	0.61	0.21	< 0.0001
Delayed Verbal Learning	Incorrect response rate	0.76	0.14	0.62	0.19	0.49	0.26	< 0.0001
Verbal Interference 1		0.85	0.12	0.84	0.11	0.81	0.21	0.863
Verbal Interference 2	Incorrect response rate divided by Total Reaction Time							
Total Reaction time	0.91	0.16	0.92	0.10	0.74	0.25	< 0.001	
Trail Making 1	Total Reaction Time	0.79	0.18	0.68	0.20	0.50	0.20	< 0.0001
Trail Making 2	Total Reaction Time	0.79	0.17	0.64	0.25	0.32	0.31	< 0.0001
Digit Span Forward	Incorrect response rate	0.67	0.15	0.57	0.20	0.44	0.33	< 0.0001
Digit Span Reverse	Incorrect response rate	0.52	0.20	0.41	0.21	0.24	0.20	< 0.0001
Maze 1	Total number of incorrect tap	0.92	0.13	0.94	0.03	0.88	0.13	< 0.01
Maze 2	Total number of incorrect tap	0.93	0.12	0.90	0.09	0.85	0.14	< 0.0001
Complex Figure	Correct Input	0.60	0.28	0.36	0.23	0.17	0.09	< 0.0001
11 Test Total		8.59	0.86	7.64	1.08	6.04	1.1	< 0.0001
6 Domain Score Total	Sum of 6 memory-related cognitive domain scores	3.94	0.56	3.18	0.72	2.25	0.68	< 0.0001
9 Domain Score Total	Sum of 6 memory-related cognitive domain scores, Focus scores, Attention scores, and Executive Function scores	6.04	0.81	5.04	1.05	3.52	1.11	< 0.0001

Table 6: Normalized scores of Savonix 11 subtests and total scores for 265 neuronormal, 28 MCI, and 21 AD subjects. M: mean values. SD: standard deviations. P value: p values for one-way Kruskal-Wallis Chi-squared.

Screening Test	Savonix Total Score	Correlation r	P value
MMSE	11 Test Total	0.625	< 0.0001
	6 Domain Score Total	0.608	< 0.0001
	9 Domain Score Total	0.601	< 0.0001
ADAS-Cog	11 Test Total	0.646	< 0.0001
	6 Domain Score Total	0.553	< 0.0001
	9 Domain Score Total	0.545	< 0.0001
CDRSB	11 Test Total	0.438	< 0.01
	6 Domain Score Total	0.340	< 0.05
	9 Domain Score Total	0.315	< 0.05

Table 7: Convergent validity for Savonix test. r: Spearman's rho.

Discriminant ability of Savonix 11 subtests

Figure 1 shows the result of ROC curve analysis for Savonix 11 subtests, MMSE, ADAS-Cog, and CDRSB using scores adjusted with education years to discriminate MCI and AD. Table 2 shows the discriminant indices of those tests as well as the summary of logistic regressions. The sensitivities, specificities, PPVs, and NPVs are determined by the top-left method. Except for the MMSE result which was used for diagnosis, the Savonix 11 subtests shows better AUC compared to those for the ADAS-Cog and the CDRSB though there were no statistically significances in DeLong's ROC comparison tests ($p=0.763$ for 11 tests vs. ADAS-Cog, $p=0.670$ for 11 tests vs. CDRSB).

Table 3, 4, and 5 show the discriminant indices of Savonix 11 test scores to detect MCI from Normal, to detect AD from Normal, and to detect MCI or AD from Normal respectively, as well as the summary of those logistic regression models. Those results show that Savonix 11 tests, when used altogether, have very powerful discriminant ability across Normal vs. MCI, Normal vs. AD, Normal vs. MCI or AD, and MCI vs. AD.

Selective use of Savonix subtests

The above results show that each of the Savonix 11 subtests differ in the ability to discriminate between different clinical conditions. For

example, one subtest may be particularly effective for differentiating AD from Normal and MCI, but not for differentiating MCI from Normal.

We proceeded to test the hypothesis that only a subset of the Savonix subtests are sufficient for identifying clinical conditions.

To validate this hypothesis, we compared the AUCs of univariate logistic models of the 11 subtests for Normal vs. MCI, Normal vs. AD, Normal vs. MCI or AD, and MCI vs. AD against the minimum and maximum reference AUCs (MinAUC and MaxAUC).

The AUC for the univariate logistic model with Verbal Interference Part 1 was used as the MinAUC for all the three types of discrimination: MCI vs. AD, Normal vs. MCI, Normal vs. AD.

The MaxAUC was AUC for the univariate logistic regression model for one of the total test scores: the univariate models of total 6 domain score were applied to the AD vs. Normal and MCI vs. AD discrimination, while the model of total 11 subtest score was applied for MCI vs. Normal discrimination. Comparison was performed using DeLong's ROC comparison tests.

Table 8, 9, and 10 shows the AUC comparisons between 11 subtests for MCI vs. AD, Normal vs. MCI, Normal vs. AD discrimination. We selected a few of subtests whose AUCs are significantly different (at $p < 0.01$) from the MinAUC, but not from the MaxAUC for each type of discrimination. For MCI vs. AD discrimination, three subtests, Trail Making Part 1 and 2, and Complex Figure were selected; for Normal vs. MCI discrimination, Trail Making Part 2, Maze Part 2, and Complex Figure were selected; and for Normal vs. AD discrimination, two subtests, Trail Making Part 2 and Complex Figure were selected.

Total scores of those selected subsets of subtests were evaluated for their discriminant abilities as with the total 11 subtests' scores to detect AD from MCI, MCI from Normal, or AD from Normal respectively. As seen in Tables 2, 3, and 4, combinations of subtests can detect comparably to the full set of 11 Savonix subtests.

Subtests	AUC (95% CI)	Comparison with MinAUC (Verbal Interference Part 1's) p value*	Comparison with MaxAUC (AUC for Total 6 Domain score: 0.832) p value*
Immediate Verbal Learning	0.708 (0.564-0.851)	$p < 0.05$	$p < 0.05$
Delayed Verbal Learning	0.647 (0.490-0.805)	$p=0.208$ n.s.	$p < 0.001$
Verbal Interference Part 1	0.490 (0.319-0.661)	n.a.	n.a.
Verbal Interference Part 2	0.765 (0.625-0.906)	$p < 0.05$	$p=0.363$ n.s.
Trail Making Part 1	0.810 (0.688-0.931)	$P < 0.01$	$p=0.723$ n.s.
Trail Making Part 2	0.793 (0.666-0.921)	$P < 0.01$	$P=0.332$ n.s.
Digit Span Forward	0.676 (0.523-0.829)	$p=0.138$ n.s.	$p=0.127$ n.s.
Digit Span Reverse	0.666 (0.528-0.804)	$p=0.087$ n.s.	$p < 0.05$
Maze 1	0.642 (0.482-0.802)	$P=0.241$ n.s.	$P < 0.05$
Maze 2	0.577 (0.410-0.744)	$p=0.526$ n.s.	$P < 0.01$
Complex Figure	0.769 (0.634-0.903)	$P < 0.01$	$P=0.388$ n.s.

Table 8: AUC comparison of Savonix subtests with the lower and higher reference AUCs. For MCI vs. AD discrimination. Tests in bold: Selected tests to detect AD from MCI. AUC: Area under the ROC curve, CI: confidence interval. *: p value of DeLong's ROC comparison tests.

Tests	AUC (95% CI)	Comparison with MinAUC (Verbal Interference Part 1's) p value*	Comparison with MaxAUC (AUC for Total 6 Domain score: 0.800) p value
Immediate Verbal Learning	0.684 (0.571-0.798)	p < 0.05	p<001
Delayed Verbal Learning	0.720 (0.616-0.823)	p<0.01	p<0.05
Verbal Interference Part 1	0.469 (0.352-0.586)		
Verbal Interference Part 2	0.485 (0.377-0.593)	p=0.81 n.s.	p<0.001
Trail Making Part 1	0.724 (0.645-0.804)	P<0.001	p<0.05
Trail Making Part 2	0.742 (0.652-0.831)	P<0.001	p=0.111 n.s.
Digit Span Forward	0.551 (0.439-0.662)	p=0.366 n.s.	p<0.0001
Digit Span Reverse	0.557 (0.454-0.660)	p=0.223 n.s.	p<0.001
Maze 1	0.603 (0.510-0.697)	P=0.081 n.s	P<0.01
Maze 2	0.718 (0.621-0.815)	p<0.001	P=0.190 n.s.
Complex Figure	0.745 (0.665-0.825)	P<0.001	P=0.08n.s.

Table 9: AUC comparison of Savonix subtests with the lower and higher reference AUCs. For Normal vs. MCI discrimination.

Tests in bold: Selected tests to detect MCI from Normal. AUC: Area under the ROC curve. CI: confidence interval. *: p value of DeLong's ROC comparison tests.

Tests	AUC (95% CI)	Comparison with MinAUC (Verbal Interference Part 1's) p value*	Comparison with MaxAUC (AUC for Total 11 test score: 0.948) p value
Immediate Verbal Learning	0.885 (0.828-0.941)	p =0.064 n.s.	p<0.01
Delayed Verbal Learning	0.833 (0.738-0.929)	p<0.05	p<0.01
Verbal Interference Part 1	0.511 (0.367-0.655)		
Verbal Interference Part 2	0.755 (0.642-0.869)	p<0.05	p<0.001
Trail Making Part 1	0.894 (0.851-0.937)	P<0.0001	p<0.05
Trail Making Part 2	0.924 (0.880-0.969)	P<0.0001	P=0.084 n.s.
Digit Span Forward	0.627 (0.492-0.762)	p=0.276 n.s.	p<0.0001
Digit Span Reverse	0.596 (0.502-0.691)	p=0.294 n.s.	p<0.0001
Maze 1	0.720 (0.629-0.810)	p<0.05	p<0.0001
Maze 2	0.791 (0.693-0.887)	p<0.01	p<0.01
Complex Figure	0.907 (0.869-0.945)	P<0.0001	P<0.118 n.s.

Table 10: AUC comparison of Savonix subtests with the lower and higher reference AUCs. For Normal vs. AD discrimination.

Tests in bold: Selected tests to detect AD from Normal. AUC: Area under the ROC curve. CI: confidence interval. *: p value of DeLong's ROC comparison tests.

Conclusion

This study investigates the diagnostic validity of Savonix, a mobile/tablet-based neurocognitive assessment system, in a Chinese population. The investigation comprised three parts: 1) subtest-level discrimination between neuronormal, mild cognitive impairment, and Alzheimer's disease cases; 2) convergent validity between the Savonix assessment total scores and the traditional paper-and-pen screening tests; and 3) ROC analyses for the total scores of the entire Savonix assessment and the selective subtests.

Of the 11 subtests in Savonix assessment tool, seven were measured by the test variables regarding accuracy (c.f., incorrect response rates), three measured by the variables regarding speed (c.f., reaction times), and one measured by variables regarding both the accuracy and the speed.

From the comparison between neuronormal, MCI, and AD groups, all seven subtests for the accuracy and all three for the speed variables

showed significant differences between groups indicating that the tool has fine-grained discriminant abilities even at a subtest level.

Comparisons between the total scores versus the traditional paper and pencil screening tools showed a reasonable level of convergent validity. Relatively weaker correlation between CDRSB and Savonix's assessment could be due to CDRSB's scope of functions: CDRSB includes non-cognitive functions such as those related to community affairs, home, hobbies, and personal cares [32] which are not currently included in the Savonix assessment.

The results of ROC analyses indicate high discriminant ability of the Savonix assessment tool's total scores across different types of discriminations: MCI vs. neuro-normal cases; AD vs. neuro-normal cases; MCI vs. AD cases; and MCI or AD (cognitively impaired) vs. neuro-normal cases. The Max discriminant indices for the physician diagnosis such as AUCs, sensitivities and specificities for discriminations indicate that the Savonix assessment is one of the best diagnostic sensitivities among other automated [33], mobile-based

[34], app-based, or web-based [11] tools available today. As for the screening of AD vs MCI cases, the Savonix assessment also showed better performance than traditional pen-and-paper screening such as ADAS-Cog and CDRSB.

The present study also shows that selective application of the Savonix 11 subtests is possible when screening for different neurocognitive conditions. Two tests in particular, Complex Figure (which measures visuo-spatial episodic memory, general attention, planning, executive function) and Trail Making part 2 (which measures attentional switching), can manage all three screening types-AD vs MCI screening, MCI vs neuro-normal screening, and AD vs neuro-normal screening-with diagnostic ability comparable to the use of all 11 subtests.

This data-driven finding is compatible with studies suggesting that tracking of decline in executive functions is as important as tracking episodic memory when differentiating between normal ageing and neurocognitive conditions such as MCI, and AD [35, 36].

For MCI vs neuro-normal screening, Maze part 2 (which taps into visuo-spatial working memory and executive function), was selected. This is compatible with findings that deficit characteristics in visuo-spatial working memory and executive function may be a key to screen MCI from normal aging [37,38].

For AD vs. Neuro-normal screening, Trail Making part 1, which is relatively simple compared to the other two subtests (Complex Figure and Trail Making part 2), was selected. Trail Making part 1 measures information processing speed and simple attention. This is compatible with findings that show that the TMT is sensitive to amyloid burden tracking processing speed and executive function is important for identifying MCI conditions that progress towards AD [39].

The high screening accuracy of the Savonix assessment, even with selective use of subtests, strongly suggests that the assessment tool, which covers the major cognitive domains, can be adapted and abbreviated for cognitive screening at scale, which is crucial for early detection of MCI or dementia.

There are some limitations in this study. Subject sizes for MCI and AD groups were relatively small and the convergent validity examination was limited to clinical cases only. Subtypes of MCI such as amnesic, non-amnesic, or multiple domain MCI were not examined. Future investigation is needed to address these limitations. Nonetheless, this preliminary study demonstrates the diagnostic and convergent validity of a new mobile/tablet-based neurocognitive assessment, Savonix. Savonix has high potential to enhance early detection of cognitive impairment at scale, by overcoming the traditional barriers to in-person clinics and hospitals for cognitive screening. With accumulating evidence that early identification [40], combined with appropriate intervention for modifiable factors, can alter the progression of cognitive change, the possibility of prevention of MCI and possibly dementia is emerging.

Competing Interests

The authors declare the following competing interests:

KS and SC authors are/were employees of Savonix Inc at the time of writing and retain stockholding in Savonix Inc. QZ has no competing interests.

Author Contributions

KS performed data analysis and prepared early drafts of the MS. QZ oversaw conception, design, data collection and revised manuscript. SC oversaw analysis and manuscript preparation and final revision.

Funding

This research study was supported by Savonix Inc.

References

1. Clark CM, Xie S, Chittams J, Ewbank D, Peskind E, et al. (2003) Cerebrospinal fluid tau and beta-amyloid: how well do these biomarkers reflect autopsy-confirmed dementia diagnoses? *Arch Neurol* 60: 1696-702.
2. Hampel H, Teipel SJ, Fuchsberger T, Andreasen N, Wiltfang J, et al (2004) Value of CSF beta-amyloid1-42 and tau as predictors of Alzheimer's disease in patients with mild cognitive impairment. *Mol Psychiatry* 9: 705-710.
3. Andreasen N, Blennow K (2005) CSF biomarkers for mild cognitive impairment and early Alzheimer's disease. *Clin Neurol Neurosurg* 107: 165-173.
4. Dowling NM, Johnson SC, Gleason CE, Jagust WJ (2015) Alzheimer's disease Neuroimaging Initiative. The mediational effects of FDG hypometabolism on the association between cerebrospinal fluid biomarkers and neurocognitive function. *Neuroimage* 105: 357-368.
5. Bradford A, Kunik ME, Schulz P, Williams SP, Singh H, et al. (2009) Missed and delayed diagnosis of dementia in primary care: prevalence and contributing factors. *Alzheimer Dis Assoc Disord* 23: 306-314.
6. Sabbagh MN, Boada M, Borson S, Chilukuri M, Doraiswamy PM, et al (2020) Rationale for Early Diagnosis of Mild Cognitive Impairment (MCI) Supported by Emerging Digital Technologies. *J Prev Alzheimers Dis* 7: 158-164.
7. Alzheimer's Disease Association (2018) 2018 Alzheimer's disease facts and figures.
8. Cuffaro L, Di Lorenzo F, Bonavita S, Tedeschi G, Leocani L, et al. (2020) Dementia care and COVID-19 pandemic: a necessary digital revolution. *Neurol Sci* 41: 1977-1979.
9. Dong YH, Banozic A (2020) Potential Benefits of Telemedicine for the Assessment of Neurological Change in the Age of COVID. *Int J Digit Hlthc* 1: 101.
10. Germine L, Reinecke K, Chaytor NS (2019) Digital neuropsychology: Challenges and opportunities at the intersection of science and software. *Clin Neuropsychol* 33: 271-286.
11. Charalambous AP, Pye A, Yeung WK, Leroi I, Neil M, et al. (2020) Tools for App- and Web-Based Self-Testing of Cognitive Impairment: Systematic Search and Evaluation. *J Med Internet Res*. 22: e14551.
12. Björngrim S, van den Hurk W, Betancort M, Machado A, Lindau M, et al. (2019) Comparing Traditional and Digitized Cognitive Tests Used in Standard Clinical Evaluation - A Study of the Digital Application Minnemera. *Front Psychol* 10: 2327.
13. Smith A (2017) Record shares of Americans now own smartphones, have home broadband.
14. Gao S, Yang Y, Krogstie J (2015) The adoption of smartphones among older adults in China. *International Conference on Informatics and Semiotics in Organisations*. Springer, Cham.
15. The elderly are the next big growth market for Chinese tech firms (2019) *The Economist*.
16. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR Jr, et al. (2011) The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* 7: 263-269.
17. Jack CR Jr, Albert MS, Knopman DS, McKhann GM, Sperling RA, et al. (2011) Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* 7: 257-262.

18. The American Psychiatric Association (1994) *Diagnostic and Statistical Manual of Mental Disorders*. Washington, DC: The American Psychiatric Association. The *Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition.
19. Petersen RC (2004) Mild cognitive impairment as a diagnostic entity. *J Intern Med* 256: 183-194.
20. Katzman R, Zhang MY, Ouang-Ya-Qu, Wang ZY, Liu WT, et al. (1988) A Chinese version of the Mini-Mental State Examination; impact of illiteracy in a Shanghai dementia survey. *J Clin Epidemiol* 41: 971-978.
21. Wang H, Yu X, Li S, Chen Y, Li H, et al. (2004) The cognitive subscale of Alzheimer's Disease Assessment Scale, Chinese version in staging of Alzheimer disease. *Alzheimer Dis Assoc Disord* 18: 231-235.
22. Lin KN, Liu HC (2003) Clinical dementia rating (CDR), Chinese version. *Acta Neurol Taiwanica* 12: 154-165.
23. Rey A (1964) *L'examen clinique en psychologie*. Paris: Presses Universitaires de France.
24. Stroop JR (1935) Studies of interference in serial verbal reactions. *J Exp Psychol* 18: 643-662.
25. Reitan RM (1958) Validity of the Trail Making Test as an indicator of organic brain damage. *Perceptual and Motor Skills* 8: 271-276.
26. Wechsler D (1981) *WAIS-R manual: Wechsler adult intelligence scale-revised*. Psychological Corporation.
27. Milner B (1965) Visually guided maze learning in man: effects of bilateral hippocampal, bilateral frontal, and unilateral cerebral lesions. *Neuropsychol* 3: 317-338.
28. Rey A (1941) *L'examen psychologique dans les cas d'encéphalopathie traumatique*. *Arch Psychol* 37: 126-139.
29. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, et al. (2020) Package 'pROC'.
30. Revelle W, Revelle MW (2020) Package "psych". CRAN.
31. Wickham H, François R, Henry L, Müller K, R-Studio (2020) Package 'dplyr'.
32. O'Bryant SE, Waring SC, Cullum CM, Hall J, Lacritz L, et al. (2008) Staging dementia using Clinical Dementia Rating Scale Sum of Boxes scores: a Texas Alzheimer's research consortium study. *Arch Neurol* 65: 1091-1095.
33. Aslam RW, Bates V, Dundar Y, Hounsou J, Richardson M, et al. (2018) A systematic review of the diagnostic accuracy of automated tests for cognitive impairment. *Int J Geriatr Psychiatry* 33: 561-575.
34. Koo BM, Vizer LM (2019) Mobile Technology for Cognitive Assessment of Older Adults: A Scoping Review. *Innov Aging* 3: igy038.
35. Baudic S, Barba GD, Thibaudet MC, Smagghe A, Remy P, et al. (2006) Executive function deficits in early Alzheimer's disease and their relations with episodic memory. *Arch Clin Neuropsychol* 21: 15-21.
36. Belleville S, Chertkow H, Gauthier S (2007) Working memory and control of attention in persons with Alzheimer's disease and mild cognitive impairment. *Neuropsychology* 21: 458-469.
37. Kochan NA, Valenzuela M, Slavin MJ, McCraw S, Sachdev PS, et al. (2011) Impact of load-related neural processes on feature binding in visuospatial working memory. *PLoS One* 6: e23960.
38. Gu L, Chen J, Gao L, Shu H, Wang Z, et al. (2019) Deficits of visuospatial working memory and executive function in single- versus multiple-domain amnesic mild cognitive impairment: A combined ERP and sLORETA study. *Clin Neurophysiol* 130: 739-751.
39. Alves L, Cardoso S, Silva D, Mendes T, et al (2020) Neuropsychological profile of amyloid-positive versus amyloid-negative amnesic Mild Cognitive Impairment. *J Neuropsychol* 26: e12218.
40. Rosenberg A, Ngandu T, Rusanen M, Antikainen R, Bäckman L, et al. (2018) Multidomain lifestyle intervention benefits a large elderly population at risk for cognitive decline and dementia regardless of baseline characteristics: The FINGER trial. *Alzheimers Dement* 14: 263-270.