



Published in final edited form as:

*Clin Neuropsychol.* 2021 January ; 35(1): 148–164. doi:10.1080/13854046.2020.1811893.

## Construct Validity, Ecological Validity and Acceptance of Self-administered Online Neuropsychological Assessment in Adults

Naomi S. Chaytor, PhD, ABPP<sup>1</sup>, Celestina Barbosa-Leiker, PhD<sup>2</sup>, Laura T. Germine, PhD<sup>3,4</sup>, Luciana Mascarenhas Fonseca, PhD<sup>1</sup>, Sterling M. McPherson, PhD<sup>1,5,6</sup>, Katherine R. Tuttle, MD<sup>5,6</sup>

<sup>1</sup>Elson S. Floyd College of Medicine, Washington State University, Spokane, WA

<sup>2</sup>College of Nursing, Washington State University, Spokane, WA

<sup>3</sup>Institute for Technology in Psychiatry, McLean Hospital, Belmont, MA

<sup>4</sup>Psychiatry Department, Harvard Medical School, Boston, MA

<sup>5</sup>University of Washington School of Medicine, Seattle, WA

<sup>6</sup>Providence Health Care, Spokane, WA

### Abstract

**Objective:** The goal of this project was to explore the initial psychometric properties (construct and ecological validity) of self-administered online (SAO) neuropsychological assessment (using the [www.testmybrain.org](http://www.testmybrain.org) platform) compared to traditional testing in a clinical sample, as well as to evaluate participant acceptance. SAO assessment has the potential to expand the reach of in-person neuropsychological assessment approaches.

**Method:** Counterbalanced, within-subjects design comparing SAO performance to in-person performance in adults with diabetes with and without Chronic Kidney Disease (CKD). Forty-nine participants completed both assessment modalities (type 1 diabetes N=14, type 2 diabetes N=35; CKD N=18).

**Results:** Associations between SAO and analogous in-person tests were adequate to good ( $r = 0.49–0.66$ ). Association strength between divergent cognitive tests did not differ between SAO versus in-person tests. SAO testing was more strongly associated with age than in-person testing (age  $R^2=0.54$  versus 0.23), while prediction of education, HbA1c, and estimated glomerular

---

Corresponding Author: Naomi Chaytor, WSU Health Sciences Spokane, Elson S. Floyd College of Medicine, HERB 454, 412 E Spokane Falls Blvd, Spokane, WA 99202; Phone 509-368-6775; [naomic@wsu.edu](mailto:naomic@wsu.edu).

Author Contributions: NSC researched data, contributed to the study design, participated in discussion and wrote the manuscript. CBL contributed to the study design, participated in discussion, researched data and reviewed/edited the manuscript. KRT, LTG and SMM contributed to the study design, participated in discussion and reviewed/edited the manuscript. LMF participated in discussion and reviewed/edited the manuscript. All authors reviewed and approved the final manuscript for submission. NSC is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and accuracy of the data analysis.

Duality of Interest: NSC, LTG, LMF, CBL, SMM and KRT have no conflicts of interest related to the present work. NSC has received personal fees from Eli Lilly outside the submitted work. CBL has consulted for Providence Saint John's Health Center and has received research funding from the Bristol-Myers Squibb Foundation outside the submitted work. SMM has received research funding from the Bristol-Myers Squibb Foundation, Ringful Health, LLC, Managed Health Connections, LLC, the Orthopedic Specialty Institute, and consulted for Consistent Care, LLC and the Department of Justice outside the submitted work.

filtration rate (eGFR) did not differ significantly between test modalities (education  $R^2=0.37$  versus 0.30; HbA1c  $R^2=0.20$  versus 0.12; eGFR  $R^2 = 0.41$  versus 0.33). Associations with measures of everyday functioning were also similar (Functional Activities Questionnaire  $R^2=0.08$  versus 0.07; Neuro-QoL  $R^2=0.14$  versus 0.16; Diabetes Self-Management Questionnaire  $R^2=0.19$  versus 0.19).

**Conclusions:** The selected SAO neuropsychological tests had acceptable construct validity (including divergent, convergent, and criterion-related validity), and similar ecological validity to that of traditional testing. These SAO assessments were acceptable to participants and appear appropriate for use in research applications, although further research is needed to better understand the strengths and weaknesses in other clinical populations.

In recent years there has been a growing accumulation of evidence that many systemic diseases, medications, and medical procedures adversely impact cognitive function. However, several limitations of traditional neuropsychological approaches make application for clinical research a challenge. Measures traditionally in use are resource intensive and costly, as they require extensive examiner training and supervision, extended time occupying lab or clinic space, and complex scoring procedures that may introduce human error or bias. These costs are magnified in longitudinal study designs, such as those involved in clinical trials. Increasingly, researchers are exploring remote data collection (e.g., continuous glucose monitoring, GPS, physical activity tracking) and delivery of interventions (e.g., mHealth interventions for people with diabetes (1; 2). Variables that are not possible to measure remotely, such as cognition via in-person traditional neuropsychological testing, are not feasible in these novel trial designs. Therefore, there is a need for developing and validating remote cognitive assessment approaches that could allow for greater efficiency in conducting clinical trials.

Computerized cognitive assessment is often perceived as less difficult and distressing to patients (3), increases efficiency, and reduces associated costs (4). Many existing computerized cognitive tests, however, remain limited by the need for costly technology and skilled test administration in a clinical setting (e.g., NIH Toolbox Cognition Battery). Teleneuropsychology, or the use of video conferencing to administer traditional neuropsychology tests, has been shown to produce equivalent results to in person testing for many verbally administered measures (5; 6), although the evidence for non-verbal processing speed and executive functioning tests is limited (i.e., few studies and small sample sizes) (7; 8). While allowing for remote testing, teleneuropsychology requires skilled test administration and supervision via videoconferencing. Self-administered online (SAO) cognitive assessment addresses these economic and efficiency barriers by allowing patients to complete testing with no human supervision within their own homes using their personal electronic devices.

In addition to economic and logistical concerns, there are growing concerns about the representativeness of samples included in cognitive research, with particular concern about disparities in study enrollment based on dwelling location, socioeconomic status, race and ethnicity, comorbidities, gender, and age (9). Research that occurs in large urban medical centers is often not accessible to non-urban patients who may lack resources or willingness

to participate. In a review of reasons for declining research study participation in NIH intramural research (10), 33% of decliners cited inconvenience as the primary reason (e.g., inability to take time off of work; inability to travel to the Clinical Center; distance to NIH; lack of flexibility in the participant's schedule; and inability to participate during the work week and/or during work hours). This is magnified for individuals who are homebound or have limited mobility. In particular, rural communities are under-represented in research (11), with declination also being related to inconvenience and travel cost (12). Attention to participant preferences when selecting outcome measures may result in better study recruitment, retention and representativeness, assuming psychometric rigor can simultaneously be maintained. Allowing participants to complete assessments in their homes when it is convenient for them to do so (e.g., SAO testing), would reduce barriers to participation (9; 13) and may result in the inclusion of more representative research samples.

Furthermore, evaluation of the psychometric qualities of online neuropsychological measurement tools in clinical samples is particularly important given the recent coronavirus disease (COVID-19) outbreak and the current need for physical distancing. This unprecedented time has necessitated that neuropsychologists be flexible and adapt to the changing health care environment, with increasing expectations to provide services remotely (14; 15). Providers are increasingly looking for evidence to support alternative assessment approaches, including SAO assessment (e.g., see the *Evidence Based Neuropsychological Care During the Covid-19 Pandemic* report by the Inter Organizational Practice Committee, which includes a call for further validation of web-based and computerized testing platforms). In addition, in-person clinical research unrelated to COVID-19 was largely discontinued, making SAO assessment particularly attractive to researchers wanting to continue studies remotely. This public health emergency has rapidly increased the need for data on the strengths and limitations of these approaches.

Although there are some challenges in using self-administered tests, such as the lack of control over the testing environment and the need for a minimal level of familiarity with technology and access to a wifi enabled device, there are many potential advantages. Traditional examiner-administered neuropsychological tests are conducted under artificial testing conditions that have little resemblance to everyday cognitive performance conditions, which is ideal for assessing maximal cognitive performance as needed for diagnostic purposes. There can be a disconnect, however, between performance under controlled conditions and performance in typical day to day environments due to many potential factors such as test-taking anxiety, being tested earlier in the day than preferred, or living in a distracting or noisy household (16; 17). Thus, it is important to include evaluation of ecological validity in SAO assessment validation studies.

Diabetes was chosen as the test population to investigate the use of SAO. Both type 1 and type 2 diabetes are associated with small, but significant cognitive deficits relative to those without diabetes (18–23), cognitive impairments in older adults (24; 25) and an increased risk of cognitive decline (21; 26–28), mild cognitive impairment (29) and dementia (30; 31). Data also indicate a link between diabetes complications, including CKD, and increased risk for cognitive deficits (22; 25; 27; 28). Importantly, the degree of cognitive deficit is associated with glycemic control, as measured by HbA1c (23; 26) and kidney function,

as measured by estimated glomerular filtration rate (eGFR (32–34). Further, remote data collection in diabetes research is increasing with the availability of passive blood glucose monitoring technology (35). Meticulous assessment of cognition in diabetes and CKD research is uncommon, largely due to the logistical challenges associated with existing neuropsychological measures (36).

Despite the promise of the SAO approach, many existing SAO cognitive assessments lack rigorous psychometric validation (e.g., direct comparison to traditional neuropsychological testing, precise characterization of samples), testing in clinical populations with health conditions such as diabetes or CKD, the ability to scale for use in multi-site research, validation for home use, and/or evaluation of patient acceptance (13; 37–40). Further, many existing SAO assessments require the use of a specific device type or operating system (e.g., iPad), and/or consist of a packaged set of tests that may not be appropriate for all research or clinical contexts.

Our primary objective was to establish the construct (convergent, divergent, criterion-related) and ecological validity of a battery of cognitive tests delivered via a research-based SAO cognitive assessment platform ([www.testmybrain.org](http://www.testmybrain.org)) compared to 1) in-person gold standard neuropsychological measures measuring similar cognitive constructs (convergent and divergent validity), 2) demographic and medical characteristics (criterion-related validity) and 3) measures of everyday functioning (ecological validity). Our secondary aim was to demonstrate acceptance of this approach in patients with diabetes (with and without CKD), a population susceptible to disease-associated cognitive deficits. Test My Brain is a not-for-profit program developed at Harvard Medical School and McLean Hospital by Dr. Laura Germine (co-supported by the 501c3 Many Brains Project). Over the past 12 years, over 2.5 million people have completed cognitive tasks on the Test My Brain website. They have developed a wide array of neuroscience based measures for use in research applications, including in cognitive aging and psychiatric populations (41–43). The battery used in the current study has not been directly validated against traditional in-person neuropsychological measures in the same individuals.

## Methods

### Participants.

The study was reviewed and approved by Providence Health Care Institutional Review Board. Patients were recruited from outpatient clinics at Providence Medical Group clinics for Nephrology and Endocrinology (Spokane, WA) between June 2017 and February 2018. Participants were adults with type 1 or type 2 diabetes who 1) were 18 years old, 2) fluent in English and 3) had internet access. Exclusion criteria, as determined via medical record review of diagnostic codes and problem lists by the study coordinator (research nurse) in consultation with a practicing physician (KRT and the study PI, NSC), included: 1) clinical diagnosis of dementia based on medical records, 2) severe impairment in vision, hearing or manual dexterity that would preclude cognitive assessment (based on medical record or observation during consent procedures), 3) physical or mental condition that has known cognitive consequences (e.g., medical record diagnosis of traumatic brain injury of any severity, neurodegenerative disease, stroke, current substance use disorder, developmental

disorder), 4) CKD stage 5 treated by maintenance dialysis or kidney transplant. No additional screening was done to identify any of these exclusion criteria (i.e., cognitive screening for dementia or substance use questionnaire). Study candidates were identified via electronic medical record review and contacted by clinic staff. Interested patients met with a research coordinator at their next clinic visit to learn more about the study, complete the informed consent process, and complete the baseline questionnaires.

### Study Design.

**Baseline measures.**—All baseline variables were assessed via review of medical records and self-report questionnaires. The following data were collected at the baseline clinic visit: Demographic variables including age, gender, race/ethnicity, and attained educational level. Medical record variables included: Medications, diabetes duration, diabetes complications, and laboratory data (HbA1c, serum creatinine for calculation of eGFR). Participants also completed the following self-report questionnaires assessing aspects of everyday functioning at the baseline visit: Instrumental activities of daily living was assessed via the Functional Activities Questionnaire (44), everyday cognitive functioning was assessed via the Neuro-QoL Cognitive Function (45), and cognitively demanding medical management tasks were assessed via the Diabetes Self-Management Questionnaire (46).

**Randomization.**—Following the baseline visit, participants were randomized to one of two counter-balanced assessment order conditions, stratified by diabetes type and CKD status: 1) traditional in-person assessment followed by SAO assessment, or 2) SAO assessment followed by traditional in-person assessment. Permuted block randomization was used with randomly varying block order and length. The blocks varied in length between two and six. After randomization to test order, participants were contacted by a research coordinator to schedule the two assessment sessions (14 days apart).

Prior to any cognitive assessment, participants tested their capillary blood glucose with their home blood glucose meter via finger stick in order to ensure that testing did not occur during hypoglycemia (<70 mg/dL). If hypoglycemia was detected, participants were instructed to eat 15 grams of fast acting carbohydrate and wait until blood glucose was above 70 mg/dL prior to starting the testing.

**SAO testing.**—SAO testing was conducted through [TestMyBrain.org](https://www.testmybrain.org) (TMB), a web-based testing environment that has adapted/developed cognitive assessments for use in research. The SAO test battery included the following measures: TMB Digit Symbol Matching, TMB Digit Span, TMB Letter Number Sequencing, TMB Matrix Reasoning and TMB Vocabulary. These tests were selected because they were directly adapted for SAO administration from well-established clinical neuropsychological tests and have acceptable reliability based on previous on-line data collection on the [testmybrain.org](https://www.testmybrain.org) website (TMB digit symbol matching –  $\rho = 0.93$ ; TMB Letter Number Sequencing –  $\rho = 0.71$  ; TMB Forward Digit Span –  $\rho = 0.73$  ; TMB Backward Digit Span –  $\rho = 0.68$ ; TMB Matrix Reasoning –  $\rho = 0.89$ ; TMB Vocabulary –  $\rho = 0.83$ ). Reliabilities were calculated using a split-half approach (TMB Digit Symbol Matching, Matrix Reasoning, and Vocabulary) or based on correlation between span scores from an interleaved alternate

form with independent stopping rules (TMB Forward Digit Span, Backward Digit Span, and Letter Number Sequencing). The TMB Digit Span Forward test has been shown to produce psychometrically comparable results between anonymous online participants and those tested within a laboratory setting using the same measures (47). Similar patterns of age-related differences using the TMB Digit Symbol Matching and Digit Span compared to WAIS Coding and Digit Span normative data, as well as using the TMB Vocabulary test and the General Social Survey Wordsum test (48). Participants were provided with a study-specific website address and unique ID and password to access the SAO cognitive assessment battery at home. Participants were instructed to start the battery when they had time to complete the entire battery without interruption (~30–45 minutes) and when they were feeling well-rested and alert. They were asked to complete the testing without anyone else in the room and to turn off any devices that could distract or interrupt them (e.g., cell phone, TV, radio). Participants were provided with the phone number and email address for study personnel in case they had any questions. The TMB assessment platform contains standardized testing instructions, including practice trials with feedback on incorrect responses, prior to starting each test. The duration of the SAO test battery was approximately 30 minutes. Raw scores were used in all analyses. Of note, TMB tasks were based on similar tasks from the WAIS-IV (see In-Person Assessment below) but modified in order to optimize computerized administration, while maintaining the core cognitive construct being assessed. For example, TMB Vocabulary uses a multiple-choice format, while WAIS-IV Vocabulary uses an open-ended response format. Presentation modality for all TMB tests was visual and all responses utilized either a keyboard or mouse interface.

**In-Person Assessment.**—The following gold standard “pencil and paper” subtests from the Wechsler Adult Intelligence Scale – 4<sup>th</sup> edition (WAIS-IV) were administered per the published manual according to standardized testing guidelines: WAIS-IV Coding, WAIS-IV Digit Span, WAIS-IV Letter Number Sequencing, WAIS-IV Matrix Reasoning and WAIS-IV Vocabulary. All testing was completed in a private office during normal business hours. The duration of the in-person test battery was approximately 30 minutes. Examiners were trained and certified in test administration by a board certified clinical neuropsychologist (NSC). Raw scores were used in all analyses.

Participant preferences for test modality were assessed via a locally developed questionnaire. After describing the two testing modalities, pre-assessment preferences for each methodology were assessed at the baseline clinic visit by asking the following question: “What is your overall opinion of testing in-person (over the internet)?” using a 7-point Likert scale (e.g., 1 = strongly like in-person assessment, 4 = neutral, 7= strongly dislike in-person assessment). Participants were then asked to indicate which factors were important in determining preference (e.g., convenience, travel time, cost, social interaction). Participant preferences were then reassessed using the same questions after both testing modalities were completed via a RedCap survey.

### **Statistical Analyses.**

Skewness and kurtosis statistics, scatterplots, bivariate correlations, Variance Inflation Factor, and tolerance were used to assess normality, linearity, and multicollinearity,

respectively. Percentages for categorical variables and means with standard deviations for continuous variables (median and interquartile range are reported for non-normally distributed variables) were calculated. Preliminary analyses were performed to assess the statistical assumptions of our tests.

Pearson correlations were used to determine: 1) the association between the SAO tests and the corresponding in-person WAIS-IV tests (convergent validity); 2) if the associations between the SAO fluid tests (TMB Digit Symbol, TMB Digit Span Forward and Backward, TMB Letter Number Sequencing, and TMB Matrix Reasoning) and word knowledge (WAIS-IV Vocabulary) were similar to associations observed between the corresponding WAIS-IV fluid tests and WAIS-IV Vocabulary (divergent validity). Likewise, it was expected that TMB Vocabulary and WAIS-IV Vocabulary would have a similar (and low) correlation with WAIS-IV coding. It is expected that correlations using SAO tests would be similar to correlations using In-Person tests. Differences in the magnitude of two correlations (with one variable in common) were evaluated for statistical significance using Steiger's Z test (49; 50). Regression analyses were used to determine the magnitude of variance accounted for ( $R^2$ ) by the SAO versus in-person tests in everyday functioning measures (ecological validity) and demographic and medical variables (criterion-related validity). Paired-samples *t*-tests were used to compare SAO versus in-person test preference ratings at baseline and again post-testing. Pre-assessment preference ratings for each test methodology were also compared to post-assessment ratings (paired-samples *t*-tests) to determine if ratings changed after exposure to testing. *P* .05 (two-tailed) was used to indicate statistical significance. We elected not to adjust for multiple comparisons give that the goal of these analyses was to retain the null that there are no differences in the associations between SAO tests and in-person tests.

## Results

### Demographic and Clinical Characteristics

Ninety-five participants completed informed consent, medical record review and questionnaires at the time of their initial baseline clinic visit. Forty-two participants did not complete any cognitive testing: 30 (71%) could not be reached, failed to schedule, or cancelled the two assessment sessions, 7 (17%) declined further study participation, and 5 (12%) developed medical conditions that precluded further participation. Four participants completed only one testing session (1 SAO only and 3 In-Person only). Forty-nine participants (52% of those initially consented for the study) completed both cognitive assessment modalities (hereafter referred to as the validation sample). Fourteen had a diagnosis of type 1 diabetes and 35 had type 2 diabetes (Table 1) and 18 had CKD (see Supplemental Table S1 for the breakdown by CKD stage). Compared to those who only completed the baseline assessment ( $N = 42$ ), those in the validation sample ( $N = 49$ ) had better self-reported Diabetes Self-management Questionnaire Scores,  $t(89) = -2.25$ ,  $p = 0.027$  and lower HbA1c,  $t(86) = 2.27$ ,  $p = 0.026$ . There were no differences between the groups in age, education, employment status, diabetes type, diabetes duration, frequency of diabetes complications, eGFR or self-reported instrumental activities of daily living on the Functional Activities Questionnaire.

The two within-subjects testing sessions for the validation sample were an average of 14.8 (SD = 6.0) days apart, with 26 participants randomized to complete the in-person testing first and 23 randomized to complete the SAO testing first. After completing the SAO battery, participants were asked if they had any technical difficulties. Nine of the 42 (21%) participants who completed this item reported having some form of technical problem during SAO testing. The problems included difficulty logging in to the website, difficulty with using computers in general, or using a computer that was slow. These open-ended responses did not reveal any systematic problems and no one mentioned any interruptions during testing. Occasional “technical issues” also occurred infrequently during in-person testing (i.e., examiner forgetting to start the stopwatch or incorrectly reading a digit span sequence). There was no effect of test session order on performance, except for digit span backwards. Mean WAIS-IV Digit Span Backwards performance was better following prior SAO assessment (mean = 8.91, SD = 1.93) compared to when it was completed first (mean = 7.73, SD = 1.82),  $t(47) = -2.21$ ,  $p = 0.03$  (Cohen’s  $d = 0.63$ ). However, TMB Digit Span Backwards performance was poorer in those who had previously completed in-person testing (mean = 4.08, SD = 1.41) compared to when it was completed first (mean = 5.32, SD = 1.64),  $t(46) = -2.82$ ,  $p = .01$  (Cohen’s  $d = 0.81$ ). The SAO testing session occurred outside normal business hours (before 8 AM or after 4:00 PM) in 18 of 49 (37%) participants, with no adverse impact on test performance. Self-monitoring blood glucose values for in-person testing (median = 175 mg/dL; 95% CI [168, 217]) did not differ from those obtained prior to SAO testing (median = 160 mg/dL; 95% CI [149, 181]),  $Z = -1.86$ ,  $p = 0.063$ . There was no association between blood glucose value and any cognitive test performance in either testing session. In those who attended an in-person test session, there was no missing data for any individual cognitive test, while in those who initiated the SAO test session, 3 participants failed to complete all the individual subtests (1 missing Digit Span Backward, and 2 missing Digit Symbol and Letter-Number Switching).

### Construct Validity

Convergent validity, or the degree of association between each SAO test and the corresponding in-person test, was adequate (0.30–0.60) or good (>0.60); ranging from 0.49 for the TMB matrix reasoning task to 0.66 for the TMB digit symbol task (Table 2 and Supplemental Figure S2). Correction for attenuation (51) revealed correlation coefficients ranging from 0.54 to 0.75. The associations between divergent cognitive tests were similar for SAO tests and in-person tests, using Steiger’s Z test (Table 2).

Criterion-related validity was assessed via correlations between each cognitive test modality and variables known to be associated with cognitive performance (age, education, eGFR and HbA1c) (Table 3). The relationship between TMB vocabulary and age was stronger than that between WAIS-IV Vocabulary and age, while other associations were of similar magnitude across testing modalities using Steiger’s Z test. All correlations with education were similar between corresponding SAO and in-person tests. Correlation coefficients between clinical variables (HbA1c and eGFR) and cognitive performance were similar across modalities. The amount of variance accounted for in age was larger when using the SAO battery (Age  $R^2 = 0.54$ ,  $p < 0.001$ ) compared to the in-person battery (Age  $R^2 = 0.23$ ,  $p = 0.074$ ), while the variance accounted for in education, HbA1c and eGFR were similar for the SAO battery



(Education  $R^2 = 0.37$ ,  $p = 0.005$ ; HbA1c  $R^2 = 0.20$ ,  $p = 0.182$ ; eGFR  $R^2 = 0.41$ ,  $p = 0.002$ ) and in-person battery (Education  $R^2 = 0.30$ ,  $p = 0.016$ ; HbA1c  $R^2 = 0.12$ ,  $p = 0.514$ ; eGFR  $R^2 = 0.33$ ,  $p = 0.009$ ).

In terms of ecological validity (association between cognitive tests and measures of everyday functioning), the magnitude of variance accounted for in instrumental activities of daily functioning, self-reported everyday cognitive function and diabetes self-management performance by the battery of SAO tests (FAQ  $R^2 = 0.08$ ,  $p = 0.729$ ; Neuro-QoL  $R^2 = 0.14$ ,  $p = 0.459$ ; DSMQ  $R^2 = 0.19$ ,  $p = 0.188$ ) was similar to that accounted for by the battery of in-person tests (FAQ  $R^2 = 0.07$ ,  $p = 0.762$ ; Neuro-QoL  $R^2 = 0.16$ ,  $p = 0.333$ ; DSMQ  $R^2 = 0.19$ ,  $p = 0.175$ ).

### Participant Preference

The majority of participants had a favorable (“Moderately” or “strongly liked”) perception of both in-person and SAO testing, both before (In-person = 55%; SAO = 51%) and after (In-person = 67%; SAO = 54%) completing both assessment modalities, with convenience being reported as the most common factor supporting preference for SAO, and having someone to assist with testing being reported as the most common reason in support of in-person testing (Supplemental Table S2). No statistically significant differences in continuous participant-reported preference scores between test modality were reported at baseline (SAO mean = 2.67, SD = 1.55; In-Person mean = 2.53, SD = 1.39),  $t(48) = -0.48$ ,  $p = 0.63$ , or post-assessment (SAO mean = 2.86, SD = 1.54; In-Person mean = 2.28, SD = 1.47),  $t(42) = -1.66$ ,  $p = 0.10$ . Further, no change in preference from baseline to post-assessment was detected for SAO (baseline mean = 2.72, SD = 1.47; post mean = 2.86, SD = 1.54),  $t(42) = -0.46$ ,  $p = 0.65$ , or In-Person (baseline mean = 2.60, SD = 1.40; post mean = 2.28, SD = 1.47),  $t(42) = 1.31$ ,  $p = 0.20$ .

### Discussion

The TMB battery of SAO cognitive assessments produced acceptable overall construct validity (convergent, divergent, and criterion-related validity), as well as associations with self-reported everyday functioning (ecological validity) that were comparable to in-person assessments. The SAO tests were correlated with comparable traditional in-person neuropsychological tests, and associations between the SAO tests and other variables were generally similar to those obtained using the in-person tests. Importantly, SAO assessment was acceptable to participants before any exposure to the specific test/platform.

Criterion-related validity was explored by evaluating associations with demographic variables (age and education) as well as associations with laboratory measures that have been linked to cognitive performance in patients with diabetes and CKD (HbA1c and eGFR, respectively). Overall, the relationships between these variables and the SAO cognitive tests were comparable to those with the in-person tests, with one exception. TMB Vocabulary was more associated with age than WAIS-IV Vocabulary. This may be due to differences in the response format of the TMB version (multiple choice response). While this finding needs replication, a prior study using the TMB Vocabulary test showed a stronger association with the General Social Survey Wordsum Task (a vocabulary task that also uses a multiple

choice format) that the WAIS Vocabulary normative sample data (48). Overall, the SAO tests demonstrated a general pattern of associations with demographic and clinical variables of similar magnitude to that seen with in-person testing for most demographic and clinical variables. Participant-reported preference for the two testing modalities was positive and comparable, both before, and after completing them. Convenience of web-based testing was the most cited reason in support of SAO testing, while social interaction was the most common reason given in support of in-person testing. Participants, including older adults, were able to complete this SAO battery as instructed via email independently.

This study has limitations that must be considered when interpreting the results. First, we had a relatively small sample size, high variable to participant ratio, with a low rate of full study completion. However, even with the small sample it was possible to find an overall pattern of results that reveal acceptable construct validity and similar ecological validity to in-person testing, as well as good feasibility of SAO cognitive assessment in adults with diabetes, who are susceptible to subtle cognitive deficits. Of note, while study noncompletion was high, the majority was due to being unable to reach or schedule the testing sessions, while only 5/42 (12%) participants were unable to complete testing due to medical reasons. Nonetheless, a larger sample and a control group of individuals without diabetes would have made the results more robust and allow detection of potential differences between clinical and non-clinical samples. The lack of racial and ethnic diversity in our sample precluded analysis of potential differences in validity and/or preferences among racial and ethnic groups. This is an important area of further investigation. Second, in this study we did not investigate the test-retest reliability of these SAO assessments. Third, the battery of SAO tests together accounted for a similar (small to moderate) amount of variance in relevant self-reported everyday functioning (diabetes self-management, instrumental ADLs and cognitive functioning). It is possible that objective measures of everyday functioning may have produced different results. It has been suggested that the controlled clinic setting used with traditional neuropsychological assessment may be one possible reason for relatively limited correspondence between test scores and real world cognition (17; 52). Despite testing being conducted in the participant's everyday environment, we did not find that the SAO tests had superior ecological validity to that of the in-person tests. It should be noted, however, that in order to maximize convergent validity participants were instructed to minimize interruptions and complete the SAO testing in a private location, which may have resulted in the home environment being more similar to the clinic environment than is typical. Ecological validity of neuropsychological tests is a complex issue (16) and additional research is needed to determine if SAO assessment, particularly when administered in an ecological momentary assessment study design, may result in greater insights into everyday cognitive functioning beyond what is possible with traditional neuropsychological assessment. Next, convergent validity ranged from 0.49–0.66 (0.54–0.75 disattenuated). While adequate for use in research, these associations are modest. It is important to note that the test batteries were administered two weeks apart. The test-retest reliabilities for the WAIS-IV tests range from 0.81–0.94 in the standardization sample. This represents the maximum possible association, with expected lower test-retest reliability in patients with chronic health conditions such as diabetes. Further, these SAO versions of the WAIS-IV tests have key differences in assessment modality (e.g., auditory

vs. visual stimulus delivery, multiple choice versus open-ended oral response), which reduce the expected degree of association further. We were not able to evaluate other factors that might explain discrepant performance between SAO and in person testing (e.g., noise, people around, computer equipment used, not understanding instructions, true changes in cognition). Time of day and blood glucose were not associated with performance. To better understand environmental factors that might impact cognitive performance, further research with more sophisticated sensor-based data collection is needed. In addition, even though construct validity of this SAO battery was considered adequate for research purposes, it is important to note that SAO neuropsychological assessment should be used with extreme caution in clinical settings, as it does not replace the valuable qualitative observations provided by a well-trained professional. The TMB platform also does not confirm test-taker identity nor include measures of performance validity. Ideally, clinical use of SAO assessment should also include observation via simultaneous teleconference.

In conclusion, the SAO cognitive assessment battery tested here appears to be an acceptable research tool. It may be particularly useful when repeated assessment is required or to reach previously underrepresented populations such as those living in remote areas or with transportation and mobility restrictions.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

Funding: This study was supported by an Institute of Translational Health Sciences pilot award via the National Center for Advancing Translational Sciences of the National Institutes of Health under Award Number UL1 TR002319. This work was also partially supported by NIH grant R01 DK 121240-01. The content is solely the responsibility of the authors.

This project was supported by an Institute of Translational Health Sciences pilot award via the National Center for Advancing Translational Sciences of the National Institutes of Health under Award Number UL1 TR002319.

## REFERENCES

1. Cassimatis M, Kavanagh DJ, Hills AP, Smith AC, Scuffham PA, Gericke C, Parham S. The OnTrack Diabetes Web-Based Program for Type 2 Diabetes and Dysphoria Self-Management: A Randomized Controlled Trial Protocol. *JMIR Res Protoc* 2015;4:e97 [PubMed: 26242916]
2. Raiff BR, Barry VB, Ridenour TA, Jitnarin N. Internet-based incentives increase blood glucose testing with a non-adherent, diverse sample of teens with type 1 diabetes mellitus: a randomized controlled Trial. *Transl Behav Med* 2016;6:179–188 [PubMed: 27198235]
3. Collerton J, Collerton D, Arai Y, Barrass K, Eccles M, Jagger C, McKeith I, Saxby BK, Kirkwood T. A comparison of computerized and pencil-and-paper tasks in assessing cognitive function in community-dwelling older people in the Newcastle 85+ Pilot Study. *J Am Geriatr Soc* 2007;55:1630–1635 [PubMed: 17697099]
4. Bauer RM, Iverson GL, Cernich AN, Binder LM, Ruff RM, Naugle RI. Computerized neuropsychological assessment devices: joint position paper of the American Academy of Clinical Neuropsychology and the National Academy of Neuropsychology. *Clin Neuropsychol* 2012;26:177–196 [PubMed: 22394228]
5. Wadsworth HE, Galusha-Glasscock JM, Womack KB, Quiceno M, Weiner MF, Hynan LS, Shore J, Cullum CM. Remote Neuropsychological Assessment in Rural American Indians with and without Cognitive Impairment. *Arch Clin Neuropsychol* 2016;31:420–425 [PubMed: 27246957]

6. Munro Cullum C, Hynan LS, Grosch M, Parikh M, Weiner MF. Teleneuropsychology: evidence for video teleconference-based neuropsychological assessment. *J Int Neuropsychol Soc* 2014;20:1028–1033 [PubMed: 25343269]
7. Breatly TW, Shura RD, Martindale SL, Lazowski RA, Luxton DD, Shenal BV, Rowland JA. Neuropsychological Test Administration by Videoconference: A Systematic Review and Meta-Analysis. *Neuropsychol Rev* 2017;27:174–186 [PubMed: 28623461]
8. Marra D, Hamlet K, Bauer R, Bowers D. Validity of Teleneuropsychology for Older Adults in Response to COVID-19: A Systematic and Critical Review. *The Clinical neuropsychologist* 2020;
9. Coakley M, Fadiran EO, Parrish LJ, Griffith RA, Weiss E, Carter C. Dialogues on Diversifying Clinical Trials: Successful Strategies for Engaging Women and Minorities in Clinical Trials. *J Womens Health (Larchmt)* 2012;21:713–716 [PubMed: 22747427]
10. Brintnall-Karabelas J, Sung S, Cadman ME, Squires C, Whorton K, Pao M. Improving Recruitment in Clinical Trials: Why Eligible Participants Decline. *J Empir Res Hum Res Ethics* 2011;6:69–74 [PubMed: 21460590]
11. Bergeron CD, Foster C, Friedman DB, Tanner A, Kim SH. Clinical trial recruitment in rural South Carolina: a comparison of investigators' perceptions and potential participant eligibility. *Rural Remote Health* 2013;13:2567 [PubMed: 24325179]
12. Kim SH, Tanner A, Friedman DB, Foster C, Bergeron CD. Barriers to clinical trial participation: a comparison of rural and urban communities in South Carolina. *J Community Health* 2014;39:562–571 [PubMed: 24310703]
13. Koo BM, Vizer LM. Mobile Technology for Cognitive Assessment of Older Adults: A Scoping Review. *Innov Aging* 2019;3
14. Thornton J. Covid-19: how coronavirus will change the face of general practice forever. *BMJ* 2020;368:m1279
15. Kaiser UB, Mirmira RG, Stewart PM. Our Response to COVID-19 as Endocrinologists and Diabetologists. *J Clin Endocrinol Metab* 2020;105:dga148
16. Chaytor N, Schmitter-Edgecombe M. The ecological validity of neuropsychological tests: a review of the literature on everyday cognitive skills. *Neuropsychol Rev* 2003;13:181–197 [PubMed: 15000225]
17. Chaytor N, Schmitter-Edgecombe M, Burr R. Improving the ecological validity of executive functioning assessment. *Arch Clin Neuropsychol* 2006;21:217–227 [PubMed: 16554143]
18. Nandipati S, Luo X, Schimming C, Grossman HT, Sano M. Cognition in non-demented diabetic older adults. *Curr Aging Sci* 2012;5:131–135 [PubMed: 22023096]
19. Möttus R, Luciano M, Starr JM, Deary IJ. Diabetes and life-long cognitive ability. *J Psychosom Res* 2013;75:275–278 [PubMed: 23972418]
20. Brands AM, Biessels GJ, de Haan EH, Kappelle LJ, Kessels RP. The effects of type 1 diabetes on cognitive performance: a meta-analysis. *Diabetes Care* 2005;28:726–735 [PubMed: 15735218]
21. Johnston H, McCrimmon R, Petrie J, Astell A. An estimate of lifetime cognitive change and its relationship with diabetes health in older adults with type 1 diabetes: preliminary results. *Behav Neurol* 2010;23:165–167 [PubMed: 21422545]
22. Hardigan T, Ward R, Ergul A. Cerebrovascular complications of diabetes: focus on cognitive dysfunction. *Clin Sci (Lond)* 2016;130:1807–1822 [PubMed: 27634842]
23. Nunley KA, Rosano C, Ryan CM, Jennings JR, Aizenstein HJ, Zgibor JC, Costacou T, Boudreau RM, Miller R, Orchard TJ, Saxton JA. Clinically Relevant Cognitive Impairment in Middle-Aged Adults With Childhood-Onset Type 1 Diabetes. *Diabetes Care* 2015;38:1768–1776 [PubMed: 26153270]
24. Chaytor NS, Barbosa-Leiker C, Ryan CM, Germine LT, Hirsch IB, Weinstock RS. Clinically significant cognitive impairment in older adults with type 1 diabetes. *J Diabetes Complications* 2019;33:91–97 [PubMed: 29728302]
25. Umemura T, Kawamura T, Umegaki H, Kawano N, Mashita S, Sakakibara T, Hotta N, Sobue G. Association of chronic kidney disease and cerebral small vessel disease with cognitive impairment in elderly patients with type 2 diabetes. *Dement Geriatr Cogn Dis Extra* 2013;3:212–222 [PubMed: 23888167]

26. Yaffe K, Falvey C, Hamilton N, Schwartz AV, Simonsick EM, Satterfield S, Cauley JA, Rosano C, Launer LJ, Strotmeyer ES, Harris TB. Diabetes, glucose control, and 9-year cognitive decline among older adults without dementia. *Arch Neurol* 2012;69:1170–1175
27. Tonoli C, Heyman E, Roelands B, Pattyn N, Buyse L, Piacentini MF, Berthoin S, Meeusen R. Type 1 diabetes-associated cognitive decline: A meta-analysis and update of the current literature. *Journal of Diabetes* 2014;6:499–513 [PubMed: 25042689]
28. Ryan CM, Geckle MO, Orchard TJ. Cognitive efficiency declines over time in adults with Type 1 diabetes: effects of micro- and macrovascular complications. *Diabetologia* 2003;46:940–948 [PubMed: 12819900]
29. Roberts RO, Knopman DS, Geda YE, Cha RH, Pankratz VS, Baertlein L, Boeve BF, Tangalos EG, Ivnik RJ, Mielke MM, Petersen RC. Association of diabetes with amnesic and nonamnesic mild cognitive impairment. *Alzheimers Dement* 2014;10:18–26 [PubMed: 23562428]
30. Strachan MW, Reynolds RM, Marioni RE, Price JF. Cognitive function, dementia and type 2 diabetes mellitus in the elderly. *Nat Rev Endocrinol* 2011;7:108–114 [PubMed: 21263438]
31. Smolina K, Wotton CJ, Goldacre MJ. Risk of dementia in patients hospitalised with type 1 and type 2 diabetes in England, 1998–2011: a retrospective national record linkage cohort study. *Diabetologia* 2015;58:942–950 [PubMed: 25673256]
32. Bugnicourt JM, Godefroy O, Chillon JM, Choukroun G, Massy ZA. Cognitive disorders and dementia in CKD: the neglected kidney-brain axis. *J Am Soc Nephrol* 2013;24:353–363 [PubMed: 23291474]
33. Hermann DM, Kribben A, Bruck H. Cognitive impairment in chronic kidney disease: clinical findings, risk factors and consequences for patient care. *J Neural Transm* 2014;121:627–632 [PubMed: 24452609]
34. Yaffe K, Ackerson L, Kurella Tamura M, Le Blanc P, Kusek JW, Sehgal AR, Cohen D, Anderson C, Appel L, Desalvo K, Ojo A, Seliger S, Robinson N, Makos G, Go AS, Investigators CRIC. Chronic kidney disease and cognitive function in older adults: findings from the chronic renal insufficiency cohort cognitive study. *J Am Geriatr Soc* 2010;58:338–345 [PubMed: 20374407]
35. Woldaregay A, Årsand E, Walderhaug S, Albers D, Mamykina L, Botsis T, Hartvigsen G. Data-driven Modeling and Prediction of Blood Glucose Dynamics: Machine Learning Applications in Type 1 Diabetes. *Artificial intelligence in medicine* 2019;98:109–134 [PubMed: 31383477]
36. Scott RB, Farmer E, Smiton A, Tovey C, Clarke M, Carpenter K. Methodology of neuropsychological research in multicentre randomized clinical trials: a model derived from the International Subarachnoid Aneurysm Trial. *Clin Trials* 2004;1:31–39 [PubMed: 16281460]
37. Moore RC, Swendsen J, Depp CA. Applications for self-administered mobile cognitive assessments in clinical research: A systematic review. *Int J Methods Psychiatr Res* 2017;26:e1562.
38. Morrison RL, Pei H, Novak G, Kaufer DI, Welsh-Bohmer KA, Ruhmel S, Narayan VA. A computerized, self-administered test of verbal episodic memory in elderly patients with mild cognitive impairment and healthy participants: A randomized, crossover, validation study. *Alzheimers Dement (Amst)* 2018;10:647–656 [PubMed: 30456291]
39. Wong A, Fong CH, Mok VC, Leung KT, Tong RK. Computerized Cognitive Screen (CoCoSc): A Self-Administered Computerized Test for Screening for Cognitive Impairment in Community Social Centers. *J Alzheimers Dis* 2017;59:1299–1306 [PubMed: 28731437]
40. Hansen TI, Haferstrom EC, Brunner JF, Lehn H, Haberg AK. Initial validation of a web-based self-administered neuropsychological test battery for older adults and seniors. *J Clin Exp Neuropsychol* 2015;37:581–594 [PubMed: 26009791]
41. Fortenbaugh FC, DeGutis J, Germine L, Wilmer JB, Grosso M, Russo K, Esterman M. Sustained Attention Across the Life Span in a Sample of 10,000: Dissociating Ability and Strategy. *Psychol Sci* 2015;26:1497–1510 [PubMed: 26253551]
42. Germine LT, Hooker CI. Face emotion recognition is related to individual differences in psychosis-proneness. *Psychol Med* 2011;41:937–947 [PubMed: 20810004]
43. Germine LT, Duchaine B, Nakayama K. Where cognitive development and aging meet: face learning ability peaks after age 30. *Cognition* 2011;118:201–210 [PubMed: 21130422]
44. Pfeffer RI, Kurosaki TT, Harrah CH, Chance JM, Filos S. Measurement of functional activities in older adults in the community. *J Gerontol* 1982;37:323–329 [PubMed: 7069156]

45. Gershon RC, Lai JS, Bode R, Choi S, Moy C, Bleck T, Miller D, Peterman A, Cella D. Neuro-QOL: quality of life item banks for adults with neurological disorders: item development and calibrations based upon clinical and general population testing. *Quality of Life Research* 2011;21:475–486 [PubMed: 21874314]
46. Schmitt A, Gahr A, Hermanns N, Kulzer B, Huber J, Haak T. The Diabetes Self-Management Questionnaire (DSMQ): development and evaluation of an instrument to assess diabetes self-care activities associated with glycaemic control. *Health Qual Life Outcomes* 2013;11:138 [PubMed: 23937988]
47. Germine L, Nakayama K, Duchaine BC, Chabris CF, Chatterjee G, Wilmer JB. Is the Web as good as the lab? Comparable performance from Web and lab in cognitive/perceptual experiments. *Psychon Bull Rev* 2012;19:847–857 [PubMed: 22829343]
48. Hartshorne JK, Germine LT. When does cognitive functioning peak? The asynchronous rise and fall of different cognitive abilities across the life span. *Psychol Sci* 2015;26:433–443 [PubMed: 25770099]
49. Steiger JH. Testing Pattern Hypotheses On Correlation Matrices: Alternative Statistics And Some Empirical Results. *Multivariate Behav Res* 1980;15:335–352 [PubMed: 26794186]
50. Lee IA, Preacher KJ. Calculation for the test of the difference between two dependent correlations with one variable in common [Computer software]. Available from <http://quantpsy.org> 2013
51. Osborne JW. Effect Sizes and the Disattenuation of Correlation and Regression Coefficients: Lessons from Educational Psychology. *Practical Assessment, Research, and Evaluation* 2002;8:11
52. Germine L, Reinecke K, Chaytor NS. Digital neuropsychology: Challenges and opportunities at the intersection of science and software. *Clin Neuropsychol* 2019;33:271–286 [PubMed: 30614374]
53. Cohen J. *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ, Erlbaum, 1988.

**Table 1.**

Demographic and Clinical Characteristics of the Baseline Only (non-completers) and Validation Samples (completers).

		Baseline (N=42)	Validation (N=49)	p-value
Age (years)	Mean (SD)	59.6 (15.4)	57.7 (13.8)	0.540 <sup>†</sup>
Gender	Women	N (%) 20 (45)	26 (53)	0.605 <sup>‡</sup>
Race/Ethnicity	Non-Hispanic White	N (%) 41 (98)	48 (98)	0.277 <sup>‡</sup>
Education (years)	Mean (SD)	14.0 (3.0)	14.8 (2.5)	0.179 <sup>†</sup>
Diabetes Type	Type 1	N (%) 8 (19)	14 (29)	0.290 <sup>‡</sup>
Complications <sup>*</sup>	CKD	N (%) 13 (31)	18 (37)	0.562 <sup>‡</sup>
	Hypertension	N (%) 30 (71)	36 (73)	0.828 <sup>‡</sup>
	Dyslipidemia	N (%) 30 (71)	35 (71)	0.816 <sup>‡</sup>
	Retinopathy	N (%) 8 (19)	8 (16)	0.734 <sup>‡</sup>
	Neuropathy	N (%) 14 (33)	14 (29)	0.624 <sup>‡</sup>
	CVD	N (%) 16 (38)	11 (22)	0.103 <sup>‡</sup>
HbA1c (%)	Mean (SD) mmol/mol	8.12 (1.84) 65 (20.1)	7.37 (1.25) 57 (13.7)	<b>0.026</b> <sup>†</sup>
eGFR <sup>§</sup> (mL/min/1.73m <sup>2</sup> )	Mean (SD)	68.5 (33.8)	71.9 (32.4)	0.632 <sup>†</sup>
Diabetes Duration (years)	Mean (SD)	16.8 (10.9)	17.9 (13.4)	0.680 <sup>†</sup>
DSMQ (0–10)	Mean (SD)	6.9 (1.5)	7.5 (1.3)	<b>0.027</b> <sup>†</sup>
FAQ	Mean (SD)	1.6 (2.9)	1.0 (2.2)	0.286 <sup>†</sup>

Note:

<sup>\*</sup> Complications were determined by medical record review at the baseline visit.

Significance denotes independent samples t-tests

<sup>(†)</sup> for continuous data and  $\chi^2$  tests

<sup>(‡)</sup> for categorical data comparing participants with only baseline data vs. complete assessment data (validation sample). CKD = Chronic Kidney Disease; CVD = Cardiovascular Disease; DSMQ = Diabetes Self-Management Questionnaire; eGFR = estimated glomerular filtration rate; FAQ = Functional Activities Questionnaire; HbA1c = hemoglobin A1c; NeuroQoL = Neurological Disorders Quality of Life – Cognitive Function.

<sup>§</sup> eGFR was derived from the serum creatinine concentration using CKD-Epidemiological calculation.

**Table 2.**

## Construct Validity: Convergent and Divergent Associations

Convergent Validity	Raw Score		Raw Score	Pearson $r^*$	
	Mean (SD)				
TMB Digit Symbol Matching	33.9(7.6)	WAIS-IV Coding	56.7(15.8)	<b>.66 (.74)</b>	
TMB Digit Span Forward	6.1(1.6)	WAIS-IV Digit Span Forward	10.2(2.4)	<b>.53 (.69)</b>	
TMB Digit Span Backward	4.6(1.6)	WAIS-IV Digit Span Backward	8.3(1.9)	<b>.54 (.72)</b>	
TMB LNS	6.1(1.2)	WAIS-IV LNS	19.5(2.6)	<b>.59 (.75)</b>	
TMB Matrix Reasoning	24.0(5.4)	WAIS-IV Matrix Reasoning	16.5(4.3)	<b>.49 (.54)</b>	
TMB Vocabulary	27.0(2.5)	WAIS-IV Vocabulary	41.0(9.7)	<b>.58 (.65)</b>	
Divergent Validity				Z <sup>†</sup>	p-value
WAIS-IV Vocabulary		TMB Digit Symbol Matching	.12 (.13)	0.578	0.281
		WAIS-IV Coding	.05 (.06)		
WAIS-IV Vocabulary		TMB Digit Span Forward	<b>.41 (.49)</b>	-0.564	0.286
		WAIS-IV Digit Span Forward	<b>.48 (.63)</b>		
WAIS-IV Vocabulary		TMB Digit Span Backward	<b>.40 (.50)</b>	-0.401	0.344
		WAIS-IV Digit Span Backward	<b>.45 (.58)</b>		
WAIS-IV Vocabulary		TMB LNS	.15 (.18)	-1.162	0.123
		WAIS-IV LNS	<b>.30 (.36)</b>		
WAIS-IV Vocabulary		TMB Matrix Reasoning	<b>.42 (.46)</b>	0.667	0.252
		WAIS-IV Matrix Reasoning	<b>.33 (.39)</b>		
WAIS-IV Coding		TMB Vocabulary	-.05 (-.06)	-0.741	0.229
		WAIS-IV Vocabulary	.05 (.06)		

\* Good/Very Good (>0.60), Adequate (.30-.60), Weak (<0.30) associations (53)

TMB= [TestMyBrain.org](https://www.testmybrain.org); LNS= Letter-Number Switching; WAIS-IV= Wechsler Adult Intelligence Scale – 4<sup>th</sup> edition

<sup>†</sup> Steiger's Z for comparing dependent correlation coefficients with one variable in common (49; 50).

Disattenuated correlation coefficients in parentheses (51).

N= 49; Statistically significant correlations in bold (p<0.05).



**Table 3.**

Construct Validity: Associations with demographic and clinical variables (criterion-related validity)

	<b>Age</b>	<b>Education</b>	<b>HbA1c</b>	<b>eGFR</b>
TMB Digit Symbol Matching	<b>-.59</b>	.16	-.17	<b>.54</b>
WAIS-IV Coding	<b>-.40</b>	.11	-.21	<b>.37</b>
TMB Digit Span Forward	-.01	<b>.46</b>	.20	.02
WAIS-IV Digit Span Forward	-.09	<b>.37</b>	.12	.04
TMB Digit Span Backward	-.22	.10	-.18	.04
WAIS-IV Digit Span Backward	-.10	<b>.32</b>	.04	.21
TMB LNS	-.06	.29	.18	.06
WAIS-IV LNS	-.18	.23	.01	<b>.31</b>
TMB Matrix Reasoning	-.21	.24	-.26	.19
WAIS-IV Matrix Reasoning	-.27	<b>.40</b>	-.15	<b>.34</b>
TMB Vocabulary	<b>.52 *</b>	<b>.43</b>	.14	<b>-.39</b>
WAIS-IV Vocabulary	.12*	<b>.47</b>	.17	-.20

TMB= [TestMyBrain.org](https://www.testmybrain.org); WAIS-IV= Wechsler Adult Intelligence Scale – 4<sup>th</sup> edition; LNS= Letter-Number Switching; HbA1c = hemoglobin A1c; eGFR = estimated glomerular filtration rate

N= 49; Note: HbA1c missing 2 and eGFR missing 1

\* Steiger's Z = 3.30, p=.001

Statistically significant correlations appear in bold (p<0.05).